Tecnica Attuariale delle Assicurazioni Danni

Tommaso Lenzi

2019/2020

1 Processi Stocastici per il numero di sinistri

1.1 23/09/2019

Processo Stocastico

Un processo stocastico (o aleatorio) è una famiglia di n.a., che indichiamo con $\{X(t), t \in T\}$ o $\{X_t, t \in T\}$, dove T è l'insieme degli indici del processo e $t \in T$ è il parametro del processo.
Solitamente $T \subset \mathcal{R}$ e $t \in T$ è indice temporale (o intervallo di tempo). Il processo è detto a parametro *continuo* se $t \in \mathbb{R}$ ha cardinalità nel continuo (es. un intervallo), mentre è detto a parametro *discreto* se T è *discreto* (se è finito o numerabile).

Fissato un $t \in T$, il processo stocastico $X(t)$ diventa un numero aleatorio definito in P (partizione dell'evento certo Ω) a valori in $\mathbb{R}$ ($X(t, \cdot) : P \quad \Rightarrow \quad \mathbb{R}$). Se consideriamo quindi $\{X(t, \omega); \omega \in P\}$ si hanno tutte le determinazioni possibili per il processo $X(t)$.
Fissato invece un $\omega \in P$ e considerando tutti gli $t \in T$, si ottiene $\{X(t, \omega), t \in T\}$, che è una *realizzazione* del processo stocastico, o *traiettoria*.
Ad esempio fissato un assicurato, sia $N(t)$ il numero di sinistri in $[0, t]$, ovvero il processo stocastico $\{N(t), t \geq 0\}$. Fissato t otteniamo $\{N(t, \omega); \omega \in P\}$, che corrisponde alle possibili determinazioni di $N(t, \cdot)$, ovvero $\{1, 2, \dots\}$. Fissato invece $\omega \in P$ si ottiene $\{N(t, \omega); t \geq 0\}$, ovvero una traiettoria del processo.

Funzione di Ripartizione

Per fare una valutazione probabilistica di X introduciamo la sua *funzione di ripartizione*:

$$F_X(x) : \mathbb{R} \quad \Rightarrow \quad [0, 1] \text{ tale che } F_X(x) = Pr(X \leq x)$$

Proprietà $F_X(x)$:
∘ $0 \leq F_X(x) \leq 1$, $\forall x \in \mathbb{R}$

∘ F_X è non decrescente

In ipotesi di σ-additività
∘ F_X è continua a destra $\quad \Leftrightarrow \quad \lim\limits_{x \to x_0^+} F_X(x) = F_X(x_0)$
Se consideriamo la famiglia di eventi $(X \leq x + 1/n)_{n \geq 1}$ ho che:
$\lim\limits_{n \to +\infty} Pr(X \leq x_0 + 1/n) = Pr(\bigwedge_{n \geq 1} X \leq x_0 + 1/n) = P(X \leq x_0) = F_X(x_0)$.

∘ $\lim\limits_{x \to -\infty} F_X(x) = 0$
Consideriamo $(X \leq -n)_{n \geq 0}$, allora:
$\lim\limits_{n \to +\infty} Pr(X \leq -n) = Pr(\bigwedge_{n \geq 0}(X \leq -n)) = Pr(\emptyset)$

∘ $\lim\limits_{x \to +\infty} F_X(x) = 1$
La successione $(X \leq n)_{n \geq 0}$ dà luogo ad una famiglia monotona non decrescente di eventi $\quad \Rightarrow$
$\Rightarrow \quad \lim\limits_{n \to +\infty} Pr(X \leq n) = Pr(\bigvee_{n \geq 0}(x \leq n)) = Pr(\Omega) = 1$

Funzione di Ripartizione Congiunta

Consideriamo una coppia di n.a. (X,Y) congiuntamente, introducendo $F_{X,Y}(x,y) = Pr(X \leq x, Y \leq y)$, con $(x,y) \in \mathbb{R}^2$, che gode delle seguenti proprietà:

∘ $0 \leq F_{X,Y}(x,y) \leq 1 \; \forall (x,y)$

∘ fissata $y \in \mathbb{R} \quad \Rightarrow \quad F_{\cdot,Y}$ è monotona non decrescente, continua a destra e $\lim\limits_{x \to -\infty} F_{X,Y}(x,y) = 0$.

∘ $\lim\limits_{(x,y)\to(+\infty,+\infty)} F_{X,Y}(x,y) = Pr(\bigvee\limits_{n\geq 0} (X \leq n, Y \leq n)) = Pr(\Omega,\Omega) = 1$

∘ Fissati $x_1 < x_2$ e $y_1 < y_2$ e $R = [x_1,x_2]\text{x}[y_1,y_2] \subseteq \mathbb{R}^2 \quad \Rightarrow \quad Pr((x,y) \in R) \geq 0$ e se abbiamo assegnato $F_{X,Y}(x,y)$ si ha che:
$F_{X,Y}(R) = F_{X,Y}(x_2,y_2) - F_{X,Y}(x_1,y_2) - F_{X,Y}(x_2,y_1) + F_{X,Y}(x_1,y_1)$

Queste nozioni possono essere estese ad un vettore di n.a. $(X_1,\dots,X_n)$ a cui si assegna $\forall (x_1,\dots,x_n)$
$F_{X_1,\dots,X_n}(x_1,\dots,x_n) = Pr(X_1 \leq x_1,\dots,X_n \leq x_n)$

Valutazione Probabilistica di un Processo Stocastico

É necessario quindi assegnare una famiglia $\mathcal{F}$ di funzioni di ripartizione, per ogni possibile scelta del vettore, quindi $\mathcal{F} = \{F_{X(t_1),\dots,X(t_n)}; n \geq 1; t_1,\dots,t_n \in T\}$. Ovvero bisogna assegnare tutte le funzioni di ripartizione congiunte finite dimensionali del processo $(\forall n, \forall t_1,\dots,t_n)$ e $F_{X(t_1),\dots,X(t_n)}$ deve essere funzione di ripartizione n-dimensionale. Inoltre sia $(j_1,\dots,j_n)$ una permutazione di $(1,\dots,n)$, allora $F_{X(t_1),\dots,X(t_n)}(x_1,\dots,x_n)$ deve essere uguale a $F_{X(t_{j_1}),\dots,X(t_{j_n})}(x_{j_1},\dots,x_{j_n})$.
Se il processo $X(t)$ assume determinazioni in $E \; \forall t$, con E discreto, allora si fissa la famiglia delle funzioni di probabilità finito dimensionali $\{Pr(X(t_1) = i_1,\dots,X(t_n) = i_n), n \geq 1, \forall t_1 < t_2 < \dots < t_n, \forall i_1,\dots,i_n \in E\}$

1.2 26/09/2019

Considerato quindi il processo stocastico $\{X(t), t \in T\}$, sarà assegnata come valutazione probabilistica $\{F_{X(t_1),\dots,X(t_n)}; n \geq 1; t_1,\dots,t_n \in T\}$ nel caso in cui $X(t)$ abbia determinazioni continue. Mentre dovremo assegnare come valutazione probabilistica $\{P(X(t_1) = i_1,\dots,X(t_n) = i_n); n \geq 1; t_1,\dots,t_n \in T\}$ nel caso in cui $X(t)$ abbia determinazioni discrete $\forall t$.
La valutazione probabilistica può essere semplificata se abbiamo delle condizioni particolari, ad esempio nel caso in cui si abbia un processo $\{X(t), t \in T\}$ di n.a. stocasticamente indipendenti, ovvero se e solo se:

$$\forall n \geq 2, \forall t_1,\dots,t_n \in T \text{ si ha che } X(t_1),\dots,X(t_n) \text{ sono stocasticamente indipendenti} \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \forall i = 1,\dots,n \text{ si ha che } X(t_i) \overset{d}{=} X(t_i)|H, \forall H = \bigwedge_{j=1}^{n} (X(t_j) \in A_j), A_j \subset \mathbb{R}$$

Se consideriamo quindi la funzione di ripartizione congiunta, in ipotesi di indipendenza stocastica dei n.a. del processo, abbiamo che:

$$F_{X(t_1),\dots,X(t_n)}(x_1,\dots,x_n) = Pr(X(t_1) \leq x_1,\dots,X(t_n) \leq x_n) = \prod_{i=1}^{n} P(X(t_i) \leq x_i) = \prod_{i=1}^{n} F_{X(t_i)}(x_i)$$

Vediamo quindi che grazie all'ipotesi di indipendenza possiamo possiamo assegnare la legge del processo, con le sole funzioni di ripartizione unidimensionali $F_{X(t)}$, $\forall t \in T$. Se i n.a. del processo sono indipendenti e identicamente distribuiti, allora si ha che per assegnare la legge del processo basta assegnare un'unica funzione di ripartizione $F \sim X(t), \forall t \in T$.
Abbiamo già assegnato una valutazione probabilistica in un processo stocastico, ovvero in un processo sottostante la valutazione di un risarcimento di una polizza o un portafoglio collettivo. Consideriamo $X = \sum_{i=1}^{N} y_i$, con $X = 0$ se $N = 0$, dove N numero di sinistri nel portafoglio e y_i il danno derivante dall'i-esimo sinistro. Abbiamo quindi definito un processo stocastico $\{N, Y_1, Y_2, \dots\}$ e abbiamo introdotto un metodo di valutazione in ipotesi di distribuzione composta. In questo modo per valutare il processo basta assegnare $P(N = n)$, $\forall n \in \mathbb{N}$ e F_y, che è la funzione di ripartizione del danno provocato dal sinistro, nell'ipotesi che il sinistro si verifichi.

Definizioni: Sia $T \subset \mathbb{R}$, T intervallo e $\{X(t), t \in T\}$ allora:

∘ Dati $s, t \in T$ con $s < t$, $X(t) - X(s)$ è detto *incremento* del processo relativamente a $]s, t]$

∘ Il processo è detto ad *incremento stazionario*, se $\forall s, t \in T$ con $s < t$, $X(t) - X(s)$ ha legge che dipende solo dall'ampiezza $t - s$ e non da t, s separatamente.

∘ Il processo è detto ad *incrementi indipendenti* se gli incrementi relativi ad un numero finito di intervalli disgiunti sono stocasticamente indipendenti. Ovvero se e solo se:

$\forall n \geq 2, \forall s_1 < t_1 \leq \dots \leq s_n < t_n \in T$ si ha che $X(t_1) - X(s_1), \dots, X(t_n) - X(s_n)$ sono indipendenti $\Leftrightarrow$

$$\Leftrightarrow \quad \forall i = 1, \dots, n \text{ si ha che } [X(t_i) - X(s_i)] \stackrel{d}{=} [X(t_i) - X(s_i) | H], \forall H = \bigwedge_{j=1}^{n} (X(t_j) - X(s_j) \in A_j) \text{ con } A_j \subset \mathbb{R}$$

Processo di conta

Un processo stocastico $\{N(t), t \geq 0\}$ è detto *processo di conta*, se $N(t)$ rappresenta il numero di manifestazioni di un fissato evento nell'intervallo $[0, t]$.
Proprietà:
∘ $\forall t$, $N(t) \geq 0$
∘ $\forall t$, $N(t) \in \mathbb{N}$
∘ $\forall s, t$ se $s < t$ allora $N(s) \leq N(t)$ e $N(s) - N(t)$ rappresenta il numero di manifestazioni nell'intervallo $]s, t]$.

Processo di Poisson

Un processo di conta $\{N(t), t \geq 0\}$ è detto *processo di Poisson* di intensità λ, con $\lambda > 0$, se sono soddisfatti i seguenti assiomi:

I Il processo è ad incrementi indipendenti, ovvero se $\forall n \geq 2, \forall s_1 < t_1 \leq \dots \leq s_n < t_n \in T$ si ha che $N(t_1) - N(s_1), \dots, X(t_n) - X(s_n)$ sono indipendenti.

II $\forall t \geq 0, \forall \Delta t > 0$ si ha che $Pr(N(t + \Delta t) - N(t) = 1) = \lambda \Delta t + o(\Delta t)$, dove $o(\Delta t)$ è una funzione infinitesima per Δt.

III $\forall t \geq 0, \forall \Delta t > 0$ si ha che $Pr(N(t + \Delta t) - N(t) \geq 2) = o(\Delta t)$

IV $Pr(N(0) = 0) = 1$

Osservazione 1: Se $t \geq 0, \Delta t > 0$ allora:
$Pr(N(t+\Delta t) - N(t) = 0) = 1 - Pr(N(t+\Delta t) - N(t) = 1) - Pr(N(t+\Delta t) - N(t) \geq 2) =$
$= 1 - (\lambda\Delta t + o(\Delta t) + o(\Delta t)) = 1 - \lambda\Delta t + o(\Delta t)$

Osservazione 2: Equivalentemente l'assioma II si può scrivere come limite:

$$\lim_{\Delta t \to 0^+} \frac{Pr(N(t+\Delta t) - N(t) = 1)}{\Delta t} = \lambda$$

Ovvero dato II possiamo vederlo come:

$$\frac{Pr(N(t+\Delta t) - N(t) = 1)}{\Delta t} = \frac{\lambda\Delta t}{\Delta t} + \frac{o(\Delta t)}{\Delta t}$$

E notando che λ è funzione costante con limite per Δt tendente a 0 uguale a λ e che $\frac{o(\Delta t)}{\Delta t}$ ha limite per Δt tendente a 0 uguale a 0 (per definizione di $o(\cdot)$). Allora poiché hanno entrambe limite finito, $\exists$ limite della somma ed è uguale alla somma dei limiti. Inoltre $\exists$ limite anche dell'argomento di sinistra ed è uguale all'argomento a destra.

Osservazione 2*: Supponiamo che valga il limite e consideriamo $\forall \Delta t > 0$:

$$\frac{Pr(N(t+\Delta t) - N(t) = 1)}{\Delta t} - \lambda \triangleq \alpha(\Delta t)$$

E poiché i due addendi tendono per Δt che tende a 0 a λ e $-\lambda$, allora $\alpha(\cdot)$ ha limite 0 per $\Delta t \to 0$.

$$\Rightarrow \quad Pr(N(t+\Delta t) - N(t) = 1) = \lambda\Delta t + \alpha(\Delta t)\Delta t$$

Abbiamo riscritto quindi la probabilità di arrivo di un evento come somma di $\lambda\Delta t$, che è proporzionale a Δt e $\alpha(\Delta t)\Delta t$ che è $o(\Delta t)$, poiché il limite del rapporto è 0 e quindi otteniamo che:

$$Pr(N(t+\Delta t) - N(t) = 1) = \lambda\Delta t + o(\Delta t)$$

Teorema 1
Sia $\{N(t), t \geq 0\}$ processo di Poisson di intensità λ, allora $\forall t$ $N(t) \sim Poi(\lambda t)$, ovvero

$$\forall t \geq 0 \; Pr(N(t) = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$$

Dimostrazione
Fissiamo $n \geq 0$ e poniamo $P_n(t) = Pr(N(t) = n)$, funzione di t con $t \geq 0$, e proviamo che

$$P_n(t) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$$

Suddividiamo la dimostrazione in più punti:
1. Fissiamo $n = 0$, poniamo $P_0(t) = Pr(N(t) = 0)$ e proviamo che $P_0(t) = e^{-\lambda t}$
1.1 Fissato $t \geq 0$ proviamo che $P_0(\cdot)$ è derivabile a destra in t:
Sia $h > 0$ e consideriamo $P_0(t+h) = P(N(t+h) = 0)$. Essendo l'evento $(N(t+h) = 0)$ equivalente

a $(N(t) = 0 \wedge N(t+h) - N(t) = 0)$ ed essendo $(N(t) = 0)$ e $(N(t+h) - N(t) = 0)$ disgiunti, per la definizione di processo di Poisson (1.) sono stocasticamente indipendenti.
Allora $P_0(t+h) = Pr(N(t) = 0)Pr(N(t+h) - N(t) = 0) = P_0(t)(1 - \lambda h + o(h))$, quindi dividendo per h otteniamo:

$$\frac{P_0(t+h) - P_0(t)}{h} = P_0(h)\left(\frac{-\lambda h}{h} + \frac{o(h)}{h}\right)$$

Se eseguiamo il limite per $h \to 0^+$ del secondo membro otteniamo $-\lambda P_0(t)$ e 0, per definizione di $o(h)$. Quindi il limite al secondo membro dell'equazione è $-\lambda P_0(t)$, poiché somma di funzioni limitate. Ma allora dato che $\exists$ il limite al secondo membro, si ha che esiste

$$\lim_{h \to 0^+} \frac{P_0(t+h) - P_0(t)}{h} = -\lambda P_0(t)$$

1.2 Proviamo che $P_0(t)$ è derivabile a sinistra in t:
Consideriamo $P_0(t) = Pr(N(t) = 0) = Pr(N(t-h) = 0 \wedge N(t) - N(t-h) = 0) =$
$= Pr(N(t-h) = 0)Pr(N(t) - N(t-h) = 0) = P_0(t-h)(1 - \lambda h + o(h)) \quad \Rightarrow \quad P_0(t-h) = P_0(t)\frac{1}{1-\lambda h+o(h)}$.
Ma allora:

$$\lim_{h \to 0^+} P_0(t-h) = \lim_{h \to 0^+} P_0(t)\frac{1}{1 - \lambda h + o(h)} = P_0(t)$$

Abbiamo quindi che:

$$\frac{P_0(t) - P_0(t-h)}{h} = P_0(t-h)\left(\frac{-\lambda h}{h} + \frac{o(h)}{h}\right)$$

Facendo quindi il limite al secondo membro per $h \to 0^+$ si ottiene $-\lambda P_0(t)$. Ma allora esiste finito il limite del rapporto incrementale ed è uguale a $-\lambda P_0(t)$.

Quindi per 1.1 e 1.2 si conclude che $\forall t \geq 0$ si ha che P_0 è derivabile in t e

$$P_0'(t) = -\lambda P_0(t) \quad \Leftrightarrow \quad P_0'(t) + \lambda P_0(t) = 0 \quad \Leftrightarrow \quad e^{\lambda t}P_0'(t) + \lambda P_0(t) = 0 \quad \Leftrightarrow \quad \frac{\delta(e^{\lambda t}P_0(t))}{\delta t} = 0 \quad \Leftrightarrow$$
$$\Leftrightarrow \quad e^{\lambda t}P_0(t) = k \quad \Leftrightarrow \quad P_0(t) = ke^{-\lambda t}$$

Per la definizione di processo di Poisson (4.), si ha che $P(N(0) = 0) = 1 \quad \Leftrightarrow \quad P_0(0) = 1 \quad \Rightarrow \quad P_0(0) = k \cdot e^0 = k$, ma allora k deve essere uguale a 1, allora si ha $P_0(t) = e^{-\lambda t}$
2. Fissiamo $n \geq 1$, poniamo $P_n(t) = Pr(N(t) = n)$, una funzione di t, e proviamo che $P_n(t) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$
2.1 Fissiamo $n \geq 1$, $t \geq 0$ e proviamo che $P_n(\cdot)$ è derivabile a destra in t.Fissiamo $h > 0$ e consideriamo la funzione in $t+h$, ovvero $P_n(t+h) = Pr(N(t+h) = n)$. L'evento $(N(t+h) = n) = [(N(t+h) = n) \wedge \Omega]$ e possiamo partizionare l'evento certo Ω con $(N(t+h) - N(t) = 0) \vee (N(t+h) - N(t) = 1) \vee (N(t+h) - N(t) \geq 2)$. Vediamo quindi che:

$$(N(t+h) = n) = [N(t+h) = n \wedge N(t+h) - N(t) = 0] \vee [N(t+h) = n \wedge N(t+h) - N(t) = 1] \vee$$
$$\vee [N(t+h) = n \wedge N(t+h) - N(t) \geq 2] =$$

$$[N(t)=n\wedge N(t+h)-N(t)=0]\vee[N(t)=(n-1)\wedge N(t+h)-N(t)=1]\vee\underbrace{[N(t+h)=n\wedge N(t+h)-N(t)\geq 2]}_{A}$$

Quindi, notando che in ogni quadra si ha un prodotto logico tra eventi relativi ad intervalli disgiunti (fuorché in A) ed essendo ogni parentesi quadra disgiunta, possiamo calcolare la probabilità che ci siano n arrivi tra $[0,t+h]$ come:

$$P_n(t+h)=Pr(N(t)=n)Pr(N(t+h)-N(t)=0)+Pr(N(t)=n-1)Pr(N(t+h)-N(t)=1)+Pr(A)=$$

Per assioma II

$$P_n(t+h)=P_n(t)(1-\lambda h+o(h))+P_{n-1}(t)(\lambda h+o(h))+Pr(A)\quad\Leftrightarrow$$

$$\frac{P_n(t+h)-P_n(t)}{h}=\underbrace{P_n(t)\Big(-\frac{\lambda h}{h}+\frac{o(h)}{h}\Big)+P_{n-1}(t)\Big(\frac{\lambda h}{h}+\frac{o(h)}{h}\Big)}_{B}+\frac{Pr(A)}{h}$$

Poiché non siamo in grado di studiare il limite del primo membro, lo studiamo per il secondo. Otteniamo che il limite per $h\to 0^+$ del secondo membro, per B, è uguale a $-\lambda P_n(t)+\lambda P_{n-1}(t)$, mentre lasciamo integro il limite di $Pr(A)/h$.
Consideriamo adesso il solo evento A. Poiché A è prodotto logico degli eventi $[N(t+h)=n]$ e $[N(t+h)-N(t)\geq 2]$, allora $A\Rightarrow[N(t+h)-N(t)\geq 2]$ e per la monotonia della probabilità si ha che: $Pr(A)\leq Pr(N(t+h)-N(t)=\geq 2)$.
Per il assioma III sappiamo che $Pr(N(t+h)-N(t)=\geq 2)=o(h)$, quindi si ha che:

$$\underbrace{\frac{0}{h}}_{\to 0}\leq\frac{Pr(A)}{h}\leq\underbrace{\frac{o(h)}{h}}_{\to 0}\underset{\text{per teorema confronto}}{\Rightarrow}\lim_{h\to 0^+}\frac{Pr(A)}{h}=0$$

Allora il limite per $h\to 0^+$ del secondo membro è $-\lambda P_n(t)+\lambda P_{n-1}(t)$. Allora il limite del secondo membro dato che esiste ed è finito, è anche il limite del primo.

$$\lim_{h\to 0^+}\frac{P_n(t+h)-P_n(t)}{h}=-\lambda P_n(t)+\lambda P_{n-1}(t)$$

2.2 Fissato $n\geq 1$, $t\geq 0$, dimostriamo che P_n è derivabile in t da sinistra. Il procedimento è l'unione dei punti 1.2 e 2.1.
La conseguenza dei punti 2.1 e 2.2 è che $\forall n\geq 1, t\geq 0$ P_n è derivabile in t e si ha il *Sistema di Kolmogorov*:

$$P_n'(t)=-\lambda P_n(t)+\lambda P_{n-1}(t)\quad\Leftrightarrow$$

$$P_n'(t)+\lambda P_n(t)=\lambda P_{n-1}(t)\quad\Leftrightarrow\quad\underbrace{e^{\lambda t}\Big(P_n'(t)+\lambda P_n(t)\Big)}_{\frac{\delta}{\delta t}(e^{\lambda t}P_n(t))}=e^{\lambda t}\Big(\lambda P_{n-1}(t)\Big)\quad\Leftrightarrow\quad\frac{\delta}{\delta t}\Big(e^{\lambda t}P_n(t)\Big)=e^{\lambda t}\Big(\lambda P_{n-1}(t)\Big)$$

3 Proviamo che la famiglia di funzioni $(P_n(\cdot))_{n\geq 1}$, con $\frac{\delta}{\delta t}(e^{\lambda t}P_n(t))=e^{-\lambda t}(\lambda P_{n-1}(t))$, è tale che

$$P_n(t)=\frac{e^{-\lambda t}(\lambda t)^n}{n!}$$

Lo facciamo per induzione su n. Proviamo prima che la proposizione è vera per $n = 1$, quindi che $P_1(t) = e^{-\lambda t}\lambda t$. Abbiamo:

$$\frac{\delta}{\delta t}\left(e^{\lambda t}P_1(t)\right) = \lambda e^{\lambda t}\underbrace{P_0(t)}_{e^{-\lambda t}} = \lambda \quad \Leftrightarrow \quad e^{\lambda t}P_1(t) = \lambda t + k \quad \Leftrightarrow \quad P_1(t) = e^{-\lambda t}\lambda(t+k), \text{ con } k \in \mathbb{R}$$

Per l'assioma IV si ha che $P_1(0) = Pr(N(0) = 1) = 0$, per l'equazione precedente, k deve essere uguale a 0. Applichiamo ora il processo induttivo, ipotizziamo che la proposizione sia vera per $n-1$ e proviamo che $P_n(t) = e^{-\lambda t}(\lambda t)^n/n!$. Partendo dall'equazione precedente:

$$\frac{\delta}{\delta t}\left(e^{\lambda t}P_n(t)\right) = \lambda e^{\lambda t}\underbrace{P_{n-1}(t)}_{\frac{e^{-\lambda t}(\lambda t)^{n-1}}{(n-1)!}} \quad \Leftrightarrow \quad \frac{\delta}{\delta t}\left(e^{\lambda t}P_n(t)\right) = \frac{\lambda(\lambda t)^{n-1}}{(n-1)!} \quad \Leftrightarrow \quad P_n(t) = \frac{(\lambda t)^n}{n}\cdot\frac{e^{-\lambda t}}{(n-1)!} + ke^{-\lambda t}, \; k \in \mathbb{R}$$

Per assioma IV si ha che $P_n(0) = Pr(N(0) = n) = 0$, quindi per l'equazione precedente si ha che $P_n(0) = (0 + k)$ ed affinché questa sia uguale a 0, k deve essere uguale a 0.
Abbiamo quindi dimostrato che $\forall n \geq 0, t \geq 0$ si ha che $Pr(N(t) = n) = e^{-\lambda t}(\lambda t)^n/n!$, che è la tesi.

1.3 27/09/2019

Teorema 2

Se $\{N(t), t \geq 0\}$ è un processo di Poisson di intensità λ, allora è ad incrementi stazionari. Ovvero che $\forall s, t$; $N(t) - N(s)$ ha legge che dipende da $t - s$ e non da t e s separatamente.

Teorema 3

Se $\{N(t), t \geq 0\}$ è processo di Poisson con intensità λ, sia $\tau > 0$ arbitrario, e poniamo $N^{(\tau)}(t) = N(\tau + t) - N(\tau)$. Il processo $\{N^{(\tau)}(t), t \geq 0\}$ è processo di Poisson di intensità λ.

Dimostrazione Indichiamo con 1 $\{N(t), t \geq 0\}$ e con 2 $\{N^{(\tau)}(t), t \geq 0\}$. Vogliamo dimostrare che per 2 valgono i quattro assiomi.
Osservazione 1: se $s < t$, $N^{(\tau)}(t) - N^{(\tau)}(s) = N(\tau+t) - N(\tau) - N(\tau+s) + N(\tau) = N(\tau+t) - N(\tau+s)$, ovvero ogni incremento di 2 è incremento di 1.
Osservazione 2: presi $s_1 < t_1 \leq s_2 < t_2$ e considerati $N^{(\tau)}(t_1) - N^{(\tau)}(s_1) = N(\tau + t_1) - N(\tau + s_1)$ e $N^{(\tau)}(t_2) - N^{(\tau)}(s_2) = N(\tau + t_2) - N(\tau + s_2)$, notiamo che incrementi relativi ad intervalli disgiunti per 2, sono relativi ad intervalli disgiunti per 1, poiché $\tau + s_1 < \tau + t_1 \leq \tau + s_2 < \tau + t_2$.
I Il processo 2 deve avere intervalli indipendenti, ma se noi fissiamo un numero finito di incrementi per 2, troviamo un numero finito di incrementi relativi ad intervalli disgiunti per 1. Quindi poiché per 1 vale l'indipendenza stocastica, varrà anche per 2.

II Dobbiamo verificare che $\forall t \geq 0, \Delta t > 0$ $Pr(N^{(\tau)}(t + \Delta t) - N^{(\tau)}(t) = 1) = \lambda(\Delta t) + o(\Delta t)$. Poiché $Pr(N^{(\tau)}(t + \Delta t) - N^{(\tau)}(t) = 1) = Pr(N(\tau + t + \Delta t) - N(\tau + t) = 1)$, per II di 1 si ha che $Pr(N(\tau + t + \Delta t) - N(\tau + t) = 1) = \lambda(\Delta t) + o(\Delta t)$, quindi II vale anche per 2.

III $\forall t \geq 0, \Delta t > 0$ dobbiamo dimostrare che $Pr(N^{(\tau)}(t + \Delta t) - N^{(\tau)}(t) \geq 2) = o(\Delta t)$. Poiché

$Pr(N^{(\tau)}(t+\Delta t) - N^{(\tau)}(t) \geq 2) = Pr(N(\tau+t+\Delta t) - N(\tau+t) \geq 2)$, l'assioma II per 1 dice che $Pr(N(\tau+t+\Delta t) - N(\tau+t) \geq 2) = o(\Delta t)$, quindi lo è anche per 2.

IV Dobbiamo dimostrare che $Pr(N^{(\tau)}(0) = 0) = 1$ ed è verificato, poiché $N^{(\tau)}(0) = N(\tau) - N(\tau) = 0$.

Essendo verificati i quattro assiomi, $\{N^{(\tau)}(t), t \geq 0\}$ è un processo di Poisson di intensità λ.

◦ É possibile dimostrare il teorema 2 grazie al teorema 3, infatti fissato τ e considerato $\{N^{(\tau)}(t), t \geq 0\}$, questo per il teorema 3 è un processo di Poisson di intensità λ. Per il teorema 1 si ha che:

$$N^{(\tau)}(t) \sim Poi(\lambda t) \qquad \Rightarrow \qquad N(\tau+t) - N(\tau) \sim Poi(\lambda t)$$

Ovvero l'incremento in un processo di Poisson ha distribuzione che dipende solo dall'ampiezza dell'intervallo. In particolare in intervalli di ampiezza unitaria si ha che $N(t+1) - N(t) \sim Poi(\lambda)$ e che $\lambda = \mathbb{E}(N(t+1) - N(t))$, ovvero λ indica il numero atteso di arrivi in un intervallo unitario.

Dati i quattro assiomi, possibile assegnare la legge del processo?
Dobbiamo quindi verificare se dati i quattro assiomi è assegnata la famiglia delle probabilità; $\forall n \geq 1$, $\forall t_1 < \cdots < t_n, \forall i_1 < \cdots < i_n$, è possibile calcolare $Pr(N(t_1) = i_1, \ldots, N(t_n) = i_n)$.
Si, poiché fissato $n \geq 0$, $t_1 < \cdots < t_n$ e $i_1 < \cdots < i_n$, consideriamo che $Pr(N(t_1) = i_1, \ldots, N(t_n) = i_n) = Pr(N(t_1) = i_1, N(t_2) - N(t_1) = i_2 - i_1, \ldots, N(t_n) - N(t_{n-1}) = i_n - i_{n-1})$. Ma questi sono incrementi relativi ad intervalli disgiunti, quindi.

$$Pr\Big(N(t_1) = i_1, N(t_2) - N(t_1) = i_2 - i_1, \ldots, N(t_n) - N(t_{n-1}) = i_n - i_{n-1}\Big) =$$

$$= Pr\Big(N(t_1) = i_1\Big) \prod_{h=2}^{n} Pr\Big(N(t_h) - N(t_{h-1}) = i_h - i_{h-1}\Big) =$$

$$= \frac{e^{-\lambda t_1}(\lambda t_1)^{i_1}}{i_1!} \prod_{h=2}^{n} \frac{e^{-\lambda(t_h - t_{h-1})}(\lambda(t_h - t_{h-1}))^{(i_h - i_{h-1})}}{(i_h - i_{h-1})!}$$

Quindi tramite i quattro assiomi è possibile calcolare tutte le probabilità congiunte finite dimensionali del processo.

1.4 30/09/2019

Dato $\{N(t), t \geq 0\}$ un processo di conta, viene spontaneo considerare il *Processo dei Tempi di Attesa* $\{T_0, T_1, \ldots\}$ con $T_0 = 0$ e T_n il numero aleatorio del tempo di attesa per l'n-esimo sinistro. Un altro processo che deriva dal processo di Poisson è il *Processo dei Tempi di Inter-arrivo* $\{W_1, W_2, \ldots\}$, dove W_n è il tempo di attesa tra il sinistro $(n-1)$-esimo e il sinistro n-esimo.

Osserviamo: ◦ $T_1 = W_1$ ◦ $\sum_{i=1}^{n} W_n = T_n$ ◦ $T_n - T_{n-1} = W_n$
◦ Se $N(t) = n$ si ha che l'evento $\big(N(t) = n\big) = \big(T_n \leq t\big) \wedge \big(T_{n+1} > t\big)$, ovvero troviamo che $N(t)$ corrisponde al numero di $T_n < t$, ovvero :

$$N(t) = card\{n \geq 1 | T_n \leq t\} = card\{n \geq 1 | w_1 + \cdots + w_n \leq t\}$$

Sia $\{N(t), t \geq 0\}$ un processo di Poisson di intensità λ e ricaviamo la funzione di ripartizione di $W_1 = T_1$. Preso $t \geq 0$ abbiamo che:

$$F_{W_1}(t) = Pr(W_1 \leq t) = Pr(T_1 \leq t) = 1 - Pr(T_1 > t) = \underbrace{1 - Pr(N(t) = 0)}_{\text{poichè } (T_1 > t) \equiv (N(t)=0)} = 1 - e^{-\lambda t}$$

Ripasso

Se $X \sim exp(\rho) \equiv \Gamma(1, \rho)$ si ha che:
- $f_X(x) = \rho e^{-\rho x}$
- $F_X(x) = 1 - e^{-\rho x}$, $x \geq 0$
- $\mathbb{E}(x) = \frac{1}{\rho}$
- $m_x = \frac{\rho}{\rho - t}$, $t < \rho$

Si ha quindi che $W_1 = T_1 \sim exp(\lambda)$. Si prova che se $\{N(t), t \geq 0\}$ è un processo di Poisson di intensità λ, allora il processo degli inter-tempi $\{W_1, W_2, \dots\}$ è un processo di n.a. *i.i.d* e $W_n \sim exp(\lambda)$. Si ha quindi una nuova interpretazione di λ, che oltre ad essere il numero atteso di sinistri per un intervallo unitario, il suo reciproco $(1/\lambda)$ indica il tempo medio di attesa tra un sinistro e il suo successivo. Ad esempio preso $\lambda = 6\%$, si ha $1/\lambda = 17$ anni, che indica il tempo medio tra due sinistri.

Altri metodi per definire un Processo di Poisson

Un processo di conta $\{N(t), t \geq 0\}$ è un processo di Poisson, se:

A valgono gli assiomi I, II, III, IV

B valgono i seguenti assiomi:
1. É ad incrementi indipendenti ($\equiv$ I)
2. $\forall s < t$, $N(t) - N(s) \sim Poi(\lambda(t - s))$
3. $P(N(0) = 0) = 1$ ($\equiv$ IV)

Abbiamo già dimostrato A$\Rightarrow$B, dimostriamo B$\Rightarrow$A, con gli assiomi I e IV già dimostrati.

Dimostrazione

II Consideriamo $Pr(N(t + \Delta t) - N(t) = 1)$, Per ipotesi di B si ha:

$$Pr(N(t + \Delta t) - N(t) = 1) = \lambda \Delta t \cdot e^{-\lambda \Delta t} \quad \Leftrightarrow \quad \frac{Pr(N(t + \Delta t) - N(t) = 1)}{\Delta t} = \frac{\lambda \Delta t \cdot \overbrace{e^{-\lambda \Delta t}}^{\to 1}}{\Delta t}$$

Il limite del secondo membro per $\Delta t \to 0^+$ è uguale a λ e dato che questo esiste ed è finito, deve esistere ed essere uguale anche il limite del primo membro. Abbiamo già dimostrato che se il limite del primo membro è uguale a λ, questo equivale ad avere l'assioma II.

III Consideriamo $Pr(N(t+\Delta t) - N(t) \geq 2) = 1 - Pr(N(t+\Delta t) - N(t) = 0) - Pr(N(t+\Delta t) - N(t) = 1)$ e sappiamo, grazie a 2, che è uguale a $1 - e^{-\lambda \Delta t} - \lambda \Delta t e^{-\lambda \Delta t}$. Dividendo per Δt e facendo il limite per $\Delta t \to 0^+$ otteniamo:

$$\lim_{\Delta t \to 0^+} \frac{Pr(N(t + \Delta t) - N(t) \geq 2)}{\Delta t} = \lim_{\Delta t \to 0^+} \underbrace{\frac{1 - e^{-\lambda \Delta t}}{\Delta t}}_{A} - \underbrace{\lambda e^{-\lambda \Delta t}}_{B}$$

Il limite per $\Delta t \to 0^+$ di B è λ, come abbiamo visto al punto precedente. Anche A ha limite uguale a λ, considerando che:

$$\frac{1-e^{-\lambda\Delta t}}{\Delta t} = \lambda\frac{e^{-\lambda\Delta t}-1}{-\lambda\Delta t} \underset{\Delta t\to 0^+}{\to} \lambda\cdot 1 \text{ dato che } \boxed{\lim_{x\to 0}\frac{e^x-1}{x}=1}$$

Quindi il limite del secondo membro è differenza di limiti finiti ed è uguale a 0, ma allora esiste ed è finito anche il limite per il secondo membro ed è uguale a 0. Il che equivale all'assioma III.

C Il processo degli inter-tempi $\{W_1, W_2, \dots\}$è un processo di numeri aleatori *i.i.d.* con $W_n \sim exp(\lambda)$.

Problemi processo di Poisson

Se consideriamo $\{N(t), t \geq 0\}$ un processo di Poisson, allora abbiamo:

◦ **Incrementi indipendenti:** questa ipotesi è molto forte e non può essere presa in considerazione nel caso in cui si studino problemi come gli incendi boschivi o il numero di malati contagiosi per una determinata malattia. E questo perché si ha una dipendenza fra gli eventi, che comporterebbe l'inadeguatezza all'utilizzo di un processo di Poisson.

◦ **Stazionarietà degli incrementi:** in un processo di Poisson la probabilità del numero di incidenti dipende solo dall'ampiezza dell'intervallo, ed è stazionaria nel tempo, ma guardando all'andamento dell'incidenza della RCA notiamo che è diminuita dal 10% del 2010 al 6% del 2017, andando in contrapposizione con l'ipotesi di stazionarietà degli incrementi. Inoltre potrebbero andare in disaccordo anche problemi con effetti di ciclo, che dipendono da cicli economici o presentano effetti di *stagionalità*.

◦ **Uguaglianza tra valore atteso e varianza degli incrementi:** una delle proprietà del processo di Poisson è che $N(t) - N(s) \sim Poi(\lambda(t-s))$, che implica che $E(N(t) - N(s)) = V(N(t) - N(s))$. Questo va spesso in contrapposizione con i dati, quando c'è effetto di *sovradispersione*.

Per ottenere valutazioni probabilistiche più opportune, nel caso in cui il processo di Poisson non rispecchi la realtà, sono stati creati dei processi con alla base un processo di Poisson, ma che ne conservino il primo l'indipendenza, il secondo la stazionarietà.

Processo di Poisson non omogeneo

La stazionarietà del processo di Poisson è in realtà una conseguenza dell'assioma II, quindi viene modificato questo, affinché il processo non sia obbligatoriamente stazionario.

Definizione: Un processo di conta $\{N(t), t \geq 0\}$ è *processo di Poisson non omogeneo* se oltre agli assiomi I, III, IV, vale il seguente assioma:

II′ $Pr(N(t+\Delta t) - N(t) = 1) = \lambda(t)\Delta t + o(\Delta t)$, $\lambda(\cdot)$ è chiamata *funzione di intensità* ed è funzione di t, positiva ed integrabile per ogni intervallo limitato.

Osservazione: Se $\lambda(t) = \lambda$, $\forall t$ (se è la funzione costante), si ottiene un processo di Poisson semplice, mentre se non è costante si avrà un processo di Poisson non omogeneo.

Teorema

Se vale A′, allora $\forall s < t$, $N(t) - N(s) \sim Poi(\mu(t) - \mu(s))$, dove:

$$\mu(t) = \int_0^t \lambda(u)du \qquad \Rightarrow \qquad \mu(t) - \mu(s) = \int_s^t \lambda(u)du$$

Questo teorema, sottolinea la non stazionarietà degli incrementi, essendo questi una funzione di t ed s, e non solo di $t-s$.
Osservazione: $\forall n \geq 1$, $\forall t, s$ poniamo $P_n(s;t) = Pr(N(s+t) - N(s) = n)$ e la studiamo come funzione in t. Inoltre è possibile dimostrare che è derivabile e che

$$P_n(s;t) \sim \frac{e^{-(\mu(s+t)-\mu(s))}(\mu(s+t)-\mu(s))^n}{n!} \text{ , } t \geq 0$$

Proprietà $\mu(t)$

$\circ$ $\mu(\cdot)$ è definita su $[0, +\infty[$ $\circ$ $\mu(0) = 0$ $\circ$ $\mu(t) > 0, t > 0$
$\circ$ $\mu(\cdot)$ è assolutamente continua, poiché integrale di una funzione integrabile per ogni intervallo finito.
$\circ$ $\mu(\cdot)$ è strettamente crescente, poiché se prendo $t_1 < t_2$ ho che:

$$\underbrace{\int_0^{t_2} \lambda(u)du}_{\mu(t_2)} = \underbrace{\int_0^{t_1} \lambda(u)du}_{\mu(t_1)} + \underbrace{\int_{t_1}^{t_2} \lambda(u)du}_{>0} \quad \Rightarrow \quad \mu(t_2) > \mu(t_1)$$

Se considero $N(t) - N(s)$, il suo valore atteso è $\mathbb{E}(N(t) - N(s)) = \mu(t) - \mu(s)$. Ma allora se considero il valore atteso di $N(t)$, questo è uguale al valore atteso di $N(t) - N(0)$. Quindi avrò $\mathbb{E}(N(t)) = \mathbb{E}(N(t) - N(0)) = \mu(t) - \mu(0) = \mu(t)$. Il significato di $\mu(t)$ è quindi il valore atteso del numero aleatorio $N(t)$. La funzione $\mu(\cdot)$ è anche chiamata funzione del valore atteso.

Analogamente al processo di Poisson, un processo di conta $\{N(t), t \geq 0\}$ è processo di Poisson non omogeneo con funzione del valore atteso $\mu(\cdot)$ se valgono i seguenti assiomi:
1 Il processo è ad intervalli indipendenti
2' $\forall s < t$ $N(t) - N(s) \sim Poi(\mu(t) - \mu(s))$
3 $Pr(N(0) = 0) = 1$

1.5 04/10/2019

Proposizione 1
Sia $\underline{N} = \{N(t), t \geq 0\}$ un processo di Poisson di intensità λ sia $\mu(\cdot)$ una funzione definita in $[0, +\infty[$ tale che $\mu(0) = 0$, che $\mu(t) > 0$ per ogni $t > 0$, che $\mu(\cdot)$ sia strettamente crescente e che $\mu(\cdot)$ sia assolutamente continua. Allora posto $\widetilde{N} = N(\mu(t))$, il processo $\underline{\widetilde{N}} = \{\widetilde{N}(t), t \geq 0\}$ è un processo di Poisson non omogeneo, con funzione del valore atteso $\lambda\mu(\cdot)$.
Dimostrazione
Osservazione 1: verifichiamo l'adeguatezza della definizione, ovvero che $\forall t \geq 0$, $\exists \mu(t)$ e $\mu(t) \geq 0$. E questa è sempre vera poiché $\mu : [0, +\infty[\to [0, +\infty[$, quindi $\forall t \geq 0$ $\exists \mu(t) \geq 0$.
Osservazione 2: fissati $s < t \in [0, +\infty[$ $\widetilde{N}(t) - \widetilde{N}(s) = N(\mu(t)) - N(\mu(s))$. Si ha che $\mu(t) > \mu(s)$ perché $\mu(\cdot)$ è strettamente crescente. Ma allora ogni incremento di $\underline{\widetilde{N}}$ è incremento di $\underline{N}$.
Osservazione 3: fissati $s_1 < t_1 \leq s_2 < t_2 \in [0, +\infty[$ e considerati $\widetilde{N}(t_1) - \widetilde{N}(s_1) = N(\mu(t_1)) - N(\mu(s_1))$ e $\widetilde{N}(t_2) - \widetilde{N}(s_2) = N(\mu(t_2)) - N(\mu(s_2))$. Ma allora se $]s_1, t_1]$ e $]s_2, t_2]$ sono disgiunti, anche $]\mu(s_1), \mu(t_1)]$

e $]\mu(s_2), \mu(t_2)]$ sono disgiunti. Quindi gli incrementi relativi ad intervalli disgiunti del processo $\underline{\widetilde{N}}$ corrispondono a incrementi relativi ad intervalli disgiunti per il processo $\underline{N}$.

Le ipotesi della proposizione sono che $\underline{N}$ si processo di Poisson, ovvero che valga B e dobbiamo dimostrare che $\underline{\widetilde{N}}$ soddisfa B$'$.

1 Preso un numero finito di intervalli disgiunti su $[0, +\infty[$, gli incrementi relativi a $\underline{\widetilde{N}}$ corrispondono ad incrementi relativi ad intervalli disgiunti per $\underline{N}$, che sono indipendenti, quindi lo saranno anche gli incrementi per $\underline{\widetilde{N}}$.

2$'$ Sia $s < t$, si ha che $\widetilde{N}(t) - \widetilde{N}(s) = N(\mu(t)) - N(\mu(s)) \sim Poi(\lambda(\mu(t) - \mu(s)))$, poiché incremento di un processo di Poisson semplice.

3 $Pr(\widetilde{N}(0) = 0) = Pr(N(\mu(0)) = 0) = Pr(N(0) = 0) = 1$ per 3 di B.

Proposizione 2

Sia $\underline{\widetilde{N}} = \{\widetilde{N}(t), t \geq 0\}$ un processo di Poisson non omogeneo, con funzione del valore atteso $\lambda\mu(\cdot)$, dove λ è positivo e $\mu(\cdot) : [0, +\infty[\to [0, +\infty[$ tale che $\mu(0) = 0$, che $\mu(t) \geq 0$, che $\mu(\cdot)$ sia strettamente crescente, che $\mu(\cdot)$ sia assolutamente continua e che $\lim_{t\to+\infty} \mu(t) = +\infty$. Allora posto $N(t) = \widetilde{N}(\mu^{-1}(t))$, il processo $\{N(t), t \geq 0\}$ è un processo di Poisson di intensità λ.

Dimostrazione

Osservazione 1: $\forall t \geq 0$, $\exists \mu^{-1}(t)$ e $\mu^{-1}(t) \geq 0$, sappiamo che $\mu(\cdot)$ è definita in $[0, +\infty[$ ed ha come codominio $[0, +\infty[$. Proviamo che $[0, +\infty[$ è immagine di $\mu(\cdot)$. Chiamiamo il codominio di μ $I = [0, +\infty[$. μ è definita in I, μ è continua, ma allora l'insieme immagine per il teorema di connessione è un intervallo. L'estremo inferiore dell'intervallo corrisponde al minimo che è uguale a 0. L'estremo superiore non esiste poiché per ipotesi il limite per $t \to +\infty$ di $\mu(t)$ è $+\infty$. Allora $\mu(I) = [0, +\infty[$ ed inoltre è strettamente crescente, quindi è suriettiva ed esiste sempre $\mu^{-1}(t) : [0, +\infty[\to [0, +\infty[$ ed il valore appartiene a I ed è maggiore di 0, quindi $\forall t \geq 0$ $\exists \mu^{-1}(t)$ e $\mu^{-1}(t) \geq 0$.

Osservazione 2: sia $s < t$ e $N(t) - N(s) = \widetilde{N}(\mu^{-1}(t)) - \widetilde{N}(\mu^{-1}(s))$. Poiché $\mu(\cdot)$ è strettamente crescente, si ha che $\mu^{-1}(\cdot)$ è strettamente crescente, quindi $\mu^{-1}(s) < \mu^{-1}(t)$.

Osservazione 3: siano $s_1 < t_1 \leq s_2 < t_2$, consideriamo $N(t_1) - N(s_1) = \widetilde{N}(\mu^{-1}(t_1)) - \widetilde{N}(\mu^{-1}(s_1))$ e $N(t_2) - N(s_2) = \widetilde{N}(\mu^{-1}(t_2)) - \widetilde{N}(\mu^{-1}(s_2))$, ma quindi incrementi relativi ad intervalli disgiunti per $\underline{N}$ sono incrementi relativi ad intervalli disgiunti per $\underline{\widetilde{N}}$, dato che $\mu^{-1}(s_1) < \mu^{-1}(t_1) \leq \mu^{-1}(s_2) < \mu^{-1}(t_2)$.

L'ipotesi della proposizione è che $\underline{\widetilde{N}}$ sia processo di Poisson non omogeneo, quindi che valga B$'$, mentre la tesi è che per $\underline{N}$ valgano B.

1 N è ad incrementi indipendenti poiché corrispondono ad incrementi di $\underline{\widetilde{N}}$, come nell'osservazione.

2$'$ sia $s < t$, consideriamo $N(t) - N(s) = \widetilde{N}(\mu^{-1}(t)) - \widetilde{N}(\mu^{-1}(s)) = \widetilde{N}(T) - \widetilde{N}(S)$. Per l'assioma secondo di $\underline{\widetilde{N}}$ questo si distribuisce come una $Poi(\lambda(\mu(T) - \mu(S))) \sim Poi(\lambda(\mu(\mu^{-1}(t)) - \mu(\mu^{-1}(s)))) \sim Poi(\lambda(t - s))$

3 $Pr(N(0) = 0) = Pr(\widetilde{N}(\mu^{-1}(0)) = 0) = Pr(\widetilde{N}(0) = 0) = 1$ per il terzo assioma di $\underline{\widetilde{N}}$.

Grazie alle due proposizioni, è quindi possibile valutare un processo di Poisson non omogeneo attraverso un processo di Poisson semplice, applicando una trasformazione sul tempo. Si dice quindi che si passa a tempo *operativo*.

Processi misture di Poisson

Sia $\{N(t), t \geq 0\}$ un processo di conta e sia $\Lambda > 0$ un parametro aleatorio di rischio, allora avremo un processo *mistura di Poisson* se valgono i seguenti assiomi:
a) $\forall x$ (possibile determinazione di Λ), consideriamo $\{N(t)|\Lambda = x, t \geq 0\}$ un processo di Poisson di intensità x;
b) É assegnata la legge di Λ tramite la funzione di ripartizione F_Λ.

É assegnata la legge per il processo?

Affinché sia assegnata la legge per il processo, deve essere possibile calcolare tutte le probabilità congiunte finite dimensionali, ovvero $\forall n \geq 2$, $\forall t_1 < \cdots < t_n$, $\forall i_1 \leq \cdots \leq i_n$ calcolare la $Pr(N(t_1) = i_1, \ldots, N(t_n) = i_n)$.
○ Supponiamo Λ un n.a. finito con determinazioni $x_1, \ldots, x_n$ e supponiamo siano assegnate le $P_i = Pr(\Lambda = x_i)$, per ogni i. Allora per la disintegrabilità della probabilità si ha che:

$$Pr(N(t_1) = i_1, \ldots, N(t_n) = i_n) = \sum_{i=1}^{n} Pr(\Lambda = x_i) Pr(N(t_1) = i_1, \ldots, N(t_n) = i_n|\Lambda = x_i)$$

Siamo allora in grado di calcolare la legge del processo. Inoltre poiché $P_i \geq 0$ e $\sum P_i = 1$ e poiché il secondo addendo è un processo di Poisson, abbiamo una combinazione lineare convessa di processi di Poisson, ovvero una mistura di processi di Poisson.
○ Sia $\Lambda \sim F_\Lambda$, consideriamo allora:

$$Pr(N(t_1) = i_1, \ldots, N(t_n) = i_n) = \int_0^{+\infty} \underbrace{Pr(N(t_1) = i_1, \ldots, N(t_n) = i_n|\Lambda = x)}_{assegnata} d\, \underbrace{F_\Lambda(x)}_{assegnata}$$

Essendo entrambi i membri noti, anche la legge è nota.

Quindi il processo $\{N(t), t \geq 0\}$ che soddisfa a) e b) è detto mistura di Poisson, con misturante F_Λ. Inoltre se F_Λ ammette densità, si può ricavare l'integrale di Riemann.

Proprietà misture di Poisson

○ É un processo ad incrementi stazionari. Ovvero presi $s < t$ e considerato:

$$Pr(N(t) - N(s) = n) = \int_0^{+\infty} Pr(N(t) - N(s) = n|\Lambda = x) dF_\Lambda(x)$$

Condizionatamente a $\Lambda = x$ si ha un processo di Poisson di intensità x, quindi:

$$Pr(N(t) - N(s) = n) = \int_0^{+\infty} e^{-x(t-s)} \frac{(x(t-s))^n}{n!} dF_\Lambda(x)$$

L'integrale dipende da $t - s$ e on da t o s separatamente. Ed in particolare $N(t+1) - N(t) \stackrel{d}{=} N(1)$.
<u>Esempio:</u> sia $\{N(t), t \geq 0\}$ mistura di Poissoniani con misturante $\Gamma(\alpha, \rho)$. Allora poiché la distribuzione Γ è dotata di densità avremo :

$$Pr(N(t)-N(s)=n)=\int_0^{+\infty} e^{-x(t-s)}\frac{(x(t-s))^n}{n!}\underbrace{\frac{\rho^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\rho x}}_{\text{densità } \Gamma(\alpha,\rho)}dx=$$

$$=\frac{\rho^\alpha(t-s)^n}{\Gamma(\alpha)n!}\int_0^{+\infty}\underbrace{x^{\alpha+n-1}e^{-(\rho+(t-s))x}}_{\text{nucleo } \Gamma(\alpha+n,\rho+(t-s))}dx=\frac{\rho^\alpha(t-s)^n}{\Gamma(\alpha)n!}\cdot\frac{\Gamma(\alpha+n)}{(\rho+(t-s))^{\alpha+n}}=$$

$$\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)n!}\underbrace{\left(\frac{\rho}{\rho+(t-s)}\right)}_{\text{p}}^\alpha\underbrace{\left(\frac{t-s}{\rho+(t-s)}\right)}_{\text{q}}^n$$

Che corrisponde alla densità di una Binomiale Negativa di parametri $(\alpha,\frac{\rho}{\rho+t-s})$. Quindi:

$$N(t+1)-N(t)\sim BN(\alpha,\frac{\rho}{\rho+1})\quad \text{con}\quad \mathbb{E}(N(t+1)-N(t))=\frac{\frac{\alpha}{\rho+1}}{\frac{\rho}{\rho+1}}=\frac{\alpha}{\rho}=\mathbb{E}(\Lambda)\quad \text{e}$$

$$V(N(t+1)-N(t))=\frac{\alpha}{\rho+1}\cdot\left(\frac{\rho+1}{\rho}\right)^2=\alpha\frac{\rho+1}{\rho}>\mathbb{E}(N(t+1)-N(t))$$

In questo modo abbiamo un modello che rispecchi di più la realtà, nel caso in cui ci sia sovradispersione.
∘ Le misture non sono ad incrementi indipendenti.
Esempio: siano $s<t$ e consideriamo:

$$Pr(N(t)-N(s)=n|N(s)=m)=\int_0^{+\infty}\underbrace{Pr(N(t)-N(s)=n|N(s)=m\wedge\Lambda=x)}_{Pr(N(t)-N(s)=n|\Lambda=x)}dF_{\Lambda|N(s)=m}(x)$$

L'argomento dell'integrale non dipende da m, ma la funzione integranda sì, quindi non si ha l'indipendenza tra gli incrementi.

Osservazione: sia $\{N(t),t\geq 0\}$ una mistura di Poissoniani con F_Λ funzione misturante. Allora $\mathbb{E}(N(t)|\Lambda=x)=tx$ ed inoltre:

$$\mathbb{E}(N(t))=\int_0^{+\infty}\underbrace{\mathbb{E}(N(t)|\Lambda=x)}_{xt}dF_\Lambda(x)=t\int_0^{+\infty}xdF_\Lambda(x)=t\mathbb{E}(\Lambda)$$

In particolare $N(t+1)-N(t)\stackrel{d}{=}N(1)$ e il suo valore atteso è $\mathbb{E}(\Lambda)$. Poniamo $E(\Lambda)=\lambda$ e definiamo il numero aleatorio $U=\frac{\Lambda}{\lambda}$ di speranza matematica unitaria e tale che $(U=u)=(\Lambda=\lambda u)$. Allora possiamo esprimere Λ come il prodotto tra suo valore atteso λ e un numero aleatorio U di speranza matematica unitaria.

1.6 07/04/2019

Consideriamo un processo di conta $\{N(t),t\geq 0\}$ e U un parametro aleatorio positivo tale che $\mathbb{E}(U)=1$. Allora dati i seguenti assiomi si avrà una mistura di processi di Poisson:
a $\forall u$ determinazioni di U, $\{N(t)|U=u,t\geq 0\}$ è un processo di Poisson di intensità λu con $\lambda>0$.

(b) É assegnata la legge di U (F_U o nel caso esista f_U).

Esempio: sia $U \sim \Gamma(\alpha, \alpha)$ così che $\mathbb{E}(U) = 1$ e $Var(U) = 1/\alpha$. Valutiamo allora la probabilità di incremento su un intervallo di ampiezza 1 e applichiamo il criterio della disintegrabilità della probabilità.

$$Pr(N(t+1)-N(t) = n) = \int_0^{+\infty} \underbrace{Pr(N(t+1) - N(t) = n|U = u)}_{\sim Poi(\lambda u)} \underbrace{dF_U(u)}_{f_U(u)du} = \int_0^{+\infty} \frac{e^{-\lambda u}(\lambda u)^n}{n!} \frac{\alpha^\alpha}{\Gamma(\alpha)} u^{\alpha-1} e^{-\alpha u} du =$$

$$= \frac{\alpha^\alpha \lambda^n}{\Gamma(\alpha) n!} \int_0^{+\infty} \underbrace{u^{\alpha+n-1} e^{-(\alpha+\lambda)u}}_{\text{nucleo } \Gamma(\alpha+n, \alpha+\lambda)} du = \frac{\alpha^\alpha \lambda^n}{\Gamma(\alpha) n!} \frac{\Gamma(\alpha+n)}{(\alpha+\lambda)^{\alpha+n}} = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha) n!} \left(\frac{\alpha}{\alpha+\lambda}\right)^\alpha \left(\frac{\lambda}{\alpha+\lambda}\right)^n$$

Che è la probabilità che il numero di arrivi sia pari ad n per una Binomiale Negativa. Quindi $N(t-1) - N(t) \sim BN(\alpha, \frac{\alpha}{\alpha+\lambda})$ ed inoltre $\mathbb{E}(N(t+1) - N(t)) = \lambda$ e $Var(N(t+1) - N(t)) = \frac{\lambda(\alpha+\lambda)}{\alpha}$.

Riferimenti Bibliografici:

(1) Ross S. M. *Stochastic Process*, Wiley (1983)
(2) Mikosch T. *Non-life Insurance Mathematics. An introduction to stochastic process*, Springer (2004)

2 Modello classico della teoria collettiva del rischio

Si introduce un modello nel quale viene introdotto un processo stocastico collegato ad un portafoglio assicurativo e del quale viene studiata la probabilità di rovina, su un orizzonte temporale più ampio, rispetto alle possibili azioni dell'assicuratore.

Modello classico

Consideriamo un portafoglio assicurativo di rischi *analoghi*. Sia $R > 0$ il capitale allocato al portafoglio nell'istante iniziale (in 0) ed indichiamo con $P(t)$ i premi puri incassati in $[0, t]$ e con $S(t)$ l'importo pagato dall'assicuratore come risarcimento dei sinistri in $[0, t]$. Sia $R(t)$ il n.a. in t del risultato dell'esercizio, tale che $R(t) = R + P(t) - S(t)$. Si ha quindi il processo stocastico $\{R(t), t \geq 0\}$, che è chiamato *processo del Surplus*. Si ipotizza che l'importo dei premi sia lineari nel tempo, ovvero $P(t) = ct$, con $c > 0$, quindi che siano costanti in intervalli di lunghezza costante, in particolare $P(t+1) - P(t) = c$.

Questo modello semplice non tiene conto di determinati fattori:
◦ Delle spese generali e di gestione del contratto;
◦ Delle fluttuazioni del mercato finanziario;
◦ Del momento in cui vengono risarciti i sinistri, che potrebbe non essere immediato;
◦ Dei dividendi e delle allocazioni successive di capitali;

Definizione Rovina

(1) Si dice *rovina* di un portafoglio se $\exists t \geq 0$ tale che si ha $(R(t) < 0) \Leftrightarrow \bigvee_{t \geq 0}(R(t) < 0) \Leftrightarrow \bigvee_{t \geq 0}(R + ct - S(t) < 0)$

(2) Sia $T = \inf\{t \geq 0 | R(t) < 0\}$, si dice che T è *istante di rovina* del portafoglio se $T \neq +\infty$

Inciso: Dato $A \subset \mathbb{R}$, con $A \neq \emptyset$ e inferiormente limitato, allora $\exists \inf\{A\} \in \mathbb{R}$ ed è il massimo delle limitazioni inferiori. Se $A = \emptyset$ si considera la convenzione tale che $\inf\{\emptyset\} = +\infty$, poiché presi $A \subset B \subset \mathbb{R}$, con $A \neq \emptyset$, $B \neq \emptyset$ ed inferiormente limitati, allora $\inf\{A\} \geq \inf\{B\}$. Ma allora preso $A = \emptyset$, $\emptyset \subset B$ ed avrò $\forall B \subset \mathbb{R}$ $\inf\{\emptyset\} \geq \inf\{B\} \Rightarrow \inf\{\emptyset\} = +\infty$.

Quindi preso il n.a. $T : \mathbb{P} \to \mathbb{R}_+ \cup \{+\infty\}$, ovvero preso un $\omega \in \mathbb{P}$ tale che $T = \inf\{t \geq 0 | R(t) < 0\} < +\infty$, implica che avremo la rovina del portafoglio in T. Mentre invece qualora preso un $\omega \in \mathbb{P}$ tale che $T = \inf\{\emptyset\} = +\infty$, implica la non rovina del portafoglio. Quindi la rovina del portafoglio è l'evento $(T < +\infty) = \bigvee\{\omega \in \mathbb{P} | \exists T \geq 0 \text{ tale che } R(T;\omega) < 0\}$.

Probabilità asintotica di rovina

Definiamo la *probabilità asintotica di rovina* con tempo continuo o discreto come:

$$\Psi(R) = Pr\Big(\bigvee_{t\geq 0}(R(t) < 0)\Big) \quad \text{o} \quad \Psi(R) = Pr\Big(\bigvee_{\substack{t\geq 0\\ t\in\mathbb{N}}}(R(t) < 0)\Big)$$

Nella pratica è però più interessante valutare la rovina del portafoglio in un intervallo di tempo ristretto (10/20 anni), utilizzando un t discreto. Definiamo allora la probabilità di rovina in tempo continuo o discreto, con orizzonte limitato τ e con R capitale allocato in 0 come:

$$\overline{\Psi}(R,\tau) = Pr\Big(\bigvee_{0\leq t\leq \tau}(R(t) < 0)\Big) \quad \text{o} \quad \overline{\Psi}(R,\tau) = Pr\Big(\bigvee_{\substack{0\leq t\leq \tau\\ t\in\mathbb{N}}}(R(t) < 0)\Big)$$

Siano allora $B(R,\tau)$ e $B(R)$ la somma logica degli eventi considerando il tempo discreto, mentre $A(R,\tau)$ e $A(R)$ considerando il tempo continuo. Allora poiché $B(R,\tau) \Rightarrow B(R) \Rightarrow A(R)$, per la monotonia della probabilità si ha che $Pr(B(R,\tau)) \leq Pr(B(R)) \leq Pr(A(R))$. Ed inoltre vale anche che $B(R,\tau) \Rightarrow A(R,\tau) \Rightarrow A(R)$, quindi $Pr(B(R,\tau)) \leq Pr(A(R,\tau)) \leq Pr(A(R))$.
Siamo interessati quindi a $\Psi(R) = Pr(\bigvee_{t\geq 0}(R + ct + S(t) < 0))$, che dipende da R, c e da $\{S(t), t \geq 0\}$.
Sia $S(t) = \sum_{h=1}^{N(t)} Y_h$, dove $N(t)$ è il numero aleatorio di sinistri che colpiscono il portafoglio nell'intervallo $[0,t]$ e dove Y_h è il risarcimento per l'h-esimo sinistro. Quindi $Y_h = 0$ se il sinistro h-esimo non si verifica, mentre $Y_h = y$ è il risarcimento per l'h-esimo sinistro.
Quindi il processo $\{S(t), t \geq 0\}$ dipende sia dal processo $\{N(t), t \geq 0\}$ che dal processo $\{Y_1, Y_2, \dots\}$. Per il processo $\{Y_1, Y_2, \dots\}$ si ha la seguente ipotesi:
∘ $\forall H$ evento logicamente dipendente da $\{N(t), t \geq 0\}$, che implichi il verificarsi dei sinistri di indici $1, \dots, n$, si assume che $Y_1|H, \dots, Y_n|H$ siano indipendenti ed identicamente distribuiti e che $Y_i|H$ con $i \leq n$ ha legge che non dipende da n o H.
Sia F_Y, la legge di $Y_i|H$, la funzione di ripartizione per un sinistro, nell'ipotesi che il sinistro si verifichi. Indichiamo con $\mathbb{E}(Y) = \mu$ il risarcimento atteso per sinistro, nell'ipotesi che il sinistro si verifichi e con $V(Y)$, $\mathbb{E}(Y^k) \triangleq \mu_k$ e $m_Y(\cdot)$, la varianza, il momento k-esimo e la funzione generatrice dei momenti per $Y_i|H$.
Fissato quindi $t > 0$ e $S(t)$, come la sommatoria di $N(t)$ risarcimenti, poiché $\forall n > 0$ $Y_1|(N(t) = n), \dots, Y_n|(N(t) = n)$, questi sono IID con $H = (N(t) = n)$, e poiché $Y_i|(N(t) = n)$ ha legge che non dipende da n e da i, allora trovo che $S(t)$ ha distribuzione composta, e dato che $N(t)$ si distribuisce come una Pois-

Ripasso

$X = \sum_{h=1}^{N} Y_h$ ha distribuzione di tipo composto se:
∘ $\forall n > 0$ $Y_1|N = n, \dots, Y_n|N = n$ sono indipendenti ed identicamente distribuiti;
∘ $Y_i|N = n$ con $i \leq n$ ha legge che non dipende da n.

son, $S(t)$ ha distribuzione Poisson Composta ($S(t) \sim PoiComp(\lambda t, F_Y)$).
Quindi troviamo:
∘ $\mathbb{E}(S(t)) = E(N(t))E(Y) = \lambda t\mu$
∘ $V(S(t)) = \mathbb{E}(N(t))\mathbb{E}(Y^2) = \lambda t\mu_2$
∘ $F_{S(t)}(x) = \sum_{n=0}^{+\infty} Pr(N(t) = n)F_Y^{*(n)}(x)$

2.1 10/10/2019

Dato il processo stocastico $\{S(t), t \geq 0\}$ con le seguenti ipotesi:
a $\{N(t), t \geq 0\}$ è un processo di Poisson di intensità λ
b $Y_1|H, \ldots, Y_n|H$ sono IID $\forall H$ evento logicamente dipendente da $\{N(t), t \geq 0\}$
Abbiamo le seguenti conseguenze:
1 $\forall t > 0$ $S(t) \sim PoiComp(\lambda t, F_Y)$
2 $\{S(t), t \geq 0\}$ è ad incrementi stazionari
3 $\{S(t), t \geq 0\}$ è ad incrementi indipendenti
Dimostrazione 2
Vogliamo dimostrare che $\forall s < t$ la legge di $S(t) - S(s)$ dipende da $t - s$ e non da t o s separatamente.
Fissato $s < t$, si ha che $F_{S(t)-S(s)}(x) = Pr(S(t) - S(s) \leq x)$. Indichiamo con $H(m, n)$ l'evento $(N(s) = m, N(t) - N(s) = n)$ e consideriamo l'insieme $\{H(m, n); m, n \in \mathbb{N}\}$, che è composto da eventi a due a due incompatibili ed è tale che $\bigvee_{m=1}^{+\infty} \bigvee_{n=1}^{+\infty} H(m, n) = \Omega$ (è quindi una partizione di Ω). Allora possiamo applicare il principio di disintegrabilità della probabilità come segue:

$$F_{S(t)-S(s)}(x) = Pr(S(t) - S(s) \leq x) = \sum_{m,n}^{0,+\infty} \underbrace{Pr(H(m, n))}_{\mathrm{I}}\underbrace{Pr(S(t) - S(s) \leq x|H(m, n))}_{\mathrm{II}} \stackrel{\circ}{=}$$

<u>Inciso:</u> Data $\sum_{m,n}^{0,+\infty} a_{m,n}$ si prova che se $a_{m,n}$ è a termini positivi, allora $\sum_{m,n}^{0,+\infty} a_{m,n} = \sum_{n=0}^{+\infty} \left[\sum_{m=0}^{+\infty} a_{m,n} \right]$

Analizziamo ora le due componenti separatamente:
I Poiché $H(m, n)$ è un'intersezione di eventi relativi ad intervalli disgiunti ($[0, s]$ e $]s, t]$), questi sono indipendenti per l'ipotesi che gli incrementi siano un processo di Poisson. Quindi:

$$Pr(H(m, n)) = Pr(N(s) = m, N(t) - N(s) = n) = Pr(N(s) = m)Pr(N(t) - N(s) = n)$$

II Possiamo riscrivere la seconda parte, grazie all'ipotesi che $Y_1|H(m, n), \ldots, Y_{m+n}|H(m, n)$ ∼IID:

$$Pr\left(\sum_{h=1}^{N(t)} Y_h - \sum_{h=1}^{N(s)} Y_h \leq x|N(s) = m, N(t) - N(s) = n \right) = Pr\left(\sum_{h=1}^{n+m} Y_h - \sum_{h=1}^{m} Y_h \leq x|H(m, n) \right) =$$

$$Pr\left(\sum_{m+1}^{m+n} Y_h \leq x|H(m, n) \right) = F_Y^{*(n)}(x)$$

Ma allora riprendendo la sommatoria iniziale abbiamo:

$$\overset{\circ}{=} \sum_{m,n}^{0,+\infty} Pr(N(s)=m)Pr(N(t){-}N(s)=n)F_y^{*(n)}(x) = \sum_{m=0}^{+\infty}\left[Pr(N(s)=m)\underbrace{\left(\sum_{n=0}^{+\infty} Pr(N(t)-N(s)=n)F_Y^{*(n)}\right)}_{\text{costante per } m}\right] \Leftrightarrow$$

$$\Leftrightarrow \left[\sum_{n=0}^{+\infty} Pr(N(t)-N(s)=n)F_Y^{*(n)}\right] \underbrace{\left[\sum_{m=0}^{+\infty} Pr(N(s)=m)\right]}_{Pr(\bigvee_{m=0}^{+\infty}(N(s)=m))=Pr(\Omega)} = \sum_{n=0}^{+\infty} Pr(N(t)-N(s)=n)F_Y^{*(n)}(x)$$

Abbiamo quindi trovato che la legge degli incrementi di $\{S(t), t \geq 0\}$ equivale alla sommatoria alla riga precedente. Quindi dato che $F_Y^{*(n)}$ non dipende nè da t, s o $t-s$ e che $\{N(t), t \geq 0\}$ è un processo di Poisson con incrementi stazionari, ovvero che dipendono solo da $t-s$, otteniamo la tesi. Infatti abbiamo trovato che:

$$F_{S(t)-S(s)}(x) = \sum_{n=0}^{+\infty} \frac{e^{-\lambda(t-s)}(\lambda(t-s))^n}{n!} F_Y^{*(n)} \qquad \sim PoiComp(\lambda(t-s), F_Y)$$

Quindi la funzione di ripartizione degli incrementi di $S(t)$ dipende solo da $t-s$ ed inoltre $\{S(t)-S(s), t \geq 0\}$ è un processo di Poisson Composto di parametri $\lambda(t-s)$ e F_Y.

Dimostrazione 3

Fissati $n \geq 2$, $s_1 < t_1 \leq \cdots \leq s_n < t_n$, proviamo che $S(t_1)-S(s_1), \ldots, S(t_n)-S(s_n)$ sono stocasticamente indipendenti, attraverso la proprietà per cui se sono stocasticamente indipendenti allora:

$$F_{S(t_1)-S(s_1),\ldots,S(t_n)-S(s)_n}(x) = \prod_{i=1}^{n} F_{S(t_i)-S(s_i)}(x)$$

Consideriamo il membro di sinistra e suddividiamo l'intervallo $[0, t_n]$ nei sotto-intervalli di interesse $]s_1, t_1]$, $\ldots$, $]s_n, t_n]$ e nei sotto-intervalli complementari $[0, s_1]$, $\ldots$, $]t_{n-1}, s_n]$. Indichiamo con n_i il numero di eventi registrati nell'i-esimo intervallo di interesse e con m_i il numero di eventi nell'intervallo precedente. Consideriamo l'evento $H(m_1, n_1, \ldots, m_n, n_n) = (N(s_1) = m_1, N(t_1) - N(s_1) = n_1, \ldots, N(s_n) - N(t_{n-1}) = m_n, N(t_n){-}N(s_n) = n_n)$. Se prendiamo in esame l'insieme degli eventi $\{H(m_1, n_1, \ldots, m_n, n_n); m_i, n_i \in \mathbb{N}\}$, questo è una partizione di Ω. Ma allora per la disintegrabilità della probabilità troviamo:

$$F_{S(t_1)-S(s_1),\ldots,S(t_n)-S(s)_n}(x) = Pr\left(\bigwedge_{i=1}^{n}(S(t_i)-S(s_i) \leq x)\right) =$$

$$= \sum_{m_1,n_1,\ldots,m_n,n_n}^{0,+\infty} \underbrace{Pr\Big(H(m_1,n_1,\ldots,m_n,n_n)\Big)}_{1} \underbrace{Pr\left(\bigwedge_{i=1}^{n}(S(t_i)-S(s_i) \leq x_i) | H(m_1,n_1,\ldots,m_n,n_n)\right)}_{2} \overset{\circ\circ}{=}$$

Tenendo conto che sono tutti intervalli disgiunti, per l'ipotesi a possiamo fare le seguenti considerazioni:
1 Dato che sono incrementi di intervalli disgiunti si ha che :

$$Pr\Big(H(m_1,n_1,\ldots,m_n,n_n)\Big) = Pr(N(s_1)=m_1)Pr(N(t_1)-N(s_1)=n_1) \cdot \ldots \cdot Pr(N(t_n)-N(s_n)=n_n)$$

2 Considerato l'i-esimo n.a. del secondo membro del prodotto e posto $m_{(i)} = m_1 + n_1 + \cdots + m_i$ si ha:

$$S(t_i){-}S(s_i)|H(m_1, n_1, \ldots, m_n, n_n) = \sum_{h=1}^{N(t_i)} Y_h - \sum_{h=1}^{N(s_i)} Y_h|H(m_1, n_1, \ldots, m_n, n_n) = \sum_{h=m_{(i)}+1}^{m_{(i)}+n_i} Y_h|H(m_1, n_1, \ldots, m_n, n_n)$$

Quindi trovo la somma di un sotto-insieme dei n.a. $Y_1|H, \ldots, Y_{m_{(n)}+n_n}|H$ che sono IID e $\sim F_Y$. Allora applicando ad un sotto-insieme di n.a. indipendenti una funzione, se i sotto-insiemi sono disgiunti, otterrò n.a. relativi ad eventi disgiunti ed indipendenti. Possiamo allora scrivere:

$$Pr(\bigwedge_{i=1}^{n}(S(t_i) - S(s_i) \leq x_i)|H) = \prod_{i=1}^{n} Pr(S(t_i) - S(s_i) \leq x_i|H) =$$

$$\prod_{i=1}^{n} Pr(Y_{m_{(i)}+1}|H + \cdots + Y_{m_{(i)}+n_i}|H \leq x_i) = \prod_{i=1}^{n} F_Y^{*(n_i)}(x_i)$$

Tornando quindi all'espressione della funzione di ripartizione degli incrementi di $\{S(t), t \geq 0\}$:

$$\overset{\text{oo}}{=} \sum_{m_1,n_1,\ldots,m_n,n_n}^{0,+\infty} \Big[Pr(N(s_1) = m_1)Pr(N(t_1){-}N(s_1) = n_1)\cdot\ldots\cdot Pr(N(t_n){-}N(s_n) = n_n)\Big] \prod_{i=1}^{n} F_Y^{*(n_i)}(x_i) =$$

$$\sum_{m_1,n_1,\ldots,m_n,n_n}^{0,+\infty} Pr(N(s_1) = m_1)Pr(N(t_1){-}N(s_1) = n_1)F_Y^{*(n_1)}(x_1)\cdot\ldots\cdot Pr(N(t_n){-}N(s_n) = n_n)F_Y^{*(n_n)}(x_n) =$$

$$= \underbrace{\sum_{m_1=0}^{+\infty} Pr(N(s_1) = m_1)}_{Pr(\Omega)=1} \sum_{n_1=0}^{+\infty} Pr(N(t_1) - N(s_1) = n_1)F_Y^{*(n_1)}(x_1) \cdot \ldots \cdot \underbrace{\sum_{m_n=0}^{+\infty} Pr(N(s_n) - N(t_{n-1}) = m_n)\cdot}_{Pr(\Omega)=1}$$

$$\cdot \sum_{n_n=0}^{+\infty} Pr(N(t_n){-}N(s_n) = n_n)F_Y^{*(n_n)}(x_n) = \prod_{i=1}^{n} \sum_{n_i=0}^{+\infty} Pr(N(t_i){-}N(s_i) = n_i)F_Y^{*(n_i)}(x_i) = \prod_{i=1}^{n} F_{S(t_i)-S(s_i)}(x_i)$$

Abbiamo quindi trovato che la funzione di ripartizione congiunta è uguale al prodotto delle singole funzioni di ripartizione, quindi i n.a. $S(t_1) - S(s_1), \ldots, S(t_n) - S(s_n)$ sono stocasticamente indipendenti.

Definizione: un processo $\{S(t), t \geq 0\}$ che soddisfa le proprietà 1, 2, 3 è un processo di Poisson Composto con parametri λ e F_Y.

Osservazione 1

Se consideriamo $\{N(t), t \geq 0\}$ un proceso di Poisson di intensità λ e $\{Y_1', Y_2' \ldots\}$ IID, con FdR F_Y e stocasticamente indipendenti da $\{N(t), t \geq 0\}$ e posto $S'(t) = \sum_{h=1}^{N(t)} Y_h'$, allora abbiamo che $\{S'(t), t \geq 0\}$ è un processo di Poisson Composto con parametri λ e F_Y ed inoltre:

$$\{S'(t), t \geq 0\} \overset{d}{=} \{S(t), t \geq 0\}$$

Osservazione 2

Sia $S(t) - S(s) \sim PoiComp(\lambda(t - s), F_Y)$ e $\mathbb{E}(Y^k) = \mu_k$ allora:

- $\mathbb{E}(S(t) - S(s)) = \lambda(t - s)\mu_1$ ed in particolare $\mathbb{E}(S(t + 1) - S(t)) = \lambda\mu_1$
- $V(S(t) - S(s)) = \lambda(t - s)\mu_2$ ed in particolare $V(S(t + 1) - S(t)) = \lambda\mu_2$

Osservazione 3

Ripensando all'ipotesi che i premi incassati a fronte di $S(t)$ nell'intervallo $[0,t]$ siano proporzionali all'ampiezza per un coefficiente $P(t) = ct$. Se l'assicuratore applicasse il principio di equità avremmo $P(t) = \mathbb{E}(S(t)) = (\lambda\mu_1)t$, ovvero un coefficiente $(\lambda\mu_1)$ per l'ampiezza dell'intervallo. Se invece applicasse il principio della speranza matematica, avremmo $P(t) = \mathbb{E}(S(t))(1+\theta) = (\lambda\mu_1(1+\theta))t$, quindi di nuovo una situazione di $P(t) = ct$. Allo stesso modo avviene con il criterio della Varianza si trova che $P(t) = \mathbb{E}(S(t)) + \beta V(S(t)) = \lambda\mu_1 t + \beta\lambda\mu_2 t = (\lambda\mu_1 + \beta\lambda\mu_2)t$, ovvero $P(t) = ct$.

2.2 11/10/2019

Coefficiente di Aggiustamento

Sia dato il processo $\{R(t), t \geq 0\}$, nelle ipotesi del modello della teoria collettiva del rischio (*TCR*), tale che $R(t) = R + ct - S(t)$, $c > 0$. Sia $S(t) = \sum_{h=1}^{N(t)} Y_h$ e $\{S(t), t \geq 0\}$ un processo di Poisson composto con parametri (λ, F_Y). Siano le seguenti ipotesi:

a) Sia $c > \lambda\mu$, allora dato che $\mathbb{E}(S(t) - S(s)) = \lambda\mu(t-s)$ e $P(t) - P(s) = c(t-s)$, si ha che $\forall s < t$, $P(t) - P(s) > \mathbb{E}(S(t) - S(s))$. Quindi i premi hanno un caricamento di sicurezza.

b) Sia F_Y dotato di funzione generatrice dei momenti, ovvero:

$$\exists \mathrm{I}_0 : \quad m_Y(t) = \int_{-\infty}^{+\infty} e^{ty} dF_Y < +\infty$$

Sia $\gamma = \sup\{t \geq 0 | m_Y(t) < +\infty\}$, allora $[0,\gamma[$ è il più grande intervallo incluso in $\mathbb{R}_+$ tale che $m_Y(\cdot)$ è finito.

c) $\lim_{t\to\gamma^-} [\lambda m_Y(t) - (\lambda + ct)] = +\infty$.

In questo modo non consideriamo le funzioni generatrici dei momenti che ammettano un salto tra γ^- e γ, così avremo $\lim_{t\to\gamma^-} m_Y(t) = \lim_{t\to\gamma^+} m_Y(t)$.

Proposizione

Se valgono le ipotesi del modello della *TCR* per il processo di Surplus $\{R(t), t \geq 0\}$ e le ipotesi a), b) e c), allora $\lambda + ct = \lambda m_Y(t)$, con $t \in [0, +\infty[$ ammette una ed una sola radice positiva. Quindi esiste un unico $\alpha > 0$, tale che $\lambda + \alpha c = \lambda m_Y(\alpha)$. La soluzione unica α è chiamata *coefficiente di aggiustamento*. Non è possibile dare un significato reale ad α, ma è importante poiché entra in gioco nella *Disuguaglianza di Lundberg*.

Teorema (Disuguaglianza di Lundberg)

Nelle ipotesi del modello TCR e con le ipotesi a), b) e c), si ha che $\Psi(R) \leq e^{-\alpha R}$, $\forall R > 0$. Dove $\Psi(\cdot)$ è la probabilità asintotica di rovina e α è il coefficiente di aggiustamento.

Dimostrazione Proposizione

Poniamo $g_1(t) = \lambda + ct$ e $g_2(t) = \lambda m_Y(t)$. Abbiamo che $g_1(\cdot)$ è una funzione semplice lineare, ci concentreremo quindi $g_2(\cdot)$. Sappiamo che in zero $g_2(0) = \lambda$, che esiste la derivata prima in zero e $g_2'(0) = \lambda\mu$, per la proprietà della funzione generatrice dei momenti. Sappiamo inoltre che $g_2(\cdot)$ è strettamente convessa, ovvero che $\forall t_1 \neq t_2 \in \mathbb{D}$, $\forall 0 < \delta < 1$, posto $\overline{\delta} = 1 - \delta$, allora $g_2(\delta t_1 + \overline{\delta} t_2) < \delta g_2(t_1) + \overline{\delta} g_2(t_2)$. Lo si può dimostrare notando che $exp(\cdot)$ è strettamente convessa e che :

$$g_2(\delta t_1 + \overline{\delta} t_2) = \lambda \int_{-\infty}^{+\infty} e^{(\delta t_1 + \overline{\delta} t_2)y} dF_Y(y) < \lambda \left[\delta \int_{-\infty}^{+\infty} e^{t_1 y} dF_Y(y) + \overline{\delta} \int_{-\infty}^{+\infty} e^{t_2 y} dF_Y(y) \right] =$$

$$= \delta\lambda m_Y(t_1) + \overline{\delta}\lambda m_Y(t_2) = \delta g_2(t_1) + \overline{\delta} g_2(t_2)$$

Inoltre $g_2(\cdot)$ è continua, poiché considerata $g_2|_{]0,\gamma[}$, dato che questa è convessa nell'intervallo aperto $]0, \gamma[$, allora è anche continua nell'intervallo, ed essendo derivabile in zero per definizione di funzione generatrice dei momenti, allora g_2 è continua in $[0, \gamma[$.
Posto quindi $g(t) = g_2(t) - g_1(t)$, dobbiamo dimostrare che $\exists t \in \mathbb{R}$ tale che $g(t) = 0$, oltre a $t = 0$ ($g(0) = 0 + 0$). Se deriviamo $g(\cdot)$ in zero troviamo $g'(0) = \lambda\mu - c < 0$ che è negativo per ipotesi, ne segue che $g(\cdot)$ è decrescente in zero, quindi $\exists$ U_0^+ tale che $g(t) < g(0)$, $\forall t \in \mathrm{U}_0^+$. Si noti inoltre che $g(\cdot)$ è strettamente convessa, poiché somma di una funzione strettamente convessa ed una lineare, e che $\lim_{t\to\gamma} g(t) = \lim_{t\to\gamma} \lambda m_Y(t) - (\lambda + ct) = +\infty$ per ipotesi. Ma allora, poiché $g(0) = 0$, $\exists$ U_0^+ tale che $g(t) < g(0)$, $\forall t \in \mathrm{U}_0^+$ ed è continua, allora per il teorema della di connessione, $g(\cdot)$ assume tutti i punti tra $\min_{t>0}(g(t))$ e $+\infty$. Quindi $\exists t \neq 0$ tale che $g(t) = 0$. Inoltre possiamo dimostrare che è unico.
Supponiamo per assurdo che $\exists \alpha_1, \alpha_2$ con $\alpha_1 < \alpha_2$, tale che $g(\alpha_1) = g(\alpha_2) = 0$. Allora possiamo esprimere α_1 come mistura di $[0, \alpha_2]$ $\Rightarrow$ $\exists \delta \in]0, 1[$ tale che $\alpha_1 = \delta \cdot 0 + (1 - \delta)\alpha_2$. Allora dato che $0 = g(\alpha_1) = g(\delta \cdot 0 + (1 - \delta)\alpha_2) < \delta g(0) + (1 - \delta)g(\alpha_2) = \delta \cdot 0 + (1 - \delta) \cdot 0 = 0$ e otteniamo $0 < 0$, ovvero una contraddizione.

Osservazione 1
Il coefficiente α dipende da λ, c e $m_Y(\cdot)$ quindi θ e $F_Y(\cdot)$. Fissati λ e $F_Y(\cdot)$, studiamo come varia α al variare di θ. Graficamente α è il punto in cui si intersecano $g_2(\cdot)$ e $g_1(\cdot)$. Il grafico di $g_1(\cdot)$ è un retta con coefficiente angolare c ed intercetta λ, mentre $g_2(\cdot)$ è una curva minore di $g_1(\cdot)$ fino ad α e maggiore oltre. Inoltre $g_2(0) = g_1(0) = \lambda$. Quindi minore è θ, minore è c e quindi α si avvicina a zero, ovvero $\theta \to 0 \Rightarrow \alpha \to 0$.

Osservazione 2
Sia $\lambda + \lambda\mu(1+\theta)t = \lambda m_Y(t) \Leftrightarrow 1 + \mu(1+\theta) = m_Y(t)$. Allora notiamo che α non dipende da λ, ma solo da θ e $F_Y(\cdot)$. Quindi la probabilità asintotica di rovina dipende solo dal coefficiente di caricamento e dalla funzione di ripartizione del risarcimento, nel caso che il sinistro si verifichi, e non dal numero medio di sinistri nell'intervallo unitario.

Esempio
Sia $F_Y \sim exp(\rho)$, allora si ha $\mu = \frac{1}{\rho}$ e $m_Y(t) = \frac{\rho}{\rho - t}$ con $t < \rho$. Vogliamo allora trovare il t che soddisfi la seguente equazione:

$$1 + \frac{1}{\rho}(1+\theta)t = \frac{\rho}{\rho - t} \Leftrightarrow \rho(\rho - t) + (1+\theta)t(\rho - t) = \rho^2 \Leftrightarrow \rho^2 - \rho t + (1+\theta)t\rho - (1+\theta)t^2 = \rho^2 \Leftrightarrow$$

$$\Leftrightarrow (1+\theta)t^2 + \rho t(1 - 1 - \theta) = 0 \Leftrightarrow t((1+\theta)t - \rho\theta) = 0$$

Si hanno quindi due soluzioni: $t = 0$ e $t = \frac{\rho\theta}{1+\theta}$, che è il coefficiente di aggiustamento α per la distribuzione esponenziale.

In generale non sempre è facile determinare α come soluzione dell'equazione, però essendo questo in $]0, \gamma[$, è possibile applicare metodi di stima iterativi, come il metodo di bisezione, che consentono di calcolarne una stima approssimata, ma accurata.

Teorema 1

Date le ipotesi del modello della teoria collettiva del rischio e le ipotesi a, b e c. Siano R il capitale iniziale allocato nel portafoglio, T l'istante di rovina, α il coefficiente di aggiustamento e $\Psi(\cdot)$ la probabilità asintotica di rovina, allora si prova che :

$$\Psi(R) = \frac{e^{-\alpha R}}{\mathbb{E}(e^{-\alpha R}|T < +\infty)}$$

Osservazione

Fissati R e $F_Y(\cdot)$, andiamo a studiare come varia $\Psi(\cdot)$ per $\alpha \to 0$.

$$\lim_{\alpha \to 0} \Psi(R) = \lim_{\alpha \to 0} \frac{e^{-\alpha R}}{\mathbb{E}(e^{-\alpha R}|T < +\infty)} = \lim_{\alpha \to 0} \frac{\overbrace{e^{-\alpha R}}^{\to 1}}{\underbrace{m_{R(T)|T<+\infty}(-\alpha)}_{\to 1}} = 1$$

Quindi per $\theta \to 0$, si ha che $\alpha \to 0$ e quindi che la probabilità asintotica di rovina $\Psi(R) \to 1$ per ogni $R \in [0, +\infty[$.

Teorema 2

Date le ipotesi del modello della teoria collettiva del rischio e le ipotesi a, b e c, allora la probabilità asintotica di rovina $\Psi(\cdot)$ soddisfa la seguente equazione:

$$\Psi'(R) = \frac{\lambda}{c}\Psi(R) - \frac{\lambda}{c}\int_0^R \Psi(R-y)dF_Y(y) - \frac{\lambda}{c}[1 - F_Y(R)]$$

Esempio pratico

Sia p_0 la probabilità di rovina massima che un assicuratore è disposto ad accettare (per *Solvency* II p_0 è 0.5%). Si vuole quindi $e^{-\alpha R} \leq p_0$, così che $\Psi(R) \leq p_0$. Fissati R e F_Y allora $e^{-\alpha R} \leq p_0 \Leftrightarrow -\alpha R \leq \ln(p_0)$. Troviamo allora $\alpha \geq -\frac{1}{R}\ln(p_0)$, cerchiamo ora il θ corrispondente all'α trovato.

$$1 + \mu(1+\theta)\alpha = m_Y(\alpha) \Leftrightarrow 1 + \mu(1+\theta)(-\frac{1}{R}\ln(p_0)) = m_Y(-\frac{1}{R}\ln(p_0)) \Leftrightarrow$$

$$1 + \theta = \frac{m_Y(-\frac{1}{R}\ln(p_0)) - 1}{\mu(-\frac{1}{R}\ln(p_0))} \Leftrightarrow \theta = \frac{m_Y(-\frac{1}{R}\ln(p_0)) - 1}{\mu(-\frac{1}{R}\ln(p_0))} - 1$$

Se fissiamo invece θ e F_Y si ha che $-\alpha R \leq \ln(p_0) \Leftrightarrow R \geq \frac{-\ln(p_0)}{\alpha}$

2.3 14/10/2019

Per controllare $\Psi(R)$ possiamo quindi allocare una quantità adeguata di capitale al portafoglio, ma allocarlo significa che non viene distribuito ai soci e che deve essere investito secondo certe regole, su strumenti finanziari con rendimenti bassi. Allocare troppo capitale è quindi un costo per gli azionisti, avendo effetto sulla creazione di valore.

Valutiamo adesso l'effetto di una copertura riassicurativa, fissati θ, F_Y e R. La limitazione superiore dei rischi potrebbe essere troppo elevata e non garantirci di essere sotto la soglia p_0, si può perciò agire sull'α in $e^{-\alpha R}$ tramite la riassicurazione. Sia $S(t) = \sum_{h=1}^{N(t)} Y_h$ con Y_h l'importo che la cedente deve risarcire

all'assicurato per l'h-esimo sinistro, e lo scomponiamo in due parti: $Y_h = \gamma(Y_h) + (Y_h - \gamma(Y_h))$, con $\gamma(Y_h)$ la parte di risarcimento conservata dalla cedente. Di conseguenza si ha:
∘ $S(t)_{CED} = \sum_{h=1}^{N(t)} \gamma(Y_h)$ che è l'impegno della cedente, al netto della riassicurazione.
∘ $S(t)_{RIASS} = \sum_{h=1}^{N(t)} Y_h - \gamma(Y_h)$ che è l'impegno del riassicuratore.
Valutiamo un'impostazione di tipo collettivo e consideriamo le seguenti forme riassicurative:
Quota-Share: fissata un'aliquota $0 < a \leq 1$, si ha $\gamma(Y_h) = aY_h$ e $Y_h - \gamma(Y_h) = (1-a)Y_h$, quindi $S_{CED}(t) = aS(t)$ e $S_{RIASS}(t) = (1-a)S(t)$ *Excess of Loss:* fissata una priorità $L > 0$, si ha $\gamma(Y_h) = \min\{Y_h, L\}$ e $Y_h - \gamma(Y_h) = \max\{0, Y_h - L\}$, quindi $S_{CED}(t) = \sum_{h=1}^{N(t)} \min\{Y_h, L\}$ e $S_{RIASS}(t) == \sum_{h=1}^{N(t)} \max\{0, Y_h - L\}$.

Osservazione

La legge probabilistica di $\{R(t), t \geq 0\}$ viene a dipendere dalla legge di $\{S(t), t \geq 0\}$ e nella TCR abbiamo fatto ipotesi su $\{N(t), t \geq 0\}$, che sia un processo di Poisson con intensità λ, e su $\{Y_1, Y_2, \dots\}$. Consideriamo il processo $\{S_{CED}(t), t \geq 0\}$, questo dipende dal processo degli arrivi $\{N(t), t \geq 0\}$ e dal processo dei risarcimenti ritenuti $\{\gamma(Y_1), \gamma(Y_2), \dots\}$. Sia H un evento logicamente dipendente da $\{N(t), t \geq 0\}$, che implica si sono verificati i sinistri con indice $1, \dots, n$; allora $\gamma(Y_1)|H, \dots, \gamma(Y_n)|H$ sono equivalenti a $\gamma(Y_1|H), \dots, \gamma(Y_n|H)$ ed essendo trasformati tutti per la medesima funzione, mantengono l'identica distribuzione. Inoltre sono anche stocasticamente indipendenti, quindi $\gamma(Y_1|H), \dots, \gamma(Y_n|H)$ IID $\sim F_{\gamma(Y)}$. Nelle ipotesi precedenti avevamo ricavato che il processo $\{S(t), t \geq 0\}$ è un processo di Poisson Composto, una volta assegnati (λ, F_Y). Perciò analogamente $\{S_{CED}(t), t \geq 0\}$ è un processo di Poisson Composto con parametri $(\lambda, F_{\gamma(Y)})$.

Modifica del processo $\{R(t), t \geq 0\}$, tenendo conto della copertura riassicurativa

Si ha un nuovo processo $\{R(t)_{CED}, t \geq 0\}$ dove $R_{CED}(t) = R + ct - c_r t - S_{CED}(t) = R + (c - c_r)t - S_{CED}(t)$. Si ipotizza che i premi al riassicuratore siano espressi tramite una funzione lineare $P_{RIASS}(t) = c_r t$, con $c_r > 0$. Il coefficiente c_r rappresenta il montepremi richiesto dal riassicuratore per la copertura riassicurativa in un intervallo di tempo unitario ed è pari al risarcimento atteso su un intervallo unitario da parte del riassicuratore: $c_r = \mathbb{E}(S_{RIASS}(t+1) - S_{RIASS}(t))(1+\xi)$, con il coefficiente di caricamento di sicurezza per il riassicuratore $\xi > 0$.
Analogamente al discorso precedente è necessario per l'assicuratore che i premi della cedente siano caricati, perciò si pone il vincolo $c - c_r > \mathbb{E}(S_{CED}(t+1) - S_{CED}(t)) = \lambda\mathbb{E}(\gamma(Y))$. ed inoltre l'equazione per il calcolo del coefficiente di aggiustamento α_{CED} diventa: $\lambda + (c - c_r)t = \lambda m_{\gamma(Y)}(t)$.
Quota-Share
∘ $\mathbb{E}(\gamma(Y)) = a\mathbb{E}(Y) = a\mu$
∘ $\mathbb{E}(Y - \gamma(Y)) = (1-a)\mathbb{E}(Y) = (1-a)\mu$
∘ $m_{\gamma(Y)}(t) = m_{aY}(t) = m_Y(at)$
∘ $c - c_r > \mathbb{E}(S_{CED}(t+1) - S_{CED}(t)) = \lambda\mathbb{E}(\gamma(Y)) \Leftrightarrow \lambda\mu(1+\theta) - \lambda\mu(1-a)(1+\xi) > \lambda a\mu \Leftrightarrow \theta + a - \xi + a\xi > a \Leftrightarrow \theta - \xi > a - a(1+\xi) \Leftrightarrow \xi - \theta < \xi a \Leftrightarrow a > 1 - \frac{\theta}{\xi} \Rightarrow a > 0$ se e solo se $\theta < \xi$.
Perciò il coefficiente di aggiustamento, rispetto a quello in assenza del riassicuratore, è aumentato dato che $0 < a \leq 1$, che comporta un maggior controllo sulla limitazione superiore e di conseguenza sulla probabilità di rovina; $\alpha_{CED} > \alpha \Rightarrow e^{-\alpha R} < e^{-\alpha_{CED} R}$. Inoltre al decrescere di a, α_{CED} aumenta.
Excess of Loss
∘ $\mathbb{E}(\gamma(Y)) = \mathbb{E}(\min(Y, L)) = \int_{-\infty}^{+\infty} 1 - F_Y(y)dy = \int_0^L \overline{F}_Y(y)dy$
∘ $\mathbb{E}(Y - \gamma(Y)) = \mathbb{E}(\max(0, Y - L)) = \int_L^{+\infty} F_Y(y)dy$
∘ $c - c_r = \lambda\mu(1+\theta) - \lambda\int_L^{+\infty} \overline{F}_Y(y)dy(1+\xi) > \lambda\int_0^L \overline{F_Y}(y)dy$.

Modello Individuale Monoperiodale della Teoria del Rischio
Sia un fissato portafoglio di rischi "analoghi" su un orizzonte monoperiodale e siano $R > 0$ il capitale allocato in zero nel portafoglio, X il risarcimento totale aleatorio pagato nell'intervallo $[0, 1]$ e P i premi incassati nell'intervallo. Il risarcimento totale è descritto dal modello secondo un approccio di tipo individuale, ovvero $X = \sum_{i=1}^{n} X_i$, dove X_i è il risarcimento totale per una singola polizza ed in forza dell'analogia si assume $X_1, \dots, X_n \sim$IID, con $\mathbb{E}(X_i) \triangleq \mu$ e $V(X_i) \triangleq \sigma^2$.
Il risultato di portafoglio alla fine del periodo sarà dato da $R + P - X$. L'evento rovina è definito come $(R + P - X < 0) \Leftrightarrow (R + G < 0) \Leftrightarrow (G < -R)$, con G n.a. del *profit-loss*, ed la probabilità di rovina sarà $\Psi(R, 1) = Pr(G < -R)$. Possiamo considerare G come somma dei guadagni sulle singole polizze G_i con $G_1, \dots, G_n \sim$IID.

Teorema di Lindeberg-Levy (Teorema del Limite Centrale): Sia una successione di n.a. $X_1, X_2, \dots \sim$IID con $\sigma^2 < +\infty$. Posti $S_n = X_1 + \dots + X_n$ la somma e $S_n* = \frac{S_n - \mathbb{E}(S_n)}{\sigma(S_n)}$ la somma standardizzata, allora la successione converge in distribuzione ad una Normale Standard. Quindi per n grande possiamo considerare $X_1, \dots, X_n \sim$IID e $S_n* \sim \Phi$.

Nel nostro caso quindi se n è grande, avremo $\frac{G - \mathbb{E}(G)}{\sigma(G)} \sim N(0, 1) \Rightarrow Pr(G < -R) \approx \Phi(-\frac{R + \mathbb{E}(G)}{\sigma(G)})$.
Quindi quando $\frac{R + \mathbb{E}(G)}{\sigma(G)}$ aumenta, $\Psi(R, 1)$ diminuisce; infatti il numero aleatorio standardizzato è detto *indice di stabilità relativa del portafoglio* e dipende da R, $\mathbb{E}(G)$ e $\sigma(G)$, sul quale possiamo agire tramite la riassicurazione, ma che ha delle conseguenze anche su $\mathbb{E}(G)$.

3 Personalizzazione a Posteriori

3.1 17/10/2019

Ripartiamo il portafoglio in classi tariffarie ed ipotizziamo di assumere all'interno di queste la stessa valutazione probabilistica. Attraverso questa valutazione si calcola il *premio a priori*. In alcuni ambiti però, come RCA, l'eterogeneità interna alla classe rimane molto alta e per questo è importante considerare la personale storia di sinistrosità per l'i-esimo assicurato. Si passa quindi da un premio a priori ad un premio basato sull'esperienza individuale, in modo tale che più tempo osservo l'assicurato, più il premio sarà commisurato alla sua effettiva sinistrosità. Per fare questo può essere utilizzato l'approccio bayesiano, che è dal punto di vista metodologico l'approccio più soddisfacente, ma che comporta una elevata complessità ed un obbligo di valutazioni probabilistiche a priori. Dei metodi efficaci e allo stesso tempo più semplici dell'approccio bayesiano sono la *teoria della credibilità* e il *sistema Bonus-Malus*.

Inciso
Siano i n.a. Y e U ed inoltre $\forall u$ siano dati la legge di $Y|U = u$ e $\mathbb{E}(Y|U = u)$, tale che $\mathbb{E}(Y) = \int \mathbb{E}(Y|U = u)dF_U(u)$. Nell'impostazione di De Finetti delle probabilità coerenti la scrittura risulta sensata, poiché si può avere come evento condizionante un evento di probabilità nulla, mentre nell'impostazione classica non è consentito. Per giustificare la scrittura ci sono diverse soluzioni:
1 Considerare per il n.a. Y una famiglia di valutazioni probabilistiche, indiciate sulle determinazioni di U $(F_U)u$, la FdR $Pr_U(Y \leq y)$ e $\mathbb{E}_U(Y)$, secondo la valutazione F_U e $\mathbb{E}(Y) = \int \mathbb{E}_u(Y)dF_U(u)$
2 Si fissa uno spazio di probabilità $(\mathbb{P}, \mathcal{A}, Pr)$ e sia Y un n.a. di questo spazio $(Y(\cdot) : \mathbb{P} \to \mathbb{R})$ tale che Y è un'applicazione $\mathcal{A}, \mathcal{B}$-misurabile $\Leftrightarrow \forall B \in \mathcal{B}$ si ha $Y^{-1}(B) \in \mathcal{A}$, ovvero se la controimmagine di ogni insieme boreliano, appartiene alla sigma algebra di partenza $\mathcal{A}$. Supponiamo che $\exists \mathbb{E}(|Y|) < +\infty$ e che sia

U un n.a. tale che $U : \mathbb{P} \to \mathcal{X} \subset \mathbb{R}$ con U applicazione $\mathcal{A}, \mathcal{B}$-misurabile. Se U è finito e $Pr(U = u) > 0\ \forall u$, nell'impostazione classica si pone $\mathbb{E}(Y|U = u) = \frac{\mathbb{E}(Y|U=u|)}{Pr(U=u)} \Rightarrow \mathbb{E}(Y) = \mathbb{E}(Y|\Omega|) = \mathbb{E}(Y|\bigvee_u (U = u)|) = \mathbb{E}(\sum_u Y|U = u|) = \sum_u \mathbb{E}(Y|U = u|) = \sum_u \mathbb{E}(Y|U = u)Pr(U = u)$.
Nel caso continuo abbiamo tutti eventi di probabilità nulla e quindi queste speranze matematiche non sono definite. Può comunque essere giustifica la disintegrabilità nell'impostazione classica, infatti fissato uno spazio di probabilità $(\mathbb{P}, \mathcal{A}, Pr)$, un n.a. Y $\mathcal{A}, \mathcal{B}$-misurabile con speranza matematica finita, allora esiste una funzione detta funzione di regressione di Y su U indicata con $\mathbb{E}(Y|U = \cdot) : \mathcal{X} \to \mathbb{R}$ definita a meno di insiemi trascurabili con Pr_U, la misura di probabilità in $\mathbb{R}$ che attribuisce ad un boreliano $Pr_U(B) = Pr(U \in B)$ ed ha le seguenti proprietà:
∘ $\mathbb{E}(Y|U = \cdot)$ è Pr_U-integrabile $\Leftrightarrow \exists \int_B \mathbb{E}(Y|U = u)Pr_U(du)$ ed è finito
∘ $\mathbb{E}(Y|U \in B|) \triangleq \int_{\{U \in B\}} Y F_Y(dy) = \int_B \mathbb{E}(Y|U = u)F_U(du) = \int_B \mathbb{E}(Y|U = u)dF_U(u)$.
Dalla seconda proprietà segue che $\mathbb{E}(Y) = \mathbb{E}(Y|\Omega|) = \int_{\mathbb{R}} \mathbb{E}(Y|U = u)Pr_u(du) = \int \mathbb{E}(Y|U = u)dF_U(u)$.
Quindi in questa ipotesi si dimostra che esiste una funzione da cui è possibile ricavare $\mathbb{E}(Y)$ tramite l'integrale di una $g(\cdot)$ reale di una variabile reale. Inoltre la proprietà di $\mathbb{E}(Y|U = \cdot)$ sono le medesime della speranza matematica:
∘ $\mathbb{E}(\alpha|U = \cdot) = \alpha$ se $\alpha \in \mathbb{R}$ certo
∘ $a \leq Y \leq b$ implica $a \leq \mathbb{E}(Y|U = \cdot) \leq b$
∘ $Y_1 = Y_2$ q.c. $\Rightarrow \mathbb{E}(Y_1|U = \cdot) = \mathbb{E}(Y_2|U = \cdot)$
∘ $\mathbb{E}(\sum_i^n \alpha_i Y_i|U = \cdot) = \sum_i^n \alpha_i \mathbb{E}(Y_i|U = \cdot)$
∘ $\mathbb{E}(\mathcal{H}(u)Y|U = \cdot) = \mathcal{H}(u)\mathbb{E}(Y|U = \cdot)$
∘ $\mathbb{E}(Y|U = \cdot) = \mathbb{E}(Y)$ se Y, U stocasticamente indipendenti.

Osservazione

La disintegrabilità può avere anche un'altra interpretazione, ovvero dati $(\mathbb{P}, \mathcal{A}, Pr)$, Y con speranza matematica finita, se $\mathcal{H}$ è una sotto σ-algebra di $\mathcal{A}$, si prova che esiste un n.a. indicato con $\mathbb{E}(Y|\mathcal{H})$ detta speranza matematica di Y condizionata alla σ-algebra $\mathcal{H}$, tale che $\mathbb{E}(Y|\mathcal{H})$ è $\mathcal{H}, \mathcal{B}$-misurabile e $\mathbb{E}(Y|\mathcal{H}|) = \int_H Y dPr = \int_H \mathbb{E}(Y|\mathcal{H})dPr$ a meno di insiemi trascurabili.
Da ciò segue che $\mathbb{E}(Y) = \mathbb{E}(Y|\Omega|) = \int Y dPr = \int \mathbb{E}(Y|\mathcal{H})dPr = \mathbb{E}(\mathbb{E}(Y|\mathcal{H}))$.
Date le ipotesi a e b, sia $\mathcal{U}$ la σ-algebra indotta dal n.a. U tale che $\mathcal{U} \subset \mathcal{A}$, allora esiste $\mathbb{E}(Y|\mathcal{U}) \triangleq \mathbb{E}(Y|U)$ e il secondo membro è una versione del primo membro.

Approccio Bayesiano

Consideriamo una collettività di rischi "analoghi" ed "indipendenti", ovvero consideriamo una classe tariffaria. Prendiamo in esame il rischio i-esimo della collettività e $\{Y_{i1}, Y_{i2}, \dots\}$ è un processo stocastico di n.a. a parametro discreto, dove Y_{it} è il n.a. di interesse nell'anno t.
Per l'eterogeneità residua dovuta a fattori non osservabili, si introduce il parametro aleatorio U_i, chiamato *parametro aleatorio di rischio*, è specifico dell'individuo, da cui dipende il processo di interesse e l'aleatorietà di questo parametro è dovuta al fatto che tiene conto di elementi non osservabili. Si accolgono le seguenti ipotesi:
1 $\forall u$ determinazione di U_i sia assegnata la legge del processo $\{Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots\}$ e si assume che siano stocasticamente indipendenti
2 Sia assegnata la legge di U_i.
Di conseguenza è assegnata la legge del processo $\{Y_{i1}, Y_{i2}, \dots\}$, poiché per ogni A evento logicamente dipendente dal processo possiamo calcolare $Pr(A) = \int Pr(A|U_i = u)dF_U(u) = \mathbb{E}(|A|) = \int \mathbb{E}(|A||U_i = u)dF_U(u) = \mathbb{E}(\mathbb{E}(|A||\mathcal{U}))$. Formalmente le ipotesi qualitative si traducono come:

∘ Analogia: I processi $\{U_i, Y_{i1}, Y_{i2}, \dots\}$ sono identicamente distribuiti al variare di $i \Leftrightarrow \forall i, j \; U_i \stackrel{d}{=} U_j$ e $\{Y_{i1}, Y_{i2}, \dots\}|U_i = u \stackrel{d}{=} \{Y_{j1}, Y_{j2}, \dots\}|U_i = u$
∘ Indipendenza: I processi $\{U_i, Y_{i1}, Y_{i2}, \dots\}$ sono stocasticamente indipendenti al variare di i.

Grazie alle alle ipotesi di IID, fissato l'i-esimo rischio nella collettività con $\{Y_1, Y_2, \dots\}$ il processo di interesse e U il parametro aleatorio di rischio, si ha la seguente ipotesi:
∘ $\forall u$ determinazione di U è assegnata la legge di $Y_1|U = u, Y_2|U = u, \dots$ e sono stocasticamente indipendenti. Questa indipendenza condizionata è però una forma di dipendenza, infatti presi i tempi $(t_1, \dots, t_n)$, la FdR $F_{Y_{t_1},\dots,Y_{t_n}}(y_1, \dots, y_n) = Pr(Y_{t_1} \leq y_1, \dots, Y_{t_n} \leq y_n) = \int Pr(Y_{t_1} \leq y_1, \dots, Y_{t_n} \leq y_n|U = u)dF_U(u) = \int \prod_{h=i}^{n} Pr(Y_{t_h} \leq y_h)dF_U(u) \neq \prod_{h=i}^{n} \int Pr(Y_{t_h} \leq y_h)dF_U(u)$.
Per fissare il premio equo in $T + 1$ possiamo considerare $\mathbb{E}(Y_{T+1})$ che è comune a tutti gli assicurati, essendo questi analoghi, ed è perciò il *premio collettivo*. Notiamo inoltre che $\mathbb{E}(Y_{T+1}) = \int \mathbb{E}(Y_{T+1}|U = u)dF_U(u) = \mathbb{E}(\mathbb{E}(Y_{T+1}|U))$. E poiché U tiene conto delle caratteristiche dell'assicurato, se conoscessimo $U = u$, sarebbe più opportuno utilizzare il *premio individuale* $\mathbb{E}(Y_{T+1}|U = u) = \mu_{T+1}(u)$.

3.2 18/10/2019

Per assegnare il premio individuale $\mu_{T+1}(U)$ che è un n.a., si parte dal premio collettivo $\mathbb{E}(Y_{T+1})$ e sfruttando le informazioni sulla sinistrosità si crea una stima basata sull'esperienza del premio individuale $\mathbb{E}(Y_{T+1}|U)$.
Nell'approccio bayesiano abbiamo come informazione l'evento $H_T = (Y_1 = y_1, \dots Y_T = y_T)$, con $Pr(H_T) > 0$, e si considera la speranza matematica $\mathbb{E}(Y_{T+1}|H_T) = \int y dF_{Y_{T+1}|H_T}(y) = \int \mathbb{E}(Y_{T+1}|H_T, U = u)dF_{U|H_T}(u) = \int \mathbb{E}(Y_{T+1}|U = u)dF_{U|H_T}(u) = \int \mu_{T+1}(u)dF_{U|H_T}(u) = \mathbb{E}(\mu_{T+1}(U)|H_T)$, ovvero il valore atteso del premio individuale, valutato tramite la distribuzione a posteriori. Quindi questa tecnica ha l'effetto di precisare la distribuzione del parametro aleatorio, tenendo conto dell'informazione sull'individuo $F_U \to F_{U|H_T}$.

Osservazione: fissati due individui della collettività, per le ipotesi fatte $U_j \stackrel{d}{=} U_i \Rightarrow F_{U_j} = F_{U_i}$, cioè hanno la medesima distribuzione a priori. Si ottengono in seguito le storie dei singoli assicurati H_{iT} e H_{jT} e i n.a. $U_j|H_{jT}$ e $U_i|H_{iT}$, che portano in generale a valutazioni diverse a posteriori, data le diverse storie degli assicurati.

Poniamo come stimatore del premio individuale del singolo individuo $\widetilde{\mu}_{T+1}(H_T) = \mathbb{E}(\mu_{T+1}(U)|Y_1, \dots, Y_T)$, interpretabile come:
1 Composizione tra l'applicazione che definisce la T-upla e la funzione di regressione di μ_{T+1}.
2 il valore atteso condizionato $\mathbb{E}(\mu_{T+1}(U)|\mathcal{U})$, dove $\mathcal{U}$ è la σ-algebra indotta da $(Y_1, \dots, Y_T)$, quindi $\widetilde{\mu}_{T+1}(H_T)$ è una versione delle speranza matematica condizionata.
É detto stimatore bayesiano del premio individuale e $\widetilde{\mu}_{T+1}(H_T) = \mathbb{E}(\mu_{T+1}(U)|Y_1, \dots, Y_T) = \mathbb{E}(\mathbb{E}(Y_{T+1}|U = u)|Y_1, \dots, Y_T) = \mathbb{E}(Y_{T+1}|Y_1, \dots, Y_T)$, grazie all'ipotesi di indipendenza stocastica condizionata ad H_T.

Inciso
Sia X un n.a. non osservabile tale che $\exists V(X) < +\infty$ e $Y_1, \dots, Y_T$ un processo di osservazione. Allora esiste una funzione del processo di osservazione $g(\cdot_1, \dots, \cdot_T)$ tale che meglio approssimi X. Introduciamo la funzione $L(X, g(\cdot_1, \dots, \cdot_T))$, detta *funzione di perdita*, e cerchiamo quindi la g che minimizzi L. La tipica funzione di perdita è la perdita quadratica attesa: sia g quadrato-integrabile, e troviamo g tale che $\min_{g \in \mathcal{G}} \mathbb{E}[(X - g(Y_1, \dots, Y_T))^2]$. Si dimostra che sotto queste ipotesi esiste ed è unica la soluzione e g coincide con la funzione di regressione di X su $(Y_1, \dots, Y_T)$.

Quindi nel nostro caso $\widetilde{\mu}_{T+1}(H_T)$ è la funzione del processo che meglio approssima il premio individuale secondo la minima perdita quadratica attesa.

Modello Poisson-Gamma per il processo di arrivo dei sinistri

Si consideri una collettività di rischi "analoghi" ed "indipendenti", fissato l'individuo i-esimo, consideriamo come processo di interesse il processo $\{N_{i1}, N_{i2}, \dots\}$ dei n.a. di sinistri per l'anno t.
Modelliamo l'eterogeneità residua tramite il parametro aleatorio di rischio U_i e ad ogni assicurato è associato un processo stocastico con le seguenti ipotesi:
a $\{U_i, N_{i1}, N_{i2}, \dots\}$ sono IID al variare di i
b $\forall i$, $\{U_i, N_{i1}, N_{i2}, \dots\}$ a livello di singolo individuo si ha:
- $\forall u$ determinazione di U_i, i numeri aleatori $N_{i1}|U_i = u, N_{i2}|U_i = u, \dots$ sono IID
- $\forall u$ determinazione di U_i, si ha che i numeri aleatori $N_{it}|U_i = u \sim Poi(\lambda u)$
- $U_i \sim \Gamma(\alpha, \alpha)$, il parametro ha quindi speranza matematica unitaria e varianza $\frac{1}{\alpha}$.

Quindi il processo per l'i-esimo individuo dipende dai parametri $\lambda > 0$ e $\alpha > 0$, che non dipendono da i, ma sono legati alla valutazione a priori.
Il processo $\{U_i, N_{i1}, N_{i2}, \dots\}$ con le precedenti ipotesi è detto processo *Poisson-Gamma*(λ, α).
Per il premio per la frequenza dell'individuo i-esimo in $T + 1$, si valuta il n.a. $N_{i,T+1}$, attraverso $\mathbb{E}(N_{i,T+1}|U_i = u) = \lambda u$, quindi la versione del premio individuale è $\mathbb{E}(N_{i,T+1}|U_i) = \lambda U_i$. Inoltre si nota che il premio collettivo $\mathbb{E}(N_{i,T+1}) = \mathbb{E}(\mathbb{E}(N_{i,T+1}|U_i = u)) = \mathbb{E}(\lambda U_i) = \lambda \mathbb{E}(U_i) = \lambda$. Quindi il premio individuale è dato dal prodotto del premio collettivo per un parametro aleatorio di speranza matematica unitaria. Inoltre $V(N_{i,T+1}) = \mathbb{E}(V(N_{i,T+1}|U_i)) + V(\mathbb{E}(N_{i,T+1}|U_i)) = \lambda \mathbb{E}(U_i) + \lambda^2 V(U_i) = \lambda + \frac{\lambda^2}{\alpha}$. E si riscontra che le valutazioni non dipendono da T, quindi $\forall t$ si ha $\mathbb{E}(N_{i,T+1}) = \lambda$ e $V(N_{i,T+1}) = \lambda + \frac{\lambda^2}{\alpha}$. Si conosce quindi la distribuzione dei n.a. poiché può essere visto come il processo di incrementi annui di un processo di misture di Poisson con misturante di speranza matematica unitaria ($\Rightarrow \forall t\ N_{i,t} \sim BN(\alpha, \frac{\alpha}{\alpha+\lambda})$).

<u>Osservazione:</u> al crescere di α, $\frac{1}{\alpha}$ diminuisce, quindi $V(U_i)$ diminuisce. Questo implica che la collettività sia piuttosto omogenea e che quindi non ci sia eterogeneità nella collettività. Viceversa per il diminuire di α.

Fissato un individuo campione, il processo $\{U, N_1, N_2, \dots\}$ e le sue ipotesi, cerchiamo la legge per il processo di arrivo dei sinistri, con $\{N_1, N_2, \dots\}$ processo scambiabile. Sia $\forall T \geq 1$ la sequenza di n.a. $(N_1, \dots, N_T) \stackrel{d}{=} (N_{j_1}, \dots, N_{j_T})$, per ogni $(j_1, \dots, j_T)$ permutazione di $1, \dots, T$, ovvero se e solo se $\forall n_1, \dots, n_T$ si ha che $Pr(N_1 = n_1, \dots, N_T = n_t) = Pr(N_{j_1} = n_1, \dots, N_{j_T} = n_t)$.
<u>Dimostrazione:</u> fissiamo T e $n_1, \dots, n_T$ e calcoliamo la legge congiunta sfruttando l'ipotesi b:

$$Pr(N_1 = n_1, \dots, N_T = n_T) = \int_0^{+\infty} Pr(N_1 = n_1, \dots, N_T = n_T|U = u) dF_U(u) =$$

$$= \int_0^{+\infty} \underbrace{\prod_{t=1}^{T} Pr(N_t = n_t|U = u)}_{1} f_U(u) du \stackrel{o}{=} \qquad \boxed{1 = \prod_{t=1}^{T} \frac{e^{-\lambda u}(\lambda u)^{n_t}}{n_t!} = \frac{e^{-\lambda T u}(\lambda u)^{\sum_{t=1}^{T} n_t}}{\prod_{t=1}^{T}(n_t!)}}$$

$$\stackrel{o}{=} \int_0^{+\infty} \frac{e^{-\lambda T u}(\lambda u)^{\sum_{t=1}^{T} n_t}}{\prod_{t=1}^{T}(n_t!)} \frac{\alpha^\alpha u^{\alpha-1} e^{-\alpha u}}{\Gamma(\alpha)} du = \frac{\alpha^\alpha (\lambda)^{\sum_{t=1}^{T} n_t}}{\Gamma(\alpha) \prod_{t=1}^{T}(n_t!)} \int_0^{+\infty} e^{-u(\lambda T + \alpha)} u^{\sum_t n_t + \alpha - 1} du =$$

$$= \left[\frac{(\alpha+\lambda T)^{\alpha+\sum n_t}}{\Gamma(\alpha+\sum n_t)} \right]^{-1} \frac{\alpha^{\alpha}(\lambda)^{\sum_{t=1}^{T} n_t}}{\Gamma(\alpha)\prod_{t=1}^{T}(n_t!)} = \frac{\Gamma(\alpha+\sum_t n_t)}{\Gamma(\alpha)\prod_t(n_t!)} \left(\frac{\alpha}{\alpha+\lambda T}\right)^{\alpha} \left(\frac{\lambda}{\alpha+\lambda T}\right)^{\sum_t n_t}$$

Sostituendo $Pr(N_1 = n_1, \dots, N_T = n_T)$ con $Pr(N_{j_1} = n_1, \dots, N_{j_T} = n_t)$, otteniamo lo stesso risultato, perciò la distribuzione congiunta di un numero finito di n.a. è la stessa.

<u>**Osservazione:**</u> si ha l'indipendenza stocastica se e solo se $Pr(N_1 = n_1, \dots, N_T = n_T) = \prod_t Pr(N_t = n_t)$. In questo caso non si ha, infatti:

$$\prod_{t=1}^{T} Pr(N_t = n_t) = \prod_{t=1}^{T} \frac{\Gamma(\alpha+\sum_t n_t)}{\Gamma(\alpha) n_t!} \left(\frac{\alpha}{\alpha+\lambda}\right)^{\alpha} \left(\frac{\lambda}{\alpha+\lambda}\right)^{n_t} = \frac{\prod_t \Gamma(\alpha+n_t)}{\prod_t(n_t!)} \left[\frac{1}{\Gamma(\alpha)}\left(\frac{\alpha}{\alpha+\lambda}\right)^{\alpha}\right]^{T} \left(\frac{\lambda}{\alpha+\lambda}\right)^{\sum_t n_t} \neq$$

$$\neq Pr(N_1 = n_1, \dots, N_T = n_T)$$

L'indipendenza si ha qualora i n.a. siano condizionati e non si traduce in indipendenza stocastica tra i n.a. del processo, si ha quindi una dipendenza dalla storia passata.

<u>**Revisione Bayesiana**</u> **(Aggiornamento)**

Siano $\mathbb{E}(N_{T+1}) = \lambda$, $\mathbb{E}(N_{T+1}|U) = \lambda U$ e $H_T(N_1 = n_1, \dots, N_T = n_T)$. Allora si ha che:

$$\mathbb{E}(N_{T+1}|H_T) = \int \mathbb{E}(N_{T+1}|H_T, U = u) dF_{U|H_T}(u) = \int \mathbb{E}(N_{T+1}|U = u) dF_{U|H_T}(u) = \lambda \int u dF_{U|H_T}(u) = \lambda \mathbb{E}(U|H_T)$$

Consideriamo la distribuzione di probabilità a posteriori:

$$f_{U|H_T} \propto f_U(u) \prod_{t=1}^{T} Pr(N_t = n_t | U = u) = \frac{\alpha^{\alpha} u^{\alpha-1} e^{-\alpha u}}{\Gamma(\alpha)} \frac{e^{-\lambda u T}(\lambda u)^{\sum_t n_t}}{\prod_t(n_t!)} \propto u^{(\alpha+\sum_t n_t - 1)} e^{-u(\alpha+\lambda T)}$$

Quindi la distribuzione a posteriori è ancora una *Gamma* con parametri aggiornati, $\Gamma(\alpha,\alpha) \to \Gamma(\alpha + \sum_t n_t, \alpha + \lambda T)$. Da cui troviamo che $\mathbb{E}(U|H_T) = \frac{\alpha+\sum_t n_t}{\alpha+\lambda T}$ e che $\mathbb{E}(N_{T+1}|H_T) = \lambda \frac{\alpha+\sum_t n_t}{\alpha+\lambda T}$.
Il premio bayesiano dipende quindi dalle valutazioni a priori di λ e α, dalla durata del processo di osservazione T e dalla sintesi della storia di sinistrosità $\sum_t n_t$. Possiamo inoltre vedere:

$$\mathbb{E}(N_{T+1}|H_T) = \lambda \underbrace{\frac{\alpha}{\alpha+\lambda T}}_{1-z_t} + \frac{\sum_t n_t}{T} \underbrace{\frac{\lambda T}{\alpha+\lambda T}}_{z_t}$$

Ovvero come una combinazione lineare convessa tra il premio collettivo al tempo $T+1$ e la frequenza sinistri, ovvero la stima empirica di $\mathbb{E}(N_{T+1})$ per l'individuo. Si ha quindi una correzione dell'informazione a priori con quella specifica.
<u>Osservazione:</u>
∘ Fissati α, λ; $T \uparrow$, $z_t \uparrow$, cioè più si osserva il rischio, maggiore è il peso della storia individuale, in particolare se $T \to +\infty$, $z_t \to 1$;
∘ Fissati λ, T; $\alpha \uparrow, \alpha^{-1} \downarrow, V(U_i) \downarrow, z_t \downarrow$ quindi il premio collettivo è un adeguata sintesi, essendo la collettività abbastanza omogenea.

3.3 21/10/2019

Inoltre possiamo considerare $C_{T+1}(H_T) = \frac{\mathbb{E}(N_{T+1}|H_T)}{\mathbb{E}(N_{T+1})}$, ovvero il rapporto tra il premio equo a posteriori e a priori, ed è detto *coefficiente di revisione bayesiana*, poiché moltiplicandolo per il premio a priori, si ottiene il premio a posteriori.

∘ Se $C_{T+1}(H_T) > 1$ allora $\alpha + \sum_t n_t > \alpha + \lambda T$, e questo se e solo se il numero dei sinistri osservati è superiore al valore atteso iniziale, quindi si avrà un aggravamento del premio. Tuttavia non risulta soddisfacente considerare la sola sintesi di sinistrosità $\sum_t n_t$, non tenendo conto della sequenza $(n_1, \dots, n_T)$. Infatti se osserviamo due individui entrambi per 10 anni con $s_1 = (1, 0, \dots, 0)$ e $s_2 = (0, \dots, 0, 1)$, non è ragionevole attribuire la medesima valutazione probabilistica al numero atteso di sinistri,dovuto al fatto di considerare un parametro aleatorio U statico, mentre sarebbe più adeguato considerarne uno che varia nel tempo $U_1, U_2, \dots$, considerando quindi l'età dei sinistri.

Esempio: si consideri un portafoglio suddiviso in due classi, siano λ_1 e $\lambda_2 = 2\lambda_1$ ed α misturante con medesima distribuzione. A parità di storia dei sinistri, l'assicurato in classe tariffaria 1 subisce una diminuzione del premio inferiore o un aggravio maggiore, rispetto all'assicurato in classe tariffaria 2. É anche possibile ricavare la distribuzione a posteriori di $N_{T+1}|H_T$:

$$Pr(N_{T+1} = n|H_T) = \int_0^{+\infty} Pr(N_{T+1} = n|H_T, U = u)dF_{U|H_T}(u) =$$

$$= \int_0^{+\infty} Pr(N_{T+1} = n|H_T, U = u)f_{U|H_t}(u)du = \int_0^{+\infty} \frac{e^{-\lambda u}(\lambda u)^n}{n!} \frac{(\alpha + \lambda T)^{\alpha+\sum_t n_t}}{\Gamma(\alpha + \sum_t n_t)} u^{\alpha+\sum_t n_t - 1} e^{-u(\alpha+\lambda T)} du =$$

$$= \frac{\lambda^n(\alpha + \lambda T)^{\alpha+\sum_t n_t}}{\Gamma(\alpha + \sum_t n_t)n!} \int_0^{+\infty} u^{n+\alpha+\sum_t n_t-1} e^{-u(\alpha+\lambda T+\lambda)} du = \frac{\lambda^n(\alpha + \lambda T)^{\alpha+\sum_t n_t}}{\Gamma(\alpha + \sum_t n_t)n!} \frac{\Gamma(n + \alpha + \sum_t n_t)}{(\alpha + \lambda T + \lambda)^{n+\alpha+\sum_t n_t}} =$$

$$= \frac{\Gamma(n + \alpha + \sum_t n_t)}{\Gamma(\alpha + \sum_t n_t)n!} \left(\frac{\lambda}{\alpha + \lambda(T + 1)}\right)^n \left(\frac{\alpha + \lambda T}{\alpha + \lambda(T + 1)}\right)^{\alpha+\sum_t n_t} \Rightarrow N_{T+1}|H_T \sim BN\left(\alpha+\sum_t n_t, \frac{\alpha + \lambda T}{\alpha + \lambda(T + 1)}\right)$$

Come si assegnano però λ e α? Se non sono dati, è necessario introdurre i n.a. A e Λ e le ipotesi del processo si mantengono, ma condizionate ai parametri aleatori, ovvero assegnata la legge congiunta $F_{\Lambda,A}$, si ha $(N_1, N_2, \dots)|\Lambda = \lambda, A = \alpha \sim Poi - Gamma(\lambda, \alpha)$. Dovremmo quindi introdurre un modello gerarchico, attraverso cui si valutano probabilisticamente i parametri aleatori e condizionatamente a questi il processo stocastico. Operativamente si applica un approccio bayesiano empirico, nel quale (α, λ) sono considerati certi ma incogniti e vengono stimati dai dati, per cui si assume che il processo $\{N_1, N_2, \dots\}$ sia un processo $Poi - Gamma(\widehat{\alpha}, \widehat{\lambda})$.

Stima dei parametri con MVS

Supponiamo di aver osservato r rischi della collettività, ciascuno per T anni e indichiamo con n_{it} il numero osservato di sinistri per l'assicurato i-esimo nell'anno t. Siano perciò N_{it} i n.a. e n_{it} le loro realizzazioni e i processi $\{U_i, N_{i1}, N_{i2}, \dots\}$.

Siano le seguenti ipotesi:

∘ $\{U_i, N_{i1}, N_{i2}, \dots\}$ IID al variarre di i

∘ $N_{i1}|U_i = u, N_{i2}|U_i = u, \dots$ sono IID $\forall u$ determinazione di U_i

∘ $N_{it}|U_i = u \sim Poi(\lambda u)$

∘ $U_i \sim \Gamma(\alpha, \alpha)$

Sia $n_i = \sum_t n_{it}$ il numero osservato di sinistri per l'i-esimo assicurato e $n = \sum_i n_i$ il numero totale di sinistri della collettività. La verosimiglianza è quindi:

$$L(\lambda,\alpha,\underline{n}) \propto \prod_{i=1}^{r} Pr(N_{i1}=n_{i1},\ldots,N_{iT}=n_{iT};\lambda,\alpha) = \prod_{i=1}^{r} \frac{\Gamma(\alpha+n_i)}{\Gamma(\alpha)\prod_t(n_{it}!)}\left(\frac{\alpha}{\alpha+\lambda T}\right)^{\alpha}\left(\frac{\lambda}{\alpha+\lambda T}\right)^{n_i}$$

$$l(\lambda,\alpha,\underline{n}) = \sum_{i=1}^{r} \ln\left(\frac{\Gamma(\alpha+n_i)}{\Gamma(\alpha)\prod_t(n_{it}!)}\right) + \alpha[\ln(\alpha)-\ln(\alpha+\lambda T)] + n_i[\ln(\lambda)-\ln(\alpha+\lambda T)] =$$

$$= \sum_{i=1}^{r} \ln\left(\frac{\Gamma(\alpha+n_i)}{\Gamma(\alpha)}\right) - \sum_t \ln(n_{it}!) + \alpha[\ln(\alpha)-\ln(\alpha+\lambda T)] + n_i[\ln(\lambda)-\ln(\alpha+\lambda T)]$$

Si fanno quindi le derivate parziali rispetto a λ e α e si eguagliano a zero; tra le soluzioni si determinano quelle per cui l'*Hessiana* è definita negativa.

$$\frac{\delta l}{\delta \lambda} = \sum_{i=1}^{r}\left[\alpha\left(\frac{-1}{\alpha+\lambda T}\right)T\right] + n_i\left(\frac{1}{\lambda}-\frac{T}{\alpha+\lambda T}\right) = -\frac{\alpha r T}{\alpha+\lambda T} + \frac{n}{\lambda} - \frac{nT}{\alpha+\lambda T} = 0 \Leftrightarrow$$

$$\Leftrightarrow -\alpha T r + n(\alpha+\lambda T) - n\lambda T = 0 \Leftrightarrow Tr\lambda = n \quad \Rightarrow \quad \widehat{\lambda} = \frac{n}{rT}$$

Quindi in questo schema si riesce a dare in forma chiusa la distribuzione a posteriori del modello $Poi - Gamma$ con $U_i \sim \Gamma(\alpha,\alpha) \quad \Rightarrow \quad U_i|H_T \sim \Gamma(\alpha+n_i, \alpha+\lambda T)$.
Più in generale non è banale determinare la distribuzione a posteriori del processo $\{Y_1, Y_2, \ldots\}$, a causa della costante di normalizzazione, perciò vi sono tecniche numeriche iterative, come il metodo Monte Carlo. L'approccio bayesiano presenta quindi delle complessità di calcolo, richiede di assegnare oltre alla legge di U, anche la legge $\forall u$ a $Y_1|U=u, Y_2|U=u,\ldots$ ed inoltre porta a problemi di comunicazione a livello commerciale.

Teoria della Credibilità Bayesiana

É un approccio semplificato che si basa sulla teoria della credibilità di base empirica, in cui per l'assicurato i-esimo $P_i^{CRED}(H_{iT}) = P(1-z_T) + \overline{x}_{iT} z_T$, che è intuitivamente una combinazione lineare convessa del premio a priori P e della media campionaria $\overline{x}_{iT}$.
Seguendo un approccio razionale, è possibile arrivare ad una formula di questo tipo tramite la credibilità bayesiana. Si consideri una collettività di rischi analoghi ed indipendenti, fissato l'i-esimo rischio si ha il processo stocastico $\{Y_{i1}, Y_{i2},\ldots\}$, con eterogeneità residua modellata dal parametro aleatorio U_i, che descrive il profilo di rischiosità dell'assicurato. Siano le seguenti ipotesi:
∘ $\forall u$ si ha che $Y_{i1}|U_i=u, Y_{i2}|U_i=u,\ldots$ sono stocasticamente indipendenti
∘ $\{U_i, Y_{i1}, Y_{i2},\ldots\}$ stocasticamente indipendenti al variare di i
Nello schema bayesiano la seconda ipotesi considera i processi anche identicamente distribuiti al variare di i, mentre qui solo indipendenti. Inoltre, sempre nello schema bayesiano, dalla grandezza $\mathbb{E}(Y_{i,T+1}|U_i)$ si ottiene lo stimatore del premio individuale in $T+1$ $\mathbb{E}(Y_{i,T+1}|Y_1,\ldots,Y_T)$. In questo caso si considera ancora una volta una funzione del processo di osservazione, ma questa volta una funzione più semplice, la funzione lineare $g(Y_{i1},\ldots Y_{iT}) = \alpha_{i0} + \sum_t \alpha_{it} y_{it}$.
Tra tutte le funzioni lineari, si seleziona $\underline{\alpha_i}$ tale che sia minima la funzione di perdita quadratica attesa:

$$\min_{\alpha_{0i},\ldots,\alpha_{Ti}} \mathbb{E}\left[\left(\mathbb{E}(Y_{i,T+1}|U_i) - (\alpha_{i0} + \sum_{t=1}^{T} \alpha_{it} y_{it})\right)^2\right]$$

Fissato un individuo i-esimo (d'ora in poi i verrà omesso), consideriamo $\{Y_1, Y_2,\ldots\}$ e U:
∘ $\forall u$ si ha che $Y_1|U=u, Y_2|U=u,\ldots$ sono stocasticamente indipendenti

Quindi:

$$\min_{\underline{\alpha}} Q(\alpha_0,\dots,\alpha_T) = \min_{\underline{\alpha}} \mathbb{E}\left[\left(\mathbb{E}(Y_{T+1}|U) - (\alpha_0 + \sum_{t=1}^{T}\alpha_t y_t)\right)^2\right] = \min_{\underline{\alpha}} \mathbb{E}\left[\left(\mu_{T+1}(U) - (\alpha_0 + \sum_{t=1}^{T}\alpha_t y_t)\right)^2\right]$$

$$\text{con} \qquad Q(\alpha_0,\dots,\alpha_T) = \int \left(\mu_{T+1}(u) - (\alpha_0 + \sum_t \alpha_t y_t)\right)^2 dF_{U,Y_1,\dots,Y_T}(u,y_1,\dots,y_T)$$

Supponiamo soddisfatte le condizioni di regolarità che permettono di derivare sotto il segno di integrale, allora:

$$\frac{\delta Q}{\delta \alpha_0} = \int -2\left(\mu_{T+1}(u) - \alpha_0 - \sum_t \alpha_t y_t\right) dF_{U,Y_1,\dots,Y_T}(u,y_1,\dots,y_T) = -2\mathbb{E}\left(\mu_{T+1}(U) - \alpha_0 - \sum_t \alpha_t Y_t\right)$$

$$\frac{\delta Q}{\delta \alpha_0} = 0 \Leftrightarrow \mathbb{E}(\mu_{T+1}(U)) - \alpha_0 - \sum_t \alpha_t \mathbb{E}(Y_t) = 0 \Leftrightarrow \mathbb{E}(\mathbb{E}(Y_{T+1}|U)) - \alpha_0 - \sum_t \alpha_t \mathbb{E}(Y_t) = 0 \Leftrightarrow$$

$$\Leftrightarrow \mathbb{E}(Y_{T+1}) = \alpha_0 + \sum_t \alpha_t \mathbb{E}(Y_t)$$

$$\frac{\delta Q}{\delta \alpha_h} = 0 \Leftrightarrow \int -2y_h\left(\mu_{T+1}(u) - \alpha_0 - \sum_t \alpha_t y_t\right) dF_{U,Y_1,\dots,Y_T}(u,y_1,\dots,y_T) = 0 \Leftrightarrow$$

$$\Leftrightarrow \underbrace{\mathbb{E}(\mu_{T+1}(U)Y_h)}_{1} - \alpha_0\mathbb{E}(Y_h) - \sum_t \alpha_t \mathbb{E}(Y_t Y_h) = 0 \Leftrightarrow \mathbb{E}(Y_{T+1}Y_h) = \alpha_0 \mathbb{E}(Y_h) + \sum_t \alpha_t \mathbb{E}(Y_t Y_h)$$

$$1 = \mathbb{E}(\mathbb{E}(\mu_{T+1}(U)Y_h|U)) = \mathbb{E}(\mu_{T+1}(U)\mathbb{E}(Y_h|U)) = \mathbb{E}(\mathbb{E}(Y_{T+1}|U)\mathbb{E}(Y_h|U)) = \mathbb{E}(\mathbb{E}(Y_{T+1}Y_h|U)) = \mathbb{E}(Y_{T+1}Y_h)$$

Quindi otteniamo il così detto sistema normale:

$$\begin{cases} \mathbb{E}(Y_{T+1}) = \alpha_0 + \sum_t \alpha_t \mathbb{E}(Y_t) \Leftrightarrow -\mathbb{E}(Y_h)\mathbb{E}(Y_{T+1}) = -\alpha_0\mathbb{E}(Y_h) - \sum_t \alpha_t \mathbb{E}(Y_t)\mathbb{E}(Y_h) \\ \mathbb{E}(Y_{T+1}Y_h) = \alpha_0\mathbb{E}(Y_h) + \sum_t \alpha_t \mathbb{E}(Y_t Y_h) \end{cases}$$

E sommando membro a membro otteniamo quindi il sistema equivalente:

$$\begin{cases} \mathbb{E}(Y_{T+1}) = \alpha_0 + \sum_t \alpha_t \mathbb{E}(Y_t) \\ \mathbb{E}(Y_{T+1}Y_h) - \mathbb{E}(Y_{T+1})\mathbb{E}(Y_h) = \sum_t \alpha_t \mathbb{E}(Y_t Y_h) - \mathbb{E}(Y_t)\mathbb{E}(Y_h) \Leftrightarrow Cov(Y_{T+1}, Y_h) = \sum_{t=1}^T \alpha_t Cov(Y_t, Y_h) \end{cases}$$

Quindi dal sistema è possibile ricavare i coefficienti, essendo composto da $T+1$ equazioni con $T+1$ incognite. Si prova che ammette un'unica soluzione ed è proprio il punto di minimo, ottenendo cosi $\alpha_o^*, \dots, \alpha_T^*$ e lo stimatore lineare di credibilità del premio individuale in $T+1$:

$$\overset{\sim}{\mu}{}^{CRED}_{T+1}(U) = \alpha_0^* + \sum_{t=1}^{T} \alpha_t^* Y_t$$

3.4 24/10/2019

<u>Osservazione 1:</u> il premio individuale in $T+1$ è $\mu_{T+1}(U) = \mathbb{E}(Y_{T+1}|U)$ e se conoscessimo il valore del parametro aleatorio, giudicheremmo i numeri stocasticamente indipendenti e sarebbe ininfluente la storia passata H_T.

Osservazione 2: data la storia $H_T = (Y_1 = y_1, \dots, Y_T = y_T)$ riusciamo a calcolare facilmente la stima di $P^{CRED}_{T+1}(H_T) = \alpha_0^* + \sum_t \alpha_t^* y_t$, a differenza dell'approccio bayesiano, in cui è necessario valutare la speranza matematica a posteriori. Per i coefficienti è necessario conoscere solo i primi due momenti dei n.a. del processo $\{Y_1, Y_2, \dots\}$, perciò non si necessita della specificazione dell'approccio bayesiano. Tuttavia ci accontentiamo di una funzione lineare per l'approssimazione, quindi si tratta di un minimo locale e non globale come l'approccio bayesiano.

Si noti che è uno stimatore non distorto:

$$\mathbb{E}(\overset{\sim}{\mu}{}^{CRED}_{T+1}(U)) = \alpha_0^* + \sum_{t=1}^{T} \alpha_t^* \mathbb{E}(Y_t) = \mathbb{E}(Y_{T+1}) = \mathbb{E}[\mathbb{E}(Y_{T+1}|U)] = \mathbb{E}(\mu_{T+1}(U))$$

Con la seconda uguaglianza data dalla prima equazione del sistema normale.

Osservazione 3: consideriamo la perdita quadratica attesa dallo stimatore bayesiano in Q_1 e dalla grandezza di interesse in Q_2:

$$Q_1(\alpha_0, \dots, \alpha_T) = \mathbb{E}\left[\left(\mathbb{E}(Y_{T+1}|Y_1, \dots, Y_T) - \alpha_0 - \sum_{t=1}^{T} \alpha_t Y_t\right)^2\right]$$

$$Q_2(\alpha_0, \dots, \alpha_T) = \mathbb{E}\left[\left(Y_{T+1} - \alpha_0 - \sum_{t=1}^{T} \alpha_t Y_t\right)^2\right]$$

Se cerchiamo gli $\alpha_0^* \dots, \alpha_T^*$ che minimizzano prima Q_1 e poi Q_2, otteniamo gli stessi coefficienti del sistema normale. Quindi lo stimatore di credibilità è la funzione lineare affine che approssima al meglio il premio individuale, il premio bayesiano ed anche la grandezza di interesse.

Osservazione 4: il sistema normale dipende da i (lo avevamo omesso) e di conseguenza i coefficienti dipendono dall'individuo. Quindi troviamo $\overset{\sim}{\mu}{}^{CRED}_{T+1}(U_i) = \alpha_{i0}^* + \sum_t \alpha_{it}^* y_{it}$. Ma il premio è tale che esso non dipende da i, nel caso in cui i momenti del sistema non dipendano dall'individuo, e in tal caso gli $\underline{\alpha}$ saranno i medesimi per tutti gli individui. Dipende quindi dalla condizione di analogia.

1 **Modello Bühlmann**

Consideriamo una collettività di rischi "analoghi" ed "indipendenti", fissato l'individuo i-esimo e consideriamo il processo $\{U_i, Y_{i1}, Y_{i2}, \dots\}$ con ipotesi probabilistiche:
- $\{U_i, Y_{i1}, Y_{i2}, \dots\}$ sono IID al variare di i
- Fissato i arbitrariamente, $\forall u$ determinazione di U_i, $Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots$ sono numeri aleatori IID

Siano $\mathbb{E}(Y_{it}|U_i = u) = \mu(u)$ e $V(Y_{it}|U_i = u) = v(u)$ ed entrambi esistano finiti. Non dipendono da t per l'ipotesi di identica distribuzione del secondo punto, mentre non compare i per l'uguale distribuzione tra i processi, come nella prima ipotesi. Di conseguenza si hanno ipotesi più stringenti rispetto al modello generale per l'aggiunta dell'identica distribuzione.
Sia $\mathbb{E}(Y_{i,T+1}|U_i) = \mu(U_i)$ il *premio individuale* in $T+1$. Sia $\mu = \mathbb{E}(\mu(U_i)) = \mathbb{E}(\mathbb{E}(Y_{i,T+1}|U_i)) = \mathbb{E}(Y_{i,T+1})$ la speranza matematica del premio individuale, non dipende da i, poiché gli U_i sono ID, ed è la comune speranza matematica della grandezza di interesse, detta premio collettivo in $T+1$ (non dipende dal tempo, quindi vale $\forall T+1$).

Poniamo $v = \mathbb{E}(v(U_i)) = \mathbb{E}(V(Y_{it}|U_i))$ e $a = V(\mu(U_i)) = V(\mathbb{E}(Y_{i,T+1}|U_i))$. Questi tre parametri non dipendono da i e sono detti *parametri di struttura del modello.* Si noti che $V(Y_{it}) = \mathbb{E}(V(Y_{it}|U_i)) + V(\mathbb{E}(Y_{it}|U_i)) = v + a$. Essi intervengono nella scomposizione della varianza, che è dovuta all'aleatorietà intrinseca dei numeri aleatori per v e dalla componente di eterogeneità residua della collettività, descritta dal parametro a, che è la varianza del premio individuale.
Inoltre i momenti del sistema normale dipendono proprio da questi parametri:
∘ $\mathbb{E}(Y_{it}) = \mu$, $\forall i, t = 1, \dots T+1$
∘ Se $t = h$, $Cov(Y_{it}, Y_{ih}) = V(Y_{it}) = v + a$;
se $t \neq h$, $Cov(Y_{it}, Y_{ih}) = \mathbb{E}(Y_{it}Y_{ih}) - \mathbb{E}(Y_{it})\mathbb{E}(Y_{ih}) = \mathbb{E}(Y_{it}Y_{ih}) - \mathbb{E}(\mu(U_i))^2 = \mathbb{E}(\mu(U_i)^2) - \mathbb{E}(\mu(U_i))^2 =$
$= V(\mu(U_i)) = a$
Dato che $\mathbb{E}(Y_{it}Y_{ih}) = \mathbb{E}(\mathbb{E}(Y_{it}Y_{ih}|U_i)) = \mathbb{E}(\mathbb{E}(Y_{it}|U_i)\mathbb{E}(Y_{ih}|U_i)) = \mathbb{E}(\mu(U_i)^2)$.

Quindi nel sistema normale, possiamo omettere i per l'identica distribuzione dei processi $\underline{\alpha}$ non dipende dal soggetto.

$$\begin{cases} \mu = \alpha_0 + \sum_{t=1}^T \alpha_t \mu \\ a = \alpha_h(v+a) + \sum_{t\neq h} \alpha_t a \end{cases} \Leftrightarrow \begin{cases} \sum_{t=1}^T \alpha_t = 1 - \frac{\alpha_0}{\mu} \\ a = \alpha_h v + a\sum_{t=1}^T \alpha_t \end{cases} \Leftrightarrow \begin{cases} \sum_{t=1}^T \alpha_t = 1 - \frac{\alpha_0}{\mu} \\ 1 = \alpha_h k + \sum_{t=1}^T \alpha_t \end{cases} \Leftrightarrow \text{ con } k \triangleq \frac{v}{a} > 0$$

$$\Leftrightarrow \begin{cases} \sum_{t=1}^T \alpha_t = 1 - \frac{\alpha_0}{\mu} \\ 1 = \alpha_h k + 1 - \frac{\alpha_0}{\mu} \end{cases} \Leftrightarrow \begin{cases} \sum_{t=1}^T \alpha_t = 1 - \frac{\alpha_0}{\mu} \\ \alpha_h = \frac{\alpha_0}{\mu k} \end{cases} \Rightarrow \text{ dato } \sum_{t=1}^T \alpha_t = \frac{T\alpha_0}{\mu k} \text{ si ha } \begin{cases} \frac{T\alpha_0}{\mu k} = 1 - \frac{\alpha_0}{\mu} \\ \alpha_h = \frac{\alpha_0}{\mu k} \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} \alpha_0\left(\frac{T}{\mu k} + \frac{k}{\mu k}\right) = 1 \\ \alpha_h = \frac{\mu k}{T+k}\frac{1}{\mu k} \end{cases} \Rightarrow \begin{cases} \alpha_0 = \frac{\mu k}{T+k} \\ \alpha_h = \frac{1}{T+k} \end{cases}$$

Secondo questo modello, lo stimatore di credibilità diventa:

$$\underset{\sim}{\mu}_{T+1}^{CRED}(H_T) = \mu\frac{k}{T+k} + \sum_{t=1}^T \frac{1}{T+k}Y_{it} = \mu\frac{k}{T+k} + \frac{T}{T+k}\sum_{t=1}^T \frac{Y_{it}}{T} = \mu(1 - z_T) + \overline{Y}_{iT} z_T$$

Si ha quindi una combinazione lineare convessa, con pesi non negativi e di somma pari a 1, tra il premio collettivo a priori e la sintesi della storia di sinistrosità dell'individuo. Si combinano quindi le due misure, che se no porterebbero ad un fenomeno di antiselezione, nel caso di μ elevato, o a premio nullo, qualora l'individuo non riporti sinistri.

Osservazione: il parametro $\alpha_t^* = \frac{1}{T+k}$ con $t = 1, \dots, T$, non dipende da t.
Per le ipotesi, $\forall u$ il processo è tale che $Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots$ sono IID e si può dimostrare che questo implica che il processo $\{Y_1, Y_2, \dots\}$ è scambiabile. Inoltre $\min_{\underline{\alpha}} \mathbb{E}[(Y_{i,T+1} - \alpha_0 - \sum_t Y_{it})^2]$ dipende dalla distribuzione del vettore aleatorio $(Y_{i1}, \dots, Y_{iT+1}) \overset{d}{=} (Y_{ij_1}, \dots, Y_{ij_{T+1}})$, per la scambiabilità ed affinché la speranza matematica sia la stessa, i parametri α_t devono essere uguali.

Stima dei parametri di struttura

Supponiamo di aver osservato r individui, ciascuno per T anni e che sia y_{it} con $i = 1, \dots, r$ e $t = 1, \dots, T$ i valori osservati della grandezza di interesse. Siano le seguenti ipotesi per il processo $\{U_i, Y_{i1}, Y_{i2}, \dots\}$:
∘ $\{U_i, Y_{i1}, Y_{i2}, \dots\}$ sono IID al variare di i

∘ Fissato i arbitrariamente, $\forall u$ determinazione di U_i, $Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots$ sono numeri aleatori IID
∘ Ed inoltre si ha che $\mathbb{E}(Y_{it}|U_i = u) = \mu(u)$ e $V(Y_{it}|U_i = u) = v(u)$

Date queste condizioni, il premio di credibilità $\widetilde{\mu}^{CRED}_{T+1}(U_i)$ dipende dai parametri di struttura μ e k, dove $\mu = \mathbb{E}(\mu(U_i))$, $v = \mathbb{E}(V(U_i))$ e $a = V(\mu(U_i))$.
L'obiettivo è quindi quello di trovare degli stimatori non distorti per i tre parametri di struttura.

μ É una "media" di "medie" a partire da quelle individuali $\overline{y}_{iT} = \frac{1}{T}\sum_t y_{it} \triangleq \widetilde{\mu}_i$. Verifichiamo infatti che lo stimatore $\widetilde{\mu} = \frac{1}{r}\sum_i \widetilde{\mu}_i$ sia corretto:

$$\mathbb{E}(\widetilde{\mu}) = \mathbb{E}(\frac{1}{r}\sum_i \widetilde{\mu}_i) = \frac{1}{r}\sum_i \mathbb{E}(\widetilde{\mu}_i) = \mathbb{E}(\widetilde{\mu}_i) = \frac{1}{T}\sum_t \mathbb{E}(Y_{it}) = \mu$$

v É il valore atteso di una varianza individuale, data infatti la varianza individuale $\widetilde{\sigma}^2_i = \frac{1}{T-1}(Y_{it} - \overline{Y}_{it})^2$, lo stimatore per v è $\widetilde{v} = \frac{1}{r}\sum_i \widetilde{\sigma}^2_i$. Verifichiamo adesso che $\widetilde{v}$ sia corretto. I $Y_{i1}, Y_{i2}, \dots$ non sono stocasticamente indipendenti, mentre fissato $U_i = u$ si ha che $Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots$ sono IID e $V(Y_{it}|U_i = u) = v(u)$.

$$\frac{1}{T-1}\sum_t (Y_{it}|U_i = u - \overline{Y}_{iT}|U_i = u)^2 = \frac{1}{T-1}\sum_t (Y_{it} - \overline{Y}_{it})^2 |U_i = u$$

$$\mathbb{E}\Bigg(\underbrace{\frac{1}{T-1}\sum_t (Y_{it} - \overline{Y}_{it})^2 |U_i = u}_{\widetilde{\sigma}^2_i}\Bigg) = v(u) \Rightarrow \mathbb{E}(\widetilde{\sigma}^2_i) = \mathbb{E}(\mathbb{E}(\widetilde{\sigma}^2_i|U_i = u)) = \mathbb{E}(v(U)) = v$$

$$\mathbb{E}(\widetilde{v}) = \mathbb{E}\left(\frac{1}{r}\sum_i \widetilde{\sigma}^2_i\right) = \frac{1}{r}\sum_i \mathbb{E}(\widetilde{\sigma}^2_i) = v$$

Quindi $\widetilde{v}$ è uno stimatore corretto di v.

a É una varianza delle medie campionarie $\frac{1}{r-1}\sum_i(\overline{Y}_{iT} - \widetilde{\mu})^2$. Siano $\overline{Y}_{1T}, \dots, \overline{Y}_{rT}$ le medie campionarie per ogni individuo. Per ipotesi i processi tra gli individui sono stocasticamente indipendenti ed in particolare ciascuno di questi n.a. è funzione di un vettore di numeri aleatori:$\overline{Y}_{iT} = f(Y_{i1}, \dots, Y_{iT})$. Essendo l'insieme essi delle T-uple stocasticamente indipendenti, anche le loro funzioni saranno stocasticamente indipendenti. Allo stesso modo essendo le T-uple identicamente distribuite, anche $\overline{Y}_{1T}, \dots, \overline{Y}_{rT}$ saranno identicamente distribuiti. Quindi abbiamo che :

$$V(\overline{Y}_{iT}) = \mathbb{E}(V(\overline{Y}_{iT}|U_i)) + V(\mathbb{E}(\overline{Y}_{iT}|U_i)) = \mathbb{E}\left[V\left(\frac{1}{T}\sum_t Y_{it}|U_i\right)\right] + V\left[\mathbb{E}\left(\frac{1}{T}\sum_t Y_{it}|U_i\right)\right] =$$

$$= \frac{1}{T^2}\mathbb{E}\left[\sum_t V\Big(Y_{it}|U_i\Big)\right] + \frac{1}{T^2}V\left[\sum_t \mathbb{E}\Big(Y_{it}|U_i\Big)\right] = \frac{1}{T^2}\mathbb{E}(Tv(U_i)) + \frac{1}{T^2}V(T\mu(U_i)) = \frac{1}{T}v + a$$

Da questo troviamo che $\widetilde{a} = \frac{1}{r-1}\sum_i(\overline{Y}_{iT} - \widetilde{\mu})^2 - \frac{1}{T}v$, che è corretto, infatti:

$$\mathbb{E}(\widetilde{a}) = \frac{1}{T}v + a - \frac{1}{T}v = a$$

3.5 25/10/2019

Date le osservazioni, è possibile calcolare:
◦ $\widehat{\mu}_i = \frac{1}{T}\sum_t y_{it}$ e $\widehat{\mu} = \frac{1}{r}\sum_i \widehat{\mu}_i$
◦ $\widehat{\sigma}_i^2 = \frac{1}{T-1}\sum_t (y_{it} - \widehat{\mu}_i)^2$ e $\widehat{v} = \frac{1}{r}\sum_i \widehat{\sigma}_i^2$
◦ $\widehat{a} = \frac{1}{r-1}\sum_i (\widehat{\mu}_i - \widehat{\mu})^2 - \frac{1}{r}\widehat{v}$
Può accadere talvolta che $\widehat{a} < 0$, dovuto al fatto che sia stato scelto un modello con sovradispersione, ma che questa non emerga dai dati. Ossia non si ha una sufficiente eterogeneità e quindi non c'è variabilità tra i premi individuali.

Consideriamo una collettività di rischi "analoghi" ed "indipendenti", fissiamo il rischio i-esimo e siamo interessati al processo di arrivo dei rischi per l'assicurato i-esimo $\{N_{i1}, N_{i2}, \dots\}$. Introduciamo il parametro U_i di eterogeneità residua e in virtù dell'analogia, accogliamo le seguenti ipotesi:
◦ $\{U_i, N_{i1}, N_{i2}, \dots\}$ sono IID
◦ fissato l'i-esimo rischio:
- $\forall u$ determinazione di U_i si ha che $N_{i1}|U_i = u, N_{i2}|U_i = u, \dots$ sono IID
- $\forall u$ determinazione di U_i si ha che $N_{it}|U_i = u \sim Poi(\lambda u)$ con $\lambda > 0$
- $\mathbb{E}(U_i) = 1$ e $V(U_i) = \sigma^2$

Il modello specificato è un modello semiparametrico detto modello *Poisson mistura semiparametrico* ed è analogo al modello Poisson-Gamma nell'approccio bayesiano, ma senza la distribuzione del parametro aleatorio. Per questo motivo non è possibile applicare l'approccio bayesiano, ma si può considerare il modello della credibilità di Bühlmann, con specificata la distribuzione condizionata.
Si considerino i parametri di struttura:
◦ $\mu = \mathbb{E}(\mu(U_i)) = \mathbb{E}(\mathbb{E}(N_{it}|U_i)) = \mathbb{E}(N_{it}) = \lambda$ che è il premio collettivo a priori
◦ $v = \mathbb{E}(v(U_i)) = \mathbb{E}(V(N_{it}|U_i)) = \mathbb{E}(\lambda U_i) = \lambda$
◦ $a = V(\mu(U_i)) = V(\mathbb{E}(N_{it}|U_i)) = \lambda^2 V(U_i) = \lambda^2\sigma^2$
E dati $z_T = \frac{T}{T+k}$ e $k = \frac{v}{a} = \frac{\lambda}{\lambda^2\sigma^2} = \frac{1}{\lambda\sigma^2}$ si ha:

$$\overset{\sim}{\mu}^{CRED}_{T+1}(H_T) = \lambda(1 - z_T) + \frac{1}{T}\sum_{t=1}^{T} N_{it} z_T$$

In questo modo è possibile stimare λ e σ^2 tramite $\widetilde{a}$ e $\widetilde{v}$.

<u>Problemi modello di Bühlmann</u>

Abbiamo osservato che il processo delle grandezze di interesse $\{Y_{i1}, Y_{i2}, \dots\}$ è un processo scambiabile, di conseguenza i n.a. sono identicamente distribuiti:

$$F_{Y_{it}}(y) = Pr(Y_{it} \leq y) = \int Pr(Y_{it} \leq y | U_i = u) dF_{U_i}(u) \quad \Rightarrow \quad F_{Y_{it}} \perp t \text{ e } Y_{i1}, Y_{i2}, \dots \text{ sono ID}$$

Tuttavia però non osserviamo i n.a. per l'intera durata, ma per un tempo di esposizione indicato con exp_{it}. Quindi il processo di osservazione non è N_{it}, ma M_{it}, ovvero il n.a. di sinistri per l'assicurato i-esimo nel tempo di esposizione, e per questo motivo non è possibile considerarli ID. Se però consideriamo:
◦ $\mathbb{E}(\frac{M_{it}}{exp_{it}})$ non dipende da t in termini attesi, ma $V(\frac{M_{it}}{exp_{it}})$ dipende da t
◦ $\mathbb{E}(\frac{X'_{it}}{exp_{it}})$ il risarcimento per unità di esposizione non dipende da t, ma $V(\frac{X'_{it}}{exp_{it}})$ dipende da t
◦ $\mathbb{E}(\frac{X_{it}}{w_{it}})$ il tasso di premio per unità di esposizione monetaria w_{it} è costante per t, ma $V(\frac{X_{it}}{w_{it}})$ dipende

da t
∘ $\mathbb{E}(\frac{X_{it}}{P_{it}})$ il risarcimento per unità di premio per una polizza collettiva con montepremi variabile P_{it} non dipende da t, ma $V(\frac{X_{it}}{P_{it}})$ dipende da t
Quindi in questi casi non è ragionevole considerare il modello di Bühlmann a causa della varianza.

2 **Modello di Bühlmann-Strauß**
Siamo del quadro della teoria della credibilità, consideriamo una collettività di rischi "analoghi" ed "indipendenti". Fissato l'i-esimo assicurato e considerato il processo di interesse $\{Y_{i1}, Y_{i2}, \dots\}$, modelliamo l'eterogeneità residua con U_i ed accogliamo le seguenti ipotesi:
∘ $\{U_i, Y_{i1}, Y_{i2}, \dots\}$ stocasticamente indipendenti al variare di i
∘ U_i identicamente distribuiti
∘ $\forall u$ determinazione di U_i si ha che:
- $Y_{i1}, |U_i = u, Y_{i2}|U_i = u, \dots$ sono stocasticamente indipendenti
- $\mathbb{E}(Y_{it}|U_i = u) = \mu(u)$ costante
- $V(Y_{it}|U_i = u) = \frac{v(u)}{m_{it}}$ dove $v(\cdot)$ non dipende da i e t e dove m_{it} è una misura data del volume dei premi

Il premio individuale nell'anno $T+1$ è $\mathbb{E}(Y_{i,T+1}|U_i) = \mu(U_i)$ e poniamo $\mu = \mathbb{E}(\mu(U_i)) = \mathbb{E}(Y_{it})$ il premio collettivo, $\frac{v}{m_{it}} = \mathbb{E}(v(U_i)) = \mathbb{E}(V(Y_{it}|U_i))$ e $a = V(\mu(U_i)) = V(\mathbb{E}(Y_{it}|U_i))$. La varianza di Y_{it} è:

$$V(Y_{it}) = \mathbb{E}(V(Y_{it}|U_i)) + V(\mathbb{E}(Y_{it}|U_i)) = \frac{v}{m_{it}} + a$$

Anche in questo caso le soluzioni del sistema normale dipendono dal parametro di struttura attraverso cui si esprimono i momenti primi:

$$\begin{cases} \mathbb{E}(Y_{i,T+1}) = \alpha_{i0} + \sum_{t=1}^T \alpha_{it}\mathbb{E}(Y_{it}) \\ Cov(Y_{i,T+1}, Y_{ih}) = \sum_{t=1}^T \alpha_{it} Cov(Y_{it}, Y_{ih}) \end{cases} \qquad \text{con } h = 1, \dots, T$$

Grazie all'indipendenza, è possibile studiare i coefficienti per ogni individuo singolarmente, che saranno però diversi da soggetto a soggetto, non essendo ID.
dato quindi che $\mathbb{E}(Y_{it}) = \mu$ e:

$$Cov(Y_{it}, Y_{ih}) = \begin{cases} V(Y_{it}) = \frac{v}{m_{it}} + a & t = h \\ V(\mu(U_i)) = a & t \neq h \end{cases}$$

$$\underbrace{Cov(Y_{it}, Y_{ih})}_{t\neq h} = \mathbb{E}(Y_{it}Y_{ih}) - \mathbb{E}(Y_{it})\mathbb{E}(Y_{ih}) = \mathbb{E}(Y_{it}Y_{ih}) - \mathbb{E}(\mu(U_i))^2 = \mathbb{E}(\mathbb{E}(Y_{it}Y_{ih}|U_i)) - \mathbb{E}(\mu(U_i))^2 =$$

$$= \mathbb{E}(\mathbb{E}(Y_{it}|U_i)\mathbb{E}(Y_{ih}|U_i)) - \mathbb{E}(\mu(U_i))^2 = \mathbb{E}(\mu(U_i)^2) - \mathbb{E}(\mu(U_i))^2 = V(\mu(U_i)) = a$$

Omettiamo i e quindi il sistema normale diventa:

$$\begin{cases} \mu = \alpha_0 + \sum_t \alpha_t \mu \\ a = \alpha_h(\frac{v}{m_h} + a) + \sum_t \alpha_t a \end{cases} \Leftrightarrow \begin{cases} \sum_t \alpha_t = 1 - \frac{\alpha_0}{\mu} \\ a = \frac{\alpha_h v}{m_h} + a\sum_t \alpha_t \end{cases} \Leftrightarrow \begin{cases} \sum_t \alpha_t = 1 - \frac{\alpha_0}{\mu} \\ 1 = \frac{\alpha_h k}{m_h} + \sum_t \alpha_t \end{cases} \Leftrightarrow$$

$$\begin{cases} \sum_t \alpha_t = 1 - \frac{\alpha_0}{\mu} \\ 1 = \frac{\alpha_h k}{m_h} + 1 - \frac{\alpha_0}{\mu} \end{cases} \Leftrightarrow \begin{cases} \sum_t \alpha_t = 1 - \frac{\alpha_0}{\mu} \\ \alpha_h = \frac{\alpha_0 m_h}{\mu k} \end{cases} \Leftrightarrow \begin{cases} \frac{\alpha_0}{\mu k} \sum_t m_t = 1 - \frac{\alpha_0}{\mu} \\ \alpha_h = \frac{\alpha_0 m_h}{\mu k} \end{cases} \Leftrightarrow \begin{cases} \alpha_0 = \frac{\mu k}{k + \sum_t m_t} \\ \alpha_h = \frac{\alpha_0 m_h}{\mu k} = \frac{m_h}{k+\sum_t m_t} \end{cases}$$

Da cui otteniamo il premio individuale:

$$\widetilde{\mu}^{CRED}_{T+1}(U_i) = \alpha^*_{i0} + \sum_t \alpha^*_{it} Y_{it} = \mu \frac{k}{k+m_{i\cdot}} + \frac{m_{i\cdot}}{m_{i\cdot}+k} \frac{\sum_t y_{it} m_{it}}{m_{i\cdot}} = \mu \cdot \underbrace{\frac{k}{k+m_{i\cdot}}}_{1-z_{iT}} + \overline{\overline{y}}_{it} \cdot \underbrace{\frac{m_{i\cdot}}{m_{i\cdot}+k}}_{z_{iT}}$$

Rispetto al modello di Bühlmann i pesi dipendono da i e i fattori dipendono in particolare da $m_{i\cdot}$ e k.

Stima dei parametri mediante stimatori non distorti

(μ) Gli stimatori corretti sono:

$$\widetilde{\mu}_i = \frac{1}{T} \sum_{t=1}^{T} \frac{m_{it}}{m_{i\cdot}} y_{it} = \overline{\overline{y}}_{iT} \qquad \widetilde{\mu} = \frac{1}{m_{\cdot\cdot}} \sum_{i=1}^{r} \overline{\overline{y}}_{iT} \qquad \text{Con } m_{\cdot\cdot} = \textstyle\sum_i m_{i\cdot} = \sum_i \sum_t m_{it}$$

(v) Lo stimatore corretto è:

$$\widetilde{v} = \frac{1}{r} \sum_i \left[\frac{1}{T-1} \sum_t m_{it} (y_{it} - \widetilde{\mu}_i)^2 \right] = \frac{1}{r} \sum_i \widetilde{\sigma}^2_i$$

(a) Lo stimatore corretto è:

$$\widetilde{a} = \frac{m_{\cdot\cdot}}{m^2_{\cdot\cdot} - \sum_i m^2_{i\cdot}} \left[\sum_{i=1}^{r} m_{i\cdot} (\widetilde{\mu}_i - \widetilde{\mu})^2 - (r-1) v \right]$$

3.6 28/10/2019

Esempio di applicazione del modello di Bülmann-Strauß

Sia una collettività di rischi "analoghi" ed "indipendenti". Consideriamo che sia fissato l'individuo i-esimo e che il processo di interesse sia il processo di arrivo dei sinistri annuo $\{N_{i1}, N_{i2}, \dots\}$, con eterogeneità modellata dal parametro aleatorio U_i, con le seguenti ipotesi: ∘ $\{U_i, Y_{i1}, Y_{i2}, \dots\}$ stocasticamente indipendenti al variare di i

∘ U_i identicamente distribuiti

∘ $\forall u$ determinazione di U_i si ha che:

- $N_{i1}, |U_i = u, N_{i2}|U_i = u, \dots$ sono stocasticamente indipendenti
- $N_{it}|U_i = u \sim Poi(\lambda_{it} u)$
- $\mathbb{E}(U_i) = 1$, $V(U_i) = \sigma^2$

Date queste ipotesi, il premio individuale in t è: $\mathbb{E}(N_{it}|U_i) = \lambda_{it} U_i$, e $\mathbb{E}(N_{it}) = \mathbb{E}(\mathbb{E}(N_{it}|U_i)) = \lambda_{it}$. Quindi il modello consente di introdurre l'ipotesi di non stazionarietà in termini attesi e di considerare l'evoluzione nel tempo degli assicurati nelle classi tariffarie, dovuta alle modifiche delle variazioni tariffarie $X_1, \dots, X_n$.

Il modello così specificato, non rientra nè nel quadro bayesiano, nè nel quadro di Bühlmann o di Bühlmann-Strauß, poiché $\mu(u) = \mathbb{E}(N_{it}|U_i = u) = \lambda_{it} u$ dipende dall'individuo.

Per correggere la specificazione, consideriamo il processo delle variabili standardizzate $Y_{it} = \frac{N_{it}}{\lambda_{it}}$ e si hanno le seguenti ipotesi:

∘ $\{U_i, Y_{i1}, Y_{i2}, \dots\}$ sono stocasticamente indipendenti al variare di i

∘ Gli U_i sono identicamente distribuiti

∘ Fissato l'individuo i-esimo:

- $\forall u$ determinazione di U_i si ha che $Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots$ sono stocasticamente indipendenti, quindi si ha, con $\mu(\cdot)$ e $v(\cdot)$ funzione identica:

$$\mathbb{E}(Y_{it}|U_i = u) = \mathbb{E}\left(\frac{N_{it}}{\lambda_{it}}\middle| U_i = u\right) = \frac{1}{\lambda_{it}}\mathbb{E}(N_{it}|U_i = u) = u = \mu(u)$$

$$V(Y_{it}|U_i = u) = V\left(\frac{N_{it}}{\lambda_{it}}\middle| U_i = u\right) = \frac{1}{\lambda_{it}^2}V(N_{it}|U_i = u) = \frac{\lambda_{it}u}{\lambda_{it}^2} = \frac{u}{\lambda_{it}} = \frac{v(u)}{m_{it}}$$

Quindi i parametri di struttura e il premio di credibilità risultano:

$$\mu = \mathbb{E}(\mu(U_i)) = \mathbb{E}(U_i) = 1 \qquad v = \mathbb{E}(v(U_i)) = \mathbb{E}(U_i) = 1 \qquad a = V(\mu(U_i)) = V(U_i) = \sigma^2$$

$$\overset{\sim CRED}{\mu_{T+1}}(U_i) = \mu(z_{iT}) + \overline{\overline{y}}_{iT} z_{iT} = \frac{\frac{1}{\sigma^2}}{\lambda_{i\cdot} + \frac{1}{\sigma^2}} + \sum_{t=1}^{T} \frac{N_{it}}{\lambda_{i\cdot} + \frac{1}{\sigma^2}} = \frac{\frac{1}{\sigma^2} + \sum_t N_{it}}{\lambda_{i\cdot} + \frac{1}{\sigma^2}} \quad \Bigg| \quad \overline{\overline{y}}_{iT} = \sum_{t=1}^{T} \frac{\lambda_{it}}{\lambda_{i\cdot}} y_{it} \ ; \ z_{iT} = \frac{\lambda_{i\cdot}}{\lambda_{i\cdot} + \frac{1}{\sigma^2}}$$

Da cui si ottiene il premio per la frequenza $\widetilde{\mathbb{E}}(N_{i,T+1}|U_i) = \widetilde{\mathbb{E}}(\lambda_{i,T+1} Y_{i,T+1}|U_i) = \lambda_{i,T+1} \cdot \overset{\sim CRED}{\mu_{T+1}}(U_i)$, con $\lambda_{i,T+1}$ osservabile, nelle ipotesi di stabilità delle classi tariffarie e della composizione del portafoglio. Anche in questo caso si stimano prima i premi a priori per ciascuna classe tariffaria, poi σ^2 tramite a ed infine si combinano con la personalizzazione a posteriori.
Il modello viene inoltre utilizzato per valutare la riserva sinistri nella pratica.

<u>Cenni sulla credibilità esatta</u>

Si parla di credibilità esatta quando *il premio bayesiano coincide con quello di credibilità* per ogni individuo.
Partiamo dallo schema bayesiano: fissiamo un individuo della collettività e consideriamo il processo di interesse a tempo discreto $Y_1, Y_2, \dots$ ed il parametro di eterogeneità residua U. Si hanno le seguenti ipotesi:
∘ $\forall u$ determinazione di U, si ha che $Y_1|U = u, Y_2|U = u, \dots$ sono stocasticamente indipendenti e di legge assegnata
∘ U è di legge assegnata
Per il premio individuale nell'anno $T+1$, $\mathbb{E}(Y_{T+1}|U) = \mu_{T+1}(U)$ e si sfrutta lo stimatore bayesiano $\mathbb{E}(\mu_{T+1}(U)|Y_1, \dots, Y_T) = \mathbb{E}(Y_{T+1}|Y_1, \dots, Y_T)$ e può accadere che sia una combinazione lineare affine del processo di osservazione, cioè: $\mathbb{E}(\mu_{T+1}(U)|Y_1, \dots, Y_T) = \alpha_0 + \sum_t \alpha_t Y_t$, derivante dall'ottimizzazione della funzione di perdita quadratica attesa. Quindi se come minimo globale di $\mathbb{E}[(\mu_{T+1}(U) - g(Y_1, \dots, Y_T))^2]$ rispetto a g, si ottiene una funzione lineare affine, ne consegue che lo stimatore bayesiano coincide con lo stimatore di credibilità, essendo questo minimo locale. Pertanto $\forall y_1, \dots, y_T$ si ha che $P_{T+1}^{BAYES}(H_T) = P_{T+1}^{CRED}(H_T)$.
Un esempio di credibilità esatta si ha con il modello Poisson-Gamma:

$$\mathbb{E}(\mu_{T+1}(U)|N_1, \dots, N_T) = \mathbb{E}(N_{T+1}|N_1, \dots, N_T) = \lambda(1 - z_T) + \sum_{t=1}^{T} \frac{N_{it}}{T} z_{it}$$

Ovvero una combinazione lineare affine del processo di osservazione. E questo fenomeno si presenta ogni qual volta si combinino distribuzioni di famiglie coniugate.
Sia $Y_t|U = u \sim f_{Y_t|U=u} \in \mathcal{F}$ dotata di funzione di densità appartenete alla famiglia esponenziale lineare e $f_{Y_t|U=u} = \exp\{\frac{yu - b(u)}{\phi/\omega_t}\} c(y, \phi, \omega_t)$, con $b(\cdot)$ funzione cumulante, u parametro canonico, ϕ parametro di dispersione, $\omega_t > 0$ il peso assegnato e $c(\cdot)$ una funzione che non dipende da u. Le distribuzioni dei

n.a. Y_t variano esclusivamente in base ai pesi ω_t e $u \sim f_U \in \mathcal{U}$ tale che $f(u) = \exp\{\frac{\mu u - b(u)}{\delta}\}d(\mu, \delta)$ con parrametri μ, δ, $b \in]u_0, u_1[$ e $d(\cdot)$ costante di normalizzazione.
Dati $\omega. = \sum_t \omega_t$ e $\overline{\overline{y}}_T = \sum_t \frac{\omega_t y_t}{\omega.}$, si determina la legge a posteriori:

$$f_{U|HT}(u) \propto f_U(u) \prod_{t=1}^{T} f_{Y_t|U=u}(u) \propto e^{\frac{\mu u}{\delta}} e^{-\frac{b(u)}{\delta}} \prod_{t=1}^{T} e^{y_t u \frac{\omega_t}{\phi}} e^{-\frac{b(u)\omega_t}{\phi}} = e^{\frac{\mu u}{\delta}} e^{-\frac{b(u)}{\delta}} e^{\frac{u}{\phi} \sum_t \omega_t y_t} e^{-\frac{b(u)}{\phi} \sum_t \omega_t} =$$

$$= e^{u(\frac{\mu}{\delta} + \frac{1}{\phi}\omega.\overline{\overline{y}}_T)} e^{-b(u)(\frac{1}{\delta} + \frac{\omega.}{\phi})} = e^{\frac{\mu^* u - b(u)}{\delta^*}} \qquad \text{con} \qquad \mu^* = \left(\frac{\phi}{\delta}\mu + \omega.\overline{\overline{y}}_T\right)\left(\frac{\phi}{\delta} + \mu\right)^{-1} \;;\; \delta^* = \phi\left(\frac{\phi}{\delta} + \omega.\right)$$

Quindi ritroviamo una distribuzione a posteriori della stessa famiglia di distribuzioni a priori, ma con parametri aggiornati.
Si dice che la famiglia $\mathcal{U}$ è coniugata naturale della famiglia $\mathcal{F}$, esempi di famiglie coniugate sono la Gamma-Gamma, Normale-Normale, Binomiale-Beta e Poisson-Gamma.

Teorema

Nelle ipotesi precedenti, se $\mathcal{U}$ è famiglia coniugata di $\mathcal{F}$ e $f_U(u_0) = f_U(u_1) = 0$, ovvero la distribuzione è nulla negli estremi, allora si prova che $\mathbb{E}(Y_{T+1}|Y_1, \ldots, Y_T) = \mu(1 - z_T) + \overline{\overline{y}}_T z_T$ dove $\mu = \mathbb{E}(Y_{T+1})$, la speranza matematica a priori, e $z_T = \frac{\omega.}{\omega. + \phi/\sigma}$. Ed inoltre si avrà la credibilità perfetta.

Sistemi Bonus-Malus in RCA

Abbiamo un portafoglio ripartito in classi tariffarie ed ognuna di queste, viene ripartita ulteriormente in *classi di merito* o *classi Bonus-Malus.* La classe BM in un fissato anno t è funzione del numero di sinistri riportati e della classe dell'anno precedente, ovvero $C_t = f(C_{t-1}, N_{t-1})$. Si utilizza il numero di sinistri N_{t-1} e non gli importi, per individuare la classe, poiché le informazioni sul numero sono più affidabili e si ipotizza che l'attenzione dell'assicurato sia più significativa verso il numero di sinistri, piuttosto che all'importo. É un metodo più semplice della credibilità e per assegnare un sistema BM si assegnano: le *classi di merito* da 1 a J, le regole *di ingresso* per i nuovi assicurati, le regolo evolutive o *di transizione* $(\Psi(C_{t-1}, n_{t-1}) \to C_t)$ e la scala dei *coefficienti di premio*, con $\pi_{\leq} \cdots \leq \pi_J$.
Il premio BM nell'anno t per un assicurato in classe tariffaria k e BM j è $P^{BM}_{t,k,j} = P^{(k)}_t \gamma_k \pi_j$, dove $P_{t,k} = P^{(k)}_t \gamma_k$ è il premio a priori della classe tariffaria k nell'anno t, ed è detto *premio di riferimento.* Quindi il premio BM è una funzione tale che $P^{BM}_{t,k,j} = f(P_{t,k}, n_{t-1}, \Psi(n_{t-T}, \ldots, n_{t-2}))$.
Esempio: il sistema ministeriale italiano prima della liberalizzazione tariffaria, nel 1994, composto da 18 classi, con la 14^a di ingresso e con le seguenti regole evolutive:

$$C_t = C_{t-1} + \begin{cases} -1 & n_{t-1} = 0 \\ +2 & n_{t-1} = 1 \\ +5 & n_{t-1} = 2 \\ +8 & n_{t-1} = 3 \\ +11 & n_{t-1} \geq 4 \end{cases} \quad \Rightarrow \quad C_t = \begin{cases} \max\{C_{t-1} - 1, 1\} & n_{t-1} = 0 \\ \min\{C_{t-1} + 3n_{t-1} - 1, 18\} & n_{t-1} = 1, \ldots, 4 \\ \min\{C_{t-1} + 11, 18\} & n_{t-1} > 4 \end{cases}$$

Problema del sistema: sia la frequenza sinistri di un assicurato $\frac{1}{3}$, ovvero un sinistro ogni tre anni, e che ci sia ipotesi di stazionarietà. Qualunque sia la classe di partenza j, l'assicurato si muoverà sempre tra le classi $j - 2$ e $j + 2$, quindi se $f > \frac{1}{3}$ l'assicurato si muoverà verso le classi di Malus, se no verso quelle di Bonus. Ma essendo prima degli anni 2000 la frequenza sinistri di 0.1 ed adesso di 0.06, gli assicurati tendevano a concentrarsi tutti nelle classi minime. E questo comportava che la differenziazione nei premi fosse insufficiente e portasse a disequilibrio tecnico finanziario , non essendoci più compensazioni tra sconti ed aggravi di premio. Un altro problema è l'autoliquidazione, che gli assicurati eseguono per non risultare sinistrosi ed avanzare di classe, di conseguenza le denunce risultano inferiori ai sinistri previsti.

3.7 31/10/2019

Alcune grandezze per la valutazione del sistema Bonus-Malus

Fissata la classe tariffaria k, che ometteremo, abbiamo una collettività di rischi "analoghi" ed "indipendenti". Sia P_t il premio di riferimento nell'anno t e sia assegnato il sistema BM, ovvero siano assegnate le classi di merito, le regole di ingresso e di transizione e i coefficienti di scala. Consideriamo un fissato assicurato nella classe e siano:

∘ N_t il numero di sinistri causati nell'anno t dall'assicurato

∘ X_t il risarcimento totale nell'anno t e sia $X_t = \sum_t \sum_{h=1}^{N_t} Y_{t,h}$, con $Y_{t,h}$ il risarcimento associato al sinistro h-esimo dell'anno t

∘ C_t è la classe BM occupata dall'assicurato nell'anno t ed è assegnata in $t-1$.

Si ipotizza:

1 Indipendenza scolastica condizionata tra il processo di arrivo dei sinistri e il processo dei risarcimenti, quindi $(N_{T+1}|N_1 = n_1, \dots, N_T = n_T, H_T) \stackrel{d}{=} (N_{T+1}|N_1 = n_1, \dots, N_T = n_T)$, con $H_T = \bigwedge_{t=1,\dots,T;h=1,\dots,N_T}(Y_{t,h} \in A_{t,h})$ l'informazione sui risarcimenti per sinistro avvenuti nel periodo $[1, T[$.

2 X_t ha distribuzione composta, quindi $\mathbb{E}(X_t) = \mathbb{E}(N_t)\mathbb{E}(Y_t)$

La classe BM C_t è funzione di $C_1, N_1, \dots, N_{t-1}$ e di conseguenza il processo delle classi BM occupate dall'assicurato nel tempo $\{C_1, C_2, \dots\}$ ha valutazione probabilistica che dipende dal processo sottostante $\{C_1, N_1, N_2, \dots\}$, che determina l'evoluzione tra le classi:

a La distribuzione probabilistica di C_t ($Pr(C_t = j)$, con $j = 1, \dots, J$) dipende dalla sinistrosità e dalle regole evolutive del sistema BM

b Il coefficiente medio di premio nell'anno t è $b_t = \sum_j \pi_j Pr(C_t = j)$ e dato il premio base P_t, si ha il numero aleatorio $P_t\pi_j$ con $C_t = j$. Se poniamo $P_t = 1$, allora b_t indica il premio medio annuo in t. Anche in questo caso siamo interessati alla sua evoluzione, che dipende dalla sinistrosità, dalle regole evolutive e dai coefficienti π_j.

Osservazione 1: sia b_t decrescente nel tempo e $\mathbb{E}(X_t)$ costante e sia fissato il premio di riferimento P_1 tale che $\sum_j P_1\pi_j Pr(C_t = j) = \mathbb{E}(X_1) = P_1 b_1$, ovvero siamo in una situazione di equilibrio tecnico con i premi attesi che coprono i risarcimenti attesi in 1.
Supponiamo che in t il premio di riferimento coincida con quello iniziale $P_t = P_1$ e che quindi il premio atteso in t sia $P_t b_t = P_1 b_t < P_1 b_1 = \mathbb{E}(X_1) = \mathbb{E}(X_t)$. Quindi il premio atteso in t risulta minore del risarcimento atteso in t, non si ha quindi equilibrio tecnico in t. Perciò affinché ci sia l'equilibrio, è necessario che il premio di riferimento in t aumenti.

Osservazione 2: supponiamo di avere due assicurati in $t = 1$ in classe BM j, con π_j e h con π_h, tale che $j < h$ e quindi $\pi_j < \pi_h$. Se l'assicurato nell'anno t riesce a passare in classe j, il sistema promette lo sconto $\frac{P_1\pi_h - P_1\pi_j}{P_1\pi_h} = 1 - \frac{\pi_j}{\pi_h}$. Tuttavia lo sconto risulta inferiore nell'ipotesi di stazionarietà di $\mathbb{E}(X_t)$, poiché dovendo incrementare P_t si ha $\frac{P_1\pi_h - P_t\pi_j}{P_1\pi_h} = 1 - \frac{P_t}{P_1}\frac{\pi_j}{\pi_h} < 1 - \frac{\pi_j}{\pi_h}$. Quindi il sistema non rispetta le promesse.

Esempio: sia il sistema ministeriale, quindi sia un assicurato in $t = 1$ in classe 14, con $\pi_{14} = 1.15$.

In 5 anni senza sinistri, $C_6 = 9$ e $\pi_9 = 0.78$, quindi gli viene promesso circa il 32% di sconto, ma se in $t = 6$ $P_6 = 1.15P_1$, riceve $1 - \frac{0.78}{1.15}\frac{1.15P_1}{P_1}$ che è circa il 22%. Inoltre se l'aggravio di premio nel tempo è superiore allo sconto, può capitare che $P_6 > P_1$, nonostante la classe sia inferiore.
ⓒ Il premio di equilibrio $P_t^{(e)}$ è la risoluzione dell'equazione $\sum_j P_t\pi_j Pr(C_t = j) = \mathbb{E}(X_t)$. Siamo interessati all'evoluzione nel tempo $(P_t^{(e)})_t$, che dipende dalla sinistrosità, dalle regole evolutive, dai coefficienti e da $\mathbb{E}(X_t)$ e dalla sua valutazione probabilistica. Perciò sarà necessario assegnare la legge del processo sottostante $\{C_1, N_1, N_2, \dots\}$ e $\mathbb{E}(Y_t)$.

Ripasso Catene Markoviane Finite
Un processo a parametro discreto $\{C_1, C_2, \dots\}$, tale che ciascun n.a. C_t ha un numero finito J di determinazioni possibili ed uguali $\forall t$, è detto processo markoviano se $\forall t, i, j, k_{t-1}$ si ha che $Pr(C_{t+1} = j|C_t = i, k_{t-1}) = Pr(C_{t+1} = j|C_t = i)$, $\forall k_{t-1} = (C_1 = i_1, \dots, C_{t-1} = i_{t-1})$.
I processi markoviani vengono utilizzati per studiare l'evoluzione nei vari strati. Se in aggiunta $Pr(C_{T+1} = j|C_t = i) \triangleq p_{ij}$, $\forall i, j, t$, ovvero se non dipende dal tempo, si parla di catena markoviana. Indichiamo con P la matrice che contiene le probabilità condizionate di transizione p_{ij} e si verificano le seguenti proprietà, che la rendono una matrice stocastica:
◦ $p_{ij} \leq 0$, $\forall i, j$
◦ $\sum_j p_{ij} = \sum_j Pr(C_{t+1} = j|C_t = i) = Pr(\bigvee_j C_{T+1} = j|C_t = i) = Pr(\Omega) = 1$
Indichiamo con $\underline{a}_t$ il vettore riga che contiene le marginali unidimensionali del processo e consideriamo:
$a_2(j) = Pr(C_2 = j) = \sum_i Pr(C_2 = j, C_1 = i) = \sum_i Pr(C_1 = i)Pr(C_2 = j|C_1 = i) = \sum_i a_1(i)p_{ij} = \underline{a}_1 \cdot P$
Per induzione si dimostra che $\underline{a}_t = \underline{a}_{t-1} \cdot P$. Inoltre posto $P^{(2)} = P \cdot P$, si ha che $\underline{a}_t = \underline{a}_1 \cdot P^{(t-1)}$, perciò c'è un legame di tipo algebrico tra le marginali di un passo e quello successivo, che vengono ottenute semplicemente assegnando la distribuzione iniziale e P.
Osservazione 1: si dimostra che non si ottengono solo le marginali, ma la legge del processo di una catena markoviana, dati $\underline{a}_1$ e P.
Osservazione 2: sia $\tau \in \mathbb{N}$ e consideriamo la potenza $P^{(\tau)}$, l'elemento $p_{ij}^{(\tau)}$ può essere interpretato come la probabilità di passare dallo stato i allo stato j in τ passi, cioè $Pr(C_{t+\tau} = j|C_t = i)$.

Una catena markoviana è detta regolare se $\exists t_0$ tale che $\forall t \geq t_0$ si ha che $p_{ij}^{(t)} > 0$, $\forall i, j$. Ovvero se, da un determinato passo, passare da un qualunque stato ad un altro ha probabilità positiva, qualunque siano lo stato di arrivo e di partenza. Inoltre vale il teorema di Markoff, il quale afferma che in una catena regolare $\forall i, j$ $\exists \lim_{t\to+\infty} p_{ij}^{(t)} = v_j$. Si ha inoltre che $v_j > 0$, $\sum_j v_j = 1$ ed il limite non dipende da i. Perciò se consideriamo $P^{(t)}$ e facciamo divergere t si ha $\lim_{t\to+\infty} P^{(t)} = V$ con V la matrice con righe tutte uguali e per colonne v_j.
Conseguenze:
◦ $\underline{a}_t = \underline{a}_1 P^{(t-1)} \Rightarrow \lim_{t\to+\infty} \underline{a}_t = \underline{a}_1 V = \underline{v}$
◦ $\underline{a}_t = \underline{a}_{t-1}P \Rightarrow \lim_{t\to+\infty} \underline{a}_t = \lim_{t\to+\infty} \underline{a}_{t-1}P \Leftrightarrow \underline{v} = \underline{v}P$ e quindi troviamo $\underline{v}$ come risoluzione del sistema lineare:
$$\begin{cases} \underline{v} = \underline{v}P \\ \sum_{j=1}^J v_j = 1 \end{cases} \Leftrightarrow \begin{cases} \underline{v}(P - I) = 0 \\ \underline{v}1_J = 1 \end{cases}$$

3.8 04/11/2019

Le valutazioni introdotte dipendono dal processo sottostante $\{C_1, N_1, N_2, \dots\}$, che determina lo spostamento dell'assicurato nell'ambito delle classi, data la classe assegnata al primo anno e l'evoluzione della sinistrosità nel tempo. Assegnato il sistema BM andremo a considerare due modelli:

Modello A
- $\forall j = 1, \dots, J$, $N_1|C_1 = j$, $N_2|C_1 = j, \dots$ sono IID
- $N_t|C_1 = j$ ha distribuzione che non dipende da j e $Pr(N_t = n|C_1 = j) = p_n$ assegnata $\forall n$
- É assegnata la distribuzione di C_1

Proprietà del Modello A

1 $Pr(N_t = n|k_t) = p_n$ $\forall t, k, n$, con $k_t = (C_1 = j_1, \dots, C_t = j_t)$ una storia di classi occupate nei t anni.

Dimostrazione: fissati $t, k_t \neq \emptyset, n$, si considera $Pr(N_t = n|k_t)$, con k_t una somma di storie relative al processo sottostante, ovvero $k_t = \bigvee(H_{t-1}) = \bigvee(C_1 = j_1, N_1 = n_1, \dots, N_{t-1} = n_{t-1})$ ed inoltre $H_t \Rightarrow k_t$.

$$Pr(N_t = n|k_t) = \sum_{H_{t-1} \Rightarrow k_t} Pr(H_{t-1}|k_t) Pr(N_t = n|H_{t-1}) = \sum_{H_{t-1} \Rightarrow k_t} Pr(H_{t-1}|k_t) Pr(N_t = n|(C_1 = j_1,$$

$$N_1 = n_1, \dots, N_{t-1} = n_{t-1})) = \sum_{H_{t-1} \Rightarrow k_t} Pr(H_{t-1}|k_t) \underbrace{Pr(N_t = n|C_1 = j_1)}_{p_n} = p_n \overbrace{\sum_{H_{t-1} \Rightarrow k_t} Pr(H_{t-1}|k_t)}^{Pr(\Omega)} = p_n$$

2 Il processo $\{C_1, C_2, \dots\}$ delle classi occupate nel tempo, risulta una catena markoviana, ovvero $\forall i, j, k_{t-1}, t$ si ha che $Pr(C_{t+1} = j|C_t = i, k_{t-1}) = Pr(C_{t+1} = j|C_t = i) = p_{ij}$.

Dimostrazione: fissati i, j, t, k_{t-1}, consideriamo:

$$Pr(C_{t+1} = j|C_t = i, k_{t-1}) = \sum_n Pr(C_{t+1} = j, N_t = n|C_t = i, k_{t-1}) =$$

$$= \sum_n \overbrace{Pr(N_t = n|\underbrace{C_t = i, k_{t-1}}_{k_t})}^{a} \overbrace{Pr(C_{t+1} = j|N_t = n, C_t = i, k_{t-1})}^{b}$$

a É pari a p_n e non dipende né da k_t, né da t

b Noti $N_t = n$ e $C_t = i$, conosciamo la classe di arrivo $C_{t+1} = j$ e quindi la probabilità è pari a 1 o a 0. Inoltre anche questo termine non dipende né da k_t, né da t.

Poiché né a, né b dipendono da t o k_t, allora neanche la loro somma dipenderà da t o k_t, poniamo quindi $Pr(C_{t+1} = j|C_t = i, k_{t-1}) \stackrel{\Delta}{=} \alpha(i, j)$.

Inoltre dato $A = \{k_{t-1} : k_{t-1} \wedge (C_t = i) \neq \emptyset\}$, si consideri:

$$Pr(C_{t+1} = j|C_t = i) = \sum_A Pr(C_{t+1} = j \wedge k_{t-1}|C_t = i) = \sum_A Pr(k_{t-1}|C_t = j) Pr(C_{t+1} = j|k_{t-1} \wedge C_t = i) =$$

$$= \alpha(i, j) \sum_A Pr(k_{t-1}|C_t = i) = \alpha(i, j) Pr(C_t = i|C_t = i) = \alpha(i, j)$$

Quindi per ottenere la legge del processo $\{C_1, C_2, \dots\}$ è necessario assegnare la distribuzione iniziale $\underline{a}_1 = (Pr(C_1 = j_1), \dots, Pr(C_1 = J))$, che si ha dalle ipotesi del modello, e P la matrice stocastica delle probabilità condizionate di transizione, che si ottiene direttamente dalle p_n e dalle regole evolutive del sistema BM.

La distribuzione di $\{C_1, C_2, \dots\}$ e p_n devono essere valutate in modo da rispettare la sinistrosità degli assicurati:
- Per $\{C_1, C_2, \dots\}$ ci si può basare sulla frequenza relativa di assicurati per ciascun classe
- Per p_n ci si può basare sulle frequenze relative del numero di sinistri sul numero totale di assicurati
- Per l'evoluzione della distribuzione $Pr(C_t = j)$, dati $\underline{a}_1$, in maniera iterativa si ha che $\underline{a}_t = \underline{a}_{t-1}P$
- I coefficienti medi di premio $b_t = \sum_j \pi_j Pr(C_t = j)$ risultano una sintesi della distribuzione, perciò possiamo rilevarne l'andamento nel tempo
- Per i premi di equilibrio $P_t^{(e)}$, il premio tale che $\sum_j P_t \pi_j Pr(C_t = j) = \mathbb{E}(X_t) = \mathbb{E}(N_t)\mathbb{E}(Y_t)$ e se $\mathbb{E}(Y_t)$ è assegnata, allora $P_t^{(e)} = \mathbb{E}(N_t)\mathbb{E}(Y_t)b_t^{-1}$ con $\mathbb{E}(N_t) = \sum_j Pr(C_1 = j)\mathbb{E}(N_t|C_1 = j) = \sum_j Pr(C_1 = j)\sum_n np_n = \sum_n np_n$

Questo modello è molto diffuso nella pratica, tuttavia presenta delle ipotesi molto forti:
a $\forall j,\ N_1|C_1 = j, N_2|C_1 = j, \dots$ sono ID. Questa ipotesi implica la stazionarietà, ma non è ragionevole pensare che l'assicurato neo-patentato abbia la stessa sinistrosità 10 anni dopo. Non comporta però grandi problematiche, qualora si facciano valutazioni sul breve periodo.
b $\forall j,\ N_1|C_1 = j, N_2|C_1 = j, \dots$ sono stocasticamente indipendenti, condizionatamente all'informazione sulla classe occupata il primo anno. E questo non è ragionevole poiché siamo nella personalizzazione a posteriori ed abbiamo solo l'informazione sulla classe occupata. La differenza sostanziale è che il parametro aleatorio di rischio modella le informazioni non osservabili e se avessimo l'informazione sul profilo di rischio, sarebbe ininfluente conoscere la storia passata.
c $N_t|C_1 = j$ ha legge che non dipende da j, ovvero si assegna la medesima distribuzione ad un assicurato in classe 1 e ad uno in classe 18.

Tuttavia il modello viene applicato molto spesso, in particolare per l'analisi dell'evoluzione della collettività rispetto alla classe di ingresso, valutandone la distribuzione asintotica che si raggiunge dopo un certo numero di anni.

<u>Modello</u> B

É un modello con ipotesi meno forti, ma più complesso e introduce il parametro aleatorio di rischio U:
- $\forall j, u\ N_1|C_1 = j, U = u, N_2|C_1 = j, U = u, \dots$ sono IID
- $\forall j, u\ N_t|C_1 = j, U = u \sim Poi(\lambda_j u)$, con $\lambda_j > 0$
- É assegnata la legge di $U|C_1 = j \sim \Gamma(\alpha, \alpha)$ con $\alpha > 0$
- É assegnata la distribuzione di C_1

Le prime tre ipotesi rivelano che il processo $\{N_1|C_1 = j, N_2|C_1 = j, \dots\}$ è un processo $Poisson - Gamma(\lambda_j, \alpha)$ e quindi $\{N_1, N_2, \dots\}$ è un processo mistura di processi $Poisson - Gamma$.
Perciò considerato il processo $\{N_1|C_1 = j, N_2|C_1 = j, \dots\}$ si ha che:

$$N_t|C_1 = j \sim BN\left(\alpha, \frac{\alpha}{\alpha + \lambda_j}\right) \quad \text{quindi} \quad \begin{cases} \mathbb{E}(N_t|C_1 = j) = \lambda_j \\ V(N_t|C_1 = j) = \lambda_j + \frac{\lambda_j^2}{\alpha} \end{cases}$$

$$Pr(N_1 = n_1, \dots, N_T = n_T|C_1 = j) = \frac{\Gamma(\alpha + \sum_t n_t)}{\Gamma(\alpha)\prod_t(n_t!)}\left(\frac{\alpha}{\alpha + \lambda_j T}\right)^{\alpha}\left(\frac{\lambda_j}{\alpha + \lambda_j T}\right)^{\sum_t n_t}$$

$$N_{T+1}|(N_1 = n_1, \dots, N_T = n_T, C_1 = j) \sim BN\left(\alpha + \sum_{t=1}^{T} n_t, \frac{\alpha + \lambda_j T}{\alpha + \lambda_j(T+1)}\right)$$

Quindi per assegnare la legge del processo è necessario assegnare $\underline{\lambda} = (\lambda_1, \dots, \lambda_J)$, α e la distribuzione

iniziale $Pr(C_1 = j)$, che ricaviamo dalle frequenze relative per classe. Per i parametri $\underline{\lambda}$ e α ci basiamo sulle stime empiriche di media e varianza, mediante le relazioni per ciascuna classe. Potremmo però avere pochi dati, rendendo necessario il raggruppamento in più classi, ma la cosa fondamentale è che i λ_j non siano tutti uguali, questo comporta che ci siano molte stime iniziali da dove calcolare.
Considerata $A = \{H_{t-1} : H_{t-1} \Rightarrow C_t = j\}$ inoltre possibile calcolare le seguenti probabilità:

$$Pr(C_t = j) = \sum_A Pr(C_1 = j_1, N_1 = n_1, \dots, N_{t-1} = n_{t-1}) = \sum_A \underbrace{Pr(C_1 = j_1)}_{\text{Assegnate}} Pr(N_1 = n_1, \dots, N_{t-1} = n_{t-1}|C_1 = j_1)$$

Quindi assegnate anche $\mathbb{E}(Y_t)$ e $\mathbb{E}(N_t)$, possiamo calcolare anche b_t e $P_t^{(e)} = \mathbb{E}(N_t)\mathbb{E}(Y_t)b_t^{-1}$. Quindi per la distribuzione è necessario tener conto di tutta la storia di sinistrosità.

3.9 07/11/2019

Analisi di sistemi BM

Modello ministeriale: si stimano le $\widehat{Pr}(C_1 = j)$ tramite le frequenze relative, siano $\mathbb{E}(N_t) = 0.1$ la frequenza sinistri, $\mathbb{E}(N_{t+1}|C_1 = j)$ stimata tramite la frequenza sinistri per ciascuna classe BM e $\mathbb{E}(Y_t) \sim$ *Pareto* pari a €1500 e costante nel tempo. Se stimiamo le $Pr(C_t = j)$ notiamo che dopo 5 anni già la maggior parte degli assicurati sarà nelle classi più basse e dopo 15 più del 60% sarà nella prima classe, che fa emergere un problema di penalizzazione insufficiente. Stimando invece i b_t, notiamo che questi sono decrescenti nel tempo, in particolare nei primi anni. Emerge quindi l'esigenza di incrementare i $P_t^{(e)}$, che aumenterebbe nei primi cinque anni del 15% e nei primi 15 del 30%. In ipotesi di portafoglio aperto, con i nuovi assicurati che entrano in classe BM 14, i $P_t^{(e)}$ aumenterebbero comunque, ma in maniera inferiore, poiché i nuovi assicurati entrano in una classe malus essendo $\pi_{14} > 1$, contribuendo a creare l'equilibrio tecnico. Anche tenendo conto dell'autoliquidazione i $P_t^{(e)}$ aumenterebbero. Il fenomeno dell'autoliquidazione viene studiato considerando l'individuo razionale, ovvero che se nell'anno t commette 2 o più sinistri denuncerà tutto, mentre se ne commette solo uno, lo denuncerà solo se l'importo del danno z è maggiore di $\overline{z}$, detta soglia di autoliquidazione, determinata, in ipotesi di non commettere sinistri nei prossimi 5 anni, dalla differenza del premio attualizzato senza denuncia e con denuncia. In questo caso l'evoluzione del processo è sul numero di denunce e non di sinistri $\{C_1, D_1, D_2, \dots\} \leq \{C_1, N_1, N_2, \dots\}$. Un altro aspetto interessante da considerare è confrontare lo sconto promesso all'assicurato, rispetto a quello effettivo. Posto P_1 in classe 14, lo sconto promesso è pari a $1 - \frac{\pi_j}{\pi_1}$, mentre invece lo sconto effettivo è $1 - \frac{P_t^{(e)}}{P_1^{(e)}}\frac{\pi_j}{\pi_1}$. In condizioni di stazionarietà quindi il sistema non premia gli assicurati quanto dovrebbe.

Funzioni per valutare l'effetto di personalizzazione e solidarietà introdotti dal sistema BM

Fissata una classe tariffaria, consideriamo un collettivo omogeneo e sia P_t il premio base della classe tariffaria al tempo t ed assumiamo che in $t = 1$ tutti gli assicurati siano inseriti nella stessa classe di ingresso j_0, perciò la distribuzione iniziale è $Pr(C_1 = j) = 0$, $\forall j \neq j_0$. L'evoluzione tra le classi è determinata quindi solo dal processo $\{N_1, N_2, \dots\}$.
Il premio BM in t è pari a $P_t\pi_{C_t}$ con C_t n.a. e P_t certo.É possibile confrontare il premio BM con altri premi tramite le funzioni:
A $\mathbb{E}((\mathbb{E}(X_t) - P_t\pi_{C_t})^2)$ lo scostamento quadratico medio tra il premio BM e il premio a priori. Minore è il valore, minore è la personalizzazione nella classe.
B $\mathbb{E}((\mathbb{E}(X_t|U) - P_t\pi_{C_t})^2)$ lo scostamento quadratico medio tra il premio BM e il premio individuale. Minore è il valore, maggiore è la personalizzazione.
C $\mathbb{E}((\mathbb{E}(X_t|C_t) - P_t\pi_{C_t})^2)$ lo scostamento quadratico medio tra il premio BM e il premio a priori per la

classe di sinistrosità occupata. Minore è il valore, maggiore è la personalizzazione.
Generalmente si richiede un bilanciamento nel sistema, per evitare il rischio di autoselezione per l'eccessiva solidarietà.
In letteratura non ci sono grandi modelli per la costruzione di sistemi BM, ma ne esistono di raffinati per il fenomeno dell'autoliquidazione e per la costruzione delle scale dei coefficienti.

Costruzione di una scala per un assegnato sistema Bonus-Malus

Consideriamo assegnati per un sistema BM le J classi e le regole di transizione ed in ingresso. Fissata una classe tariffaria, assumiamo che in $t = 1$ tutti gli assicurati siano nella classe di ingresso. Si sfrutta come modello metodologico l'approccio bayesiano, determinando i π_j, ai fini di stimare i premi per la "frequenza".
Consideriamo quindi il processo di arrivo dei sinistri $\{N_1, N_2, \dots\}$ sul fissato assicurato ed ipotizziamo:
∘ $\forall u$ determinazione di U si ha che $N_1|U = u, N_2|U = u, \dots$ sono stocasticamente indipendenti e di legge assegnata
∘ U è di legge assegnata
L'obiettivo è di approssimare il premio individuale $\mathbb{E}(N_t|U)$, attraverso i seguenti tre approcci.

1 **Norberg (1976)**

Si introduce una funzione della classe BM occupata nell'anno t $g(C_t)$ come stimatore di $\mathbb{E}(N_t|U)$. In particolare $g(\cdot)$ è una funzione che fa parte della famiglia di funzioni regolari (g^2-integrabile) tale che :

$$\min_{g\in\mathcal{G}} \mathbb{E}\left[\left(\mathbb{E}(N_t|U) - g(C_t)\right)^2\right] = \min_{g\in\mathcal{G}} \mathbb{E}\left[\left(\mu_t(U) - g(C_t)\right)^2\right] \quad \Rightarrow \quad g^*(j) = \mathbb{E}(\mu_t)$$

Inoltre abbiamo dimostrato che equivale a determinare la perdita quadratica minima attesa sulla grandezza di interesse.
Ipotesi: si introduce una classe di riferimento, cioè $\exists h : \pi_h = 1$ e si ha che $\frac{g^*(j)}{g^*(h)} = \pi_j = \frac{\mathbb{E}(N_t|C_t=j)}{\mathbb{E}(N_t|C_t=h)}$ è il premio per la classe j, che però varia al variare di t.

2 **Norberg, Borgan, Hoen (1981)**

Propongono di cambiare funzione obiettivo, considerando t su un intervallo di tempo. Si trova quindi $g(\cdot)$ come:

$$\min_{g\in\mathcal{G}} \sum_{t=1}^{T} \omega_t \mathbb{E}\left[\left(N_t - g(C_t)\right)^2\right] \quad \Rightarrow \quad g^*(j) = \frac{\sum_t Pr(C_t = j)\mathbb{E}(N_t|C_t = j)}{\sum_t \omega_t Pr(C_t = j)} \quad \text{e} \quad \pi_j = \frac{g^*(j)}{g^*(h)}$$

3 **Coene, Doney (1996)**

Nell'approccio bayesiano, data la storia $H_{t-1} = (N_1 = n_1, \dots, N_{t-1} = N_{t-1})$, $\mathbb{E}(N_t|U)$ è stimato da $\mathbb{E}(N_t|H_{t-1})$ e $\frac{\mathbb{E}(N_t|H_{t-1})}{\mathbb{E}(N_t)} = C_t(n_1, \dots, n_{t-1})$ è detto *coefficiente di revisione bayesiana.*
Nel sistema BM, l'assicurato al tempo t si troverà nella classe $C_t = J(n_1, \dots, n_{t-1})$, che è quindi funzione del numero di sinistri riportati ed il premio BM è pari a $P_t\pi_{J(n_1,\dots,n_{t-1})}$. Possiamo quindi confrontarlo con quello bayesiano $\mathbb{E}(N_t|H_{t-1}) = \mathbb{E}(N_t)C_t(n_1, \dots, n_{t-1})$ e minimizzare su un intervallo di tempo per tutte le possibili storie, ponendo vincoli sulla monotonia dei π_j:

$$\min_{\pi_1<\dots<\pi_J} \sum_{\substack{t=1 \\ n_1,\dots,n_{t-1}}}^{T} \omega_t\left(C_t(n_1, \dots, n_{t-1}) - \pi_{J(n_1,\dots,n_{t-1})}\right)^2$$

I modelli proposti sono adeguati per analisi di controllo per ogni classe, ma la scala dei coefficienti è unica su tutto il portafoglio. Non è possibile dunque fissare k, ma si tiene conto tramite $Pr(C_t = j) = \sum_k p_{k,t} Pr_k(C_t = j)$, con $p_{k,t}$ la frequenza relativa degli assicurati nella classe k al tempo t.

Riferimenti Bibliografici:

1 Lemaire J. *Bonus-malus systems in automobile insurance*, Springer science & business media (2012)
2 Denuit M. *Actuarial Modelling of Claim Counts*, Wiley (2007)
3 Kaas R. *Modern actuarial risk theory: using R*, Springer Science & Business Media (2008)

4 Misure di Rischio

4.1 08/11/2019

La misurazione dei rischi è un problema fondamentale per la valutazione dei premi, per le politiche di prezzi, per le coperture riassicurative ed ultimamente anche per le nuove norme contabili e di solvibilità.

Inversa Generalizzata di una Funzione di Ripartizione

Sia F una Funzione di Ripartizione unidimensionale, ossia una funzione reale di variabile aleatoria reale, che soddisfi le seguenti proprietà:
- $F : \mathbb{R} \to [0, 1]$
- F monotona non decrescente
- $\lim_{x\to-\infty} F(x) = 0$
- $\lim_{x\to+\infty} F(x) = 1$
- F continua a destra $\forall x \in \mathbb{R}$

In generale quindi F è monotona non decrescente, ma se F è strettamente crescente, allora $\exists F^{-1}$. In genere però non è strettamente crescente e qualora lo sia può anche capitare che l'insieme immagine non sia tutto il codominio $[0, 1]$.

Definizione

Sia F una FdR, è detta *Funzione Inversa Generalizzata* di F, la funzione $\overset{\leftarrow}{F}$ definita da:

$$\overset{\leftarrow}{F}(p) = \inf\{x \in \mathbb{R} : F(x) \geq p\}$$

Dove $p \in [0, 1]$ e per convenzione $\inf\{R\} = -\infty$ e $\inf\{\emptyset\} = +\infty$. $\overset{\leftarrow}{F}$ è anche detta *Funzione Quantile* di F.

Sia $A_p = \{x \in \mathbb{R} : F(x) \geq p\} \subset \mathbb{R}$, ha senso parlare di estremo inferiore?
Se $A_p \neq \emptyset$ e A_p è inferiormente limitato, allora se considero l'Insieme delle limitazioni inferiori, o dei minoranti, si prova che questo ha massimo, ed è detto estremo inferiore dell'insieme, quindi $\inf\{A_p\} \in \mathbb{R}$.
- Sia $p \in \mathrm{I}$ un punto in cui F è monotona crescente e quindi invertibile, allora $\overset{\leftarrow}{F}(p) = F^{-1}(p)$
- Sia $p_1 \in \mathrm{I}$ tale che $\exists x_1 \neq x_2$ con $F(x_1) = F(x_2) = p_1$, allora F non è invertibile e $\overset{\leftarrow}{F}(p_1) = \min\{x \in \mathbb{R} : F(x) = p_1\}$
- Sia $p_2 \notin \mathrm{I}$, allora F non è invertibile, ma $\overset{\leftarrow}{F}(p_2) = \overset{\leftarrow}{F}(\min\{p \in \mathrm{I} : p > p_2\})$

Quindi tramite l'inversa generalizzata troviamo $F^{-1}(p)$, dove la F è invertibile, ma riusciamo a trovare un valore, anche nel caso in cui F non lo sia, oppure il punto non appartenga all'insieme immagine I.

Osservazione 1

∘ Sia $\boxed{p=0}$, quindi stiamo considerando $A_0 = \{x \in \mathbb{R} : F(x) \geq 0\} = \mathbb{R}$ e quindi $A_0 \neq \emptyset$, non inferiormente limitato e per convenzione $\inf\{A_0\} = -\infty$, allora $\overleftarrow{F}(0) = -\infty$ e quindi l'inversa generalizzata assume valori sula retta reale ampliata.

∘ Sia $\boxed{p=1}$, allora $A_1 = \{x \in \mathbb{R} : F(x) \geq 1\}$ e questo è uguale a $\emptyset$, se $1 \notin$I e quindi $\inf\{A_1\} = +\infty$, diverso da $\emptyset$ se $1 \in$I e quindi proviamo che $\inf\{A_1\} \in \mathbb{R}$. Affinché esista l'estremo inferiore, dobbiamo verificare che l'insieme A_1 è inferiormente limitato, ovvero che $\exists x' \in \mathbb{R} : x' \leq x,\ \forall x \in A_1$. Per la proprietà della FdR, il $\lim_{x \to -\infty} F(x) = 0 \Leftrightarrow \forall I_0, \exists I_{-\infty} : \forall x \in I_{-\infty},\ F(x) \in I_0$. Fissiamo un intorno I_0 del tipo $]-\epsilon, \epsilon[$, con $0 < \epsilon < 1$. Per definizione di limite, all'intorno I_0, corrisponde un opportuno $I_{-\infty} \Leftrightarrow I_0 \to I_{-\infty}$ e sia $x' \in I_{-\infty}$. Quindi $F(x') \in I_0$ se e solo se $F(x') < \epsilon < 1$. Quindi se $F(x') < F(x)\ \forall x \in A_1$, allora $x' < x$ $\forall x \in A_1$. Ma allora, se $A_1 \neq \emptyset$, è inferiormente limitato, quindi $\exists \inf\{A_1\} \in \mathbb{R}$ e $\overleftarrow{F}(p) \in \mathbb{R}$.

∘ Sia $\boxed{0<p<1}$, allora si dimostra che $A_p \neq \emptyset$ ed è inferiormente limitato, quindi $\overleftarrow{F}(p) \in \mathbb{R}$.

<u>$A_p \neq \emptyset$</u>

Sia la proprietà della FdR $\lim_{x \to +\infty} F(x) = 1$, ovvero $\forall I_1,\ \exists I_{+\infty} : \forall x \in I_{+\infty},\ F(x) \in I_1$. Fissiamo un intorno arbitrario $I_1 =]1-\epsilon, 1+\epsilon[$ tale che $p \notin I_1$ e $p < 1-\epsilon$. Allora $F(x) \in I_1$ se e solo se $F(x) > 1-\epsilon > p$, allora per definizione di A_p, $x \in A_p$ ed essendoci almeno un elemento, $A_p \neq \emptyset$.

<u>A_p è inferiormente limitato $\Leftrightarrow \exists x' \in \mathbb{R} : x' < x\ \forall x \in A_p$</u>

Fissiamo un intorno arbitrario $I_0 =]-\epsilon, \epsilon[$ tale che $p \notin I_0$. Dato che $I_0 \to I_{-\infty}$, preso $x' \in I_{-\infty} \Rightarrow F(x') \in I_0 \Leftrightarrow F(x') < \epsilon < p$, quindi $F(x) \geq p\ \forall x \in A_p$, allora $F(x') < F(x) \Rightarrow x' < x\ \forall x \in A_p$. Quindi per $0 < p < 1\ \exists \inf\{A_p\} \in \mathbb{R}$ e quindi $\overleftarrow{F} \in \mathbb{R}$.

Quindi $\overleftarrow{F} : [0, 1] \to \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$

Osservazione 2

Sia $p \in]0, 1[$ e definiamo $\overleftarrow{F}(p) = \inf\{A_p\} \triangleq \lambda \in \mathbb{R}$. Si prova che $\forall x \in \mathbb{R} : x > \lambda \Rightarrow x \in A_p$. Fissiamo $x \in \mathbb{R}$ tale che $x > \lambda$, allora $\exists x' \in A_p : x' < x$. Quindi poiché $F(x') \geq p$ e $F(x') < F(x)$, allora $F(x) \geq p \Rightarrow x \in A_p$.

Ne segue quindi che A_p è un intervallo non degenere, superiormente illimitato e senza buchi.

Osservazione 3

Consideriamo l'insieme $B = \{x \in \mathbb{R} : 0 < F(x) < 1\}$ e sia $B \neq \emptyset$, quindi con $B^c = \{x \in \mathbb{R} : F(x) = 1 \vee F(x) = 0\}$. Consideriamo la partizione su B e sia $F_{|_B}$ strettamente crescente, troviamo quindi che $F_{|_B} : B \to F(B)$ è iniettiva e suriettiva. Quindi $F_{|_B}$ è invertibile, $\exists (F_{|_B})^{-1} : F(B) \to B$ e poniamo $(F_{|_B})^{-1} = F^{-1}$. Sia $p \in F(B)$, quindi $\exists F^{-1}(p)$ e proviamo che $F^{-1}(p) = \overleftarrow{F}(p)$, in tutti i punti in cui esiste F^{-1}. Per prima cosa dimostriamo che $A_p = \{x \in \mathbb{R} : F(x) > p\} = \{x \in \mathbb{R} : x \geq F^{-1}(p)\} \triangleq A$, poiché se sono uguali, allora gli estremi inferiori coincidono.

Verifichiamo che $A \subset A_p$: sia $x \in \mathbb{R} : x \geq F^{-1}(p) \Rightarrow x \in A$. Ma allora per la monotonia $F(x) \geq F(F^{-1}(p)) = p \Rightarrow x \in A_p$.

Verifichiamo che $A_p \subset A$: sia $x \in \mathbb{R} : F(x) \geq p \Rightarrow x \in A_p$ e possiamo avere due casi:

- Sia $x \in B$, allora $F(x) \in F(B)$, quindi possiamo calcolare F^{-1} e $F^{-1}(F(x)) \geq F^{-1}(p) \Leftrightarrow x \geq F^{-1}(p)$
- Sia $x \notin B$, allora $F(x) = 0$ o $F(x) = 1$. Ma $x \in A_p$ e per ipotesi $p \in]0, 1[$, quindi può essere solo $F(x) = 1 > p = F(F^{-1}(p))$. E quindi troviamo che $F(x) > F(F^{-1}(p)) \Leftrightarrow x > F^{-1}(p) \Rightarrow x \in A$.

Di conseguenza $x \in A \wedge x \in A_p \Rightarrow A = A_p$ e quindi troviamo che $\overleftarrow{F}(p) = \inf\{A_p\} = \inf\{A\} = F^{-1}(p)$. Quindi se $0 < F(x) < 1$ ed F è invertibile, allora $F^{-1} \equiv \overleftarrow{F}$, dove F^{-1} è definita sull'insieme immagine, mentre $\overleftarrow{F}$ è definita anche negli intervalli non appartenenti all'immagine.Inoltre se F è continua in $\mathbb{R}$ e strettamente crescente dove è diversa da 0 e da 1, allora $F^{-1} \equiv \overleftarrow{F}$, $\forall p \in]0,1[$. Di conseguenza abbiamo che l'insieme immagine è tutto $]0,1[$.

<u>**Osservazione 4**</u>

Supponiamo che F sia costante a tratti, con salti in $x_1 < x_2 < \ldots$. Consideriamo gli insiemi $]0, F(x_1)]$, $]F(x_1), F(x_2)], \ldots$, cioè gli intervalli inferiormente aperti e superiormente chiusi dei valori della FdR nei punti di salto. L'unione di questi insiemi equivale a $]0,1[$, ovvero $\forall p \in]0,1[$, $\exists! i : F(x_{i-1}) < p < F(x_i)$ e $\overleftarrow{F}(p) = x_i$, essendo gli intervalli a due a due disgiunti.

<u>**Altre proprietà dell'inversa generalizzata**</u>

i $F(\overleftarrow{F}(p)) \geq p$, $\forall p \in]0,1[$ e $\overleftarrow{F}(F(x)) \leq x$, $\forall x \in \mathbb{R}$

ii $\overleftarrow{F}$ è monotona non decrescente

iii $F(x) \geq p \Leftrightarrow x \geq \overleftarrow{F}(p)$

Dimostrazioni

i.1 Fissato arbitrariamente $p \in]0,1[$, si ha che $\overleftarrow{F}(p) \in \mathbb{R}$. Poniamo $\overleftarrow{F}(p) = \lambda$ ed essendo reale, possiamo calcolare $\lim_{x\to\lambda^+} F(x) = F(\lambda)$. É equivalente calcolare quel limite, o il limite della funzione ristretta all'insieme $A = \{x : x > \lambda\}$, ovvero $\lim_{x\to\lambda^+} F_{|A}(x)$. Dato che $x \in A$, allora $x \in A_p$, quindi $F(x) \geq p \Rightarrow F_{|A}(x) \geq p$ e per la monotonia del limite:

$$\lim_{x\to\lambda^+} F_{|A}(x) \geq \lim_{x\to\lambda^+} p \Leftrightarrow F(\lambda) \geq p \Leftrightarrow F(\overleftarrow{F}(p)) \geq p$$

i.2 Fissiamo arbitrariamente $x \in \mathbb{R}$, si ha che $F(x) \in [0,1]$ e $\forall p \in [0,1]$, $\exists\, \overleftarrow{F}(p)$. Per definizione $\overleftarrow{F}(F(x)) = \inf\{A_{F(x)}\} = \inf\{t \in \mathbb{R} : F(t) \geq F(x)\}$. Poiché $F(x) \geq F(x)$, allora $x \in A_{F(x)} \neq \emptyset$. Ogni elemento dell'insieme è maggiore o uguale all'estremo inferiore, quindi $\inf\{A_{F(x)}\} \leq x \Leftrightarrow \overleftarrow{F}(F(x)) \leq x$

ii $\overleftarrow{F}$ è monotona non decrescente se e solo se $\forall p_1, p_2 \in [0,1]$ con $p_1 < p_2$ si ha che $\overleftarrow{F}(p_1) \leq \overleftarrow{F}(p_2)$. Fissiamo arbitrariamente $p_1, p_2 \in [0,1]$ con $p_1 < p_2$. si considerino i seguenti casi:

- sia $p_1 = 0$, allora $\overleftarrow{F}(p_1) = \overleftarrow{F}(0) = -\infty < \overleftarrow{F}(p_2)$ $\forall p_2$, poiché $p_2 \neq 0$
- sia $p_2 = 1$ e $F(\cdot)$ è tale che $A_1 = \emptyset$, allora $\overleftarrow{F}(p_2) = \overleftarrow{F}(1) = +\infty \geq \overleftarrow{F}(p_1)$, $\forall p_1$
- siano $p_1, p_2 \in]0,1[$ o $p_1 \in]0,1[$ e $p_2 = 1$, con $F(\cdot)$ tale che $A_1 \neq \emptyset$. Si considerino gli insiemi $A_{p_1} = \{x \in \mathbb{R} : F(x) \geq p_1\}$ e $A_{p_2} = \{x \in \mathbb{R} : F(x) \geq p_2\}$. I due insiemi sono entrambi non vuoti e limitati inferiormente, per cui esiste l'estremo inferiore. É immediato verificare che $A_{p_2} \subset A_{p_1}$, infatti preso $x \in A_{p_2}$, allora $F(x) \geq p_2 > p_1$, quindi $x \in A_{p_1}$. Ma allora se $A_{p_2} \subset A_{p_1}$, si ha che $\inf\{A_{p_1}\} \leq \inf\{A_{p_2}\} \Leftrightarrow \overleftarrow{F}(p_1) \leq \overleftarrow{F}(p_2)$.

iii Fissiamo arbitrariamente $x \in \mathbb{R}$ e $p \in [0,1]$ e consideriamo i seguenti casi:

- sia $p = 0$, allora è sempre vero sia che $F(x) \geq 0$, sia che $x \geq \overleftarrow{F}(0) = -\infty$
- sia $p = 1$ e $F(\cdot)$ tale che $A_1 = \emptyset$, allora $F(x) \geq 1$ è sempre falso ed anche $x \geq \overleftarrow{F}(1) = +\infty$ è sempre falso
- sia $0 < p < 1$ o $p = 1$ e $F(\cdot)$ è tale che $A_1 \neq \emptyset$, allora dato $F(x) \geq p \Leftrightarrow x \in A_p$, si ha che $x \geq \inf\{A_p\} = \overleftarrow{F}(p)$, ed allo stesso modo $\overleftarrow{F}(p) = \inf\{A_p\}$ implica che $x \in A_p \Leftrightarrow F(x) \geq p$

4.2 11/11/2019

<u>**Osservazione**</u>

Per la costruzione dell'inversa generalizzata abbiamo scelto di assegnare l'estremo inferiore, ma è possibile costruirla come $\overleftarrow{F}(p)^+ = \sup\{x \in \mathbb{R} : F(x) \leq p\}$, oppure come mistura di $\overleftarrow{F}(p)$, assegnando $\alpha \in [0,1]$

e $\overleftarrow{F}(p)^{(\alpha)} = \alpha \overleftarrow{F}(p) + (1-\alpha)\overleftarrow{F}(p)^+$, per i punti in cui è definita a tratti. Inoltre si dimostra che $\overleftarrow{F}(p)^+$ è continua a destra.

Teorema

Sia $F(\cdot)$ una FdR e U un n.a. con determinazioni possibili in $]0,1[$ e $U \sim Unif(0,1)$. Se consideriamo $X' = \overleftarrow{F}(U)$, X' ha come FdR $F(\cdot)$.

Quindi data una FdR possiamo costruire un n.a. con quella $F(\cdot)$, se lo consideriamo come una trasformata di un n.a. con distribuzione uniforme tramite $\overleftarrow{F}(\cdot)$.

Dimostrazione

Si consideri $x \in \mathbb{R}$ fissato, $F_{X'}(x) = Pr(X' \leq x) = Pr(\overleftarrow{F}(U) \leq x)$. Per la proprietà iii $F(x) \geq U \Leftrightarrow x \geq \overleftarrow{F}(U)$, $\forall U \in]0,1[$, quindi gli eventi $(F(x) \geq U) = (x \geq \overleftarrow{F}(U))$ sono equivalenti ed hanno la stessa probabilità. Troviamo quindi che $Pr(F(x) \geq U) = Pr(x \geq \overleftarrow{F}(U)) = F_U(F(x)) = F(x)$.

Teorema

Sia $X \geq 0$ con $F(\cdot)$ come FdR e $\mathbb{E}(X) < +\infty$, allora $\mathbb{E}(X) = \int_0^{+\infty} x dF_X(x) = \int_0^{+\infty} \overline{F}_X(x)dx$, con $\overline{F}_X(x) = 1 - F_x(x)$.

Dimostrazione

$\mathbb{E}(X) = \int_0^{+\infty} x dF_X(x) = \int_0^{+\infty} (\int_0^x dy) dF_X(x) = \int\int_D dy dF_X(x)$, con D l'insieme del piano detto dominio normale. Vale il teorema di Fubini, quindi si può scambiare l'ordine di integrazione, ovvero $\int_0^{+\infty}(\int_0^{+\infty} dF_X(x))dy = \int_0^{+\infty}(1 - F_X(y))dy = \int_0^{+\infty} \overline{F}_X(y)dy$.

Misure di Rischio

É detto *Rischio* il n.a. $X \geq 0$, che rappresenta l'importo che un assicuratore deve pagare in un'epoca futura, per risarcire un sinistro, contratto o portafoglio in un fissato intervallo di tempo. Il flusso è $(-X, T)$.

Dato un rischio X, gli associamo un numero maggiore o uguale a zero $\rho(X)$, che ne dia una misura della rischiosità.

Altra definizione: dato $\mathcal{X}$ insieme di rischi, con $X \in \mathcal{X}$ gli importi futuri, chiamiamo *misura di rischio* in $\mathcal{X}$, un funzionale $\rho : \mathcal{X} \to \mathbb{R}_+ \cup \{+\infty\}$, tale che, dato $x \in \mathcal{X}$, $\rho(X)$ rappresenti il minimo importo di capitale di cui il detentore della posizione finanziaria deve disporre per rendere accettabile dover pagare X ad un controllore/autorità di controllo interno o esterno, affinché gli sia acconsentito gestire il rischio.

Definizione formale: fissato $\mathcal{X}$ insieme di rischi e $\mathcal{A}$ insieme di accettabilità, ovvero le posizioni accettabili, si definisce misura di rischio $\rho_{\mathcal{A}}(X) = \inf\{\alpha \geq 0 : -X + \alpha \in \mathcal{A}\}$.

Se X rappresenta il risarcimento totale che l'assicuratore deve pagare per un portafoglio, allora $\rho(X)$ è quindi il minimo capitale richiesto dall'autorità di controllo o dal risk manager, per sostenere il risarcimento. Se X rappresenta il risarcimento totale per un fissato contratto assicurativo, l'assicuratore se ne fa carico, se riceve il premio $\Pi(X)$, cioè il minimo capitale affinché sia disposto a pagare per il futuro risarcimento. Perciò i principi di calcolo del premio sono particolari misure di rischio su insiemi di rischi, che rappresentano i risarcimenti totali per ciascun contratto assicurativo, con la proprietà imprescindibile che $\Pi(X) > \mathbb{E}(X)$. In questo modo, a seconda del criterio scelto, si privilegiano dei particolari aspetti della rischiosità.

Osservazione

Alcune misure di rischio godono della proprietà che se $X \overset{d}{=} Y$, allora $\rho(X) = \rho(Y)$, cioè la misura di rischio dipende solo dalla sua valutazione probabilistica, che però potrebbe non essere soddisfacente. Se infatti supponiamo di avere due rischi del tipo $(X > M) = \emptyset$, $(Y > M) \neq \emptyset$, con $Pr(Y > M) = 0$. I due numeri aleatori hanno la stessa valutazione probabilistica, però Y potrebbe avere determinazioni maggiori di M, di conseguenza i due n.a. non sono ugualmente rischiosi.

Alcune classi di misure di rischio

1 Misure coerenti di rischio

Dato $\mathcal{X}$ insieme di rischi, una misura $\rho(\cdot)$ in $\mathcal{X}$ si dice coerente se soddisfa i seguenti assiomi:

∘ $\rho(X + c) = \rho(X) + c$, $\forall X \in \mathcal{X}, c \in \mathbb{R}$ tali che $X + c \in \mathcal{X}$

∘ $\rho(aX) = a\rho(X)$, $\forall X \in \mathcal{X}, a \in \mathbb{R}$ tali che $aX \in \mathcal{X}$

∘ $\rho(X + Y) \leq \rho(X) + \rho(Y)$ $\forall X, Y \in \mathcal{X}$ tali che $X + Y \in \mathcal{X}$

∘ Se $X \leq Y$, ovvero se $\forall \omega \in \mathcal{P}$ $X(\omega) \leq Y(\omega)$, allora $\rho(X) \leq \rho(Y)$, $\forall X, Y \in \mathcal{X}$

Il primo assioma implica che se il rischio viene aumentato o ridotto di un importo certo, è ragionevole aumentare o ridurre la misura di rischio dello stesso importo. Risulta una proprietà importante per interpretare $\rho(X)$ come un capitale: $(-X + \rho(X))$ è accettabile, se e solo se $\rho(-X + \rho(X)) = -\rho(X) + \rho(X) = 0$.

Il secondo assioma è utile per l'indipendenza in termini di unità monetaria, ma è svantaggiosa in termini di liquidità e di dispersione.

Il terzo assioma invece è importante, poiché se la copertura Z, con $\Pi(Z)$, è composta da due risarcimenti parziali $(Z = X + Y)$, si ha che $\Pi(Z) \leq \Pi(X) + \Pi(Y)$, poiché si suppone che con l'aggregazione si abbia una diversificazione, quindi che l'effetto del rischio si riduca.

Il quarto assioma può essere visto come se $Pr(X \leq Y) = 1$, allora $\rho(X) \leq \rho(Y)$.

Si noti che se $\rho(X) = \mathbb{E}(X)$, si riscontra che $\mathbb{E}(\cdot)$ soddisfa tutti gli assiomi, è perciò una misura coerente.

Misura del Rischio: Value at Risk (VaR)

Definizione 1: Dati $X \geq 0$ e $p \in]0,1[$, è detto *Value at Risk* di X a livello p il $VaR[X,p] = \overleftarrow{F}_X(p) = \inf\{x \in \mathbb{R} : F_X(x) \geq p\} = \min\{x \in \mathbb{R} : F_X(x) \geq p\}$

Definizione 2: Dati $\mathcal{X}$ un insieme di rischi e $p \in]0,1[$, è detta misura di rischio del VaR a livello p, il funzionale $\rho_p(X) = VaR[X,p]$, quindi $\rho_p(\cdot) = VaR[\cdot,p] : \mathcal{X} \to \mathbb{R}_+$.

Il VaR permette quindi di usare i quantili come misure di rischio e permette una più facile interpretazione.

Osservazione

Si ha che $Pr(X > VaR[X,p]) = 1 - F_X(VaR[X,p]) = 1 - F_X(\overleftarrow{F}_X(p)) \leq 1 - p$. Quindi la probabilità che la perdita sia superiore al capitale allocato è $1-p$ e in particolare *Solvency II* ha fissato la probabilità di rovina $1 - p = 0.005$.

Inoltre notiamo che:

∘ Il VaR è sempre definito, poiché $\overleftarrow{F}_X(p)$ esiste $\forall X, p$

∘ Il $VaR[X,p]$ è dimensionalmente equivalente a X e quindi facilmente interpretabile

∘ Se $X \overset{d}{=} Y$, allora $VaR[X,p] = VaR[Y,p]$

Proprietà del VaR

Siano $\mathcal{X}$ l'insieme dei rischi e $p \in]0,1[$ e consideriamo la misura di rischio di $\mathcal{X}$ $VaR[\cdot,p]$:

∘ **No Rip-Off:** dato $X \in \mathcal{X}$ e $M \in \mathbb{R}$ certo, se $X \leq M$, allora $VaR[X,p] \leq M$. Si ha quindi che $Pr(X \leq M) = Pr(\Omega) = 1 = F_X(M) > p$ e quindi che $M > \overleftarrow{F}_X(p) = VaR[X,p]$.

Questa proprietà è auspicabile, poiché non c'è motivo di allocare un capitale superiore a M.

○ **Traslatività:** $VaR[X+c,p] = VaR[X,p]+c$, $\forall X \in \mathcal{X}$, $c \in \mathbb{R}$ certo, tale che $X+c \in \mathcal{X}$.
Fissato $x \in \mathbb{R}$, si considera $VaR[X+c,p] \leq x \Leftrightarrow \overleftarrow{F}_{X+c}(p) \leq x \Leftrightarrow F_{X+c}(x) \geq p$. Se guardiamo la FdR troviamo che $F_{X+c}(x) = Pr(X+c \leq x) = Pr(X \leq x-c) = F_X(x-c)$, quindi $F_X(x-c) \geq p \Leftrightarrow x-c \geq \overleftarrow{F}_x(p) = VaR[X,p] \Leftrightarrow x \geq VaR[X,p]+c$. E questo implica che $VaR[X+c,p] \leq x \Leftrightarrow VaR[X,p]+c \leq x$, $\forall x \in \mathbb{R}$. Quindi si ha che $VaR[X+c,p] = VaR[X,p]+c$.

○ **Omogeneità Positiva:** $VaR[aX,p] = aVaR[X,p]$, con $x \in \mathcal{X}$ e $a > 0 \in \mathbb{R}$ tale che $aX \in \mathcal{X}$.

○ **Monotonia:** Se $X \leq Y$ $\forall \omega \in \mathcal{P}$, allora $VaR[X,p] \leq VaR[Y,p]$.
Sia $X \leq Y$, allora $F_X(x) \geq F_Y(x)$, $\forall x \in \mathbb{R}$. Fissato $x \in \mathbb{R}$ consideriamo l'evento $(Y \leq x) = (Y \leq x, X \leq Y) \Rightarrow (X \leq x)$. Dato che $(Y \leq x, X \leq Y) \subseteq (X \leq x) \Rightarrow Pr(Y \leq x, X \leq Y) \leq Pr(X \leq x) \Rightarrow Pr(X \leq x) \geq Pr(Y \leq x) \Leftrightarrow F_X(x) \geq F_Y(x)$.
Dato che $F_X(x) \geq F_Y(x)$, allora $VaR[X,p] \leq VaR[Y,p]$, cioè se la FdR di X sta sempre sopra alla FdR di Y, allora il capitale da allocare risulta inferiore a parità di p.
Consideriamo $A_p^{(X)} = \{x \in \mathbb{R} : F_X(x) \geq p\}$ e $A_p^{(Y)} = \{x \in \mathbb{R} : F_Y(x) \geq p\}$. Essendo $p \in]0,1[$, sono entrambi non vuoti ed inferiormente limitati. Inoltre si rileva che $A_p^{(Y)} \subset A_p^{(X)}$, poiché perso $x \in A_p^{(Y)}$, allora $F_X(x) \geq F_Y(x) \Rightarrow F_X(x) \geq p \Rightarrow x \in A_p^{(X)}$.
Quindi si ha che $\overleftarrow{F}_Y(p) = \inf\{A_p^{(Y)}\} \geq \inf\{A_p^{(X)}\} = \overleftarrow{F}_X(p)$ e quindi che $\overleftarrow{F}_Y(p) \geq \overleftarrow{F}_X(p)$ e che $VaR[Y,p] \geq VaR[X,p]$.

○ La misura $VaR[\cdot,p]$ non verifica la proprietà di introdurre un caricamento di sicurezza. Si dice che una misura di rischio $\rho(\cdot)$ introduce un caricamento di sicurezza se $\forall X \in \mathcal{X}$, si ha che $\rho(X) \geq \mathbb{E}(X)$.
<u>Controesempio:</u> sia $X \geq 0$ dotato di $F_X(\cdot)$ e con $\mathbb{E}(X) < +\infty$. Possiamo calcolare $F_X(\mathbb{E}(X))$ e supponiamo sia in $]0,1[$. Sia $p^* = F_X(\mathbb{E}(X))$ e prendiamo un $p < p^*$, allora $VaR[X,p] = \overleftarrow{F}_X(p) \leq \overleftarrow{F}_X(p^*) = \overleftarrow{F}_X(F(\mathbb{E}(X))) \leq \mathbb{E}(X) \Rightarrow VaR[X,p] \leq \mathbb{E}(X)$, quindi si dimostra la mancanza di caricamento. Preso invece $p > p^*$ si ha che $VaR[X,p] > \mathbb{E}(X)$, quindi per p grandi si ha un caricamento di sicurezza.
<u>Osservazione:</u> sia X il rischio, fissato $0 < \epsilon < 1$, si pone come probabilità di perdita $Pr(P-X<0) \leq \epsilon \Leftrightarrow F_X(P) \geq 1-\epsilon$ e per il principio del percentile $\Pi(X) = \inf\{P \in \mathbb{R} : F_X(P) \geq 1-\epsilon\} = \min\{P \in \mathbb{R} : F_X(P) \geq 1-\epsilon\} = \overleftarrow{F}_X(1-\epsilon) = VaR[X,1-\epsilon]$ e per $1-\epsilon$ grande, si ha il caricamento di sicurezza.

○ La misura $VaR[\cdot,p]$ non gode neanche della proprietà di subadditività, ovvero $\forall X,Y$ non si ha che $VaR[X+Y,p] \leq VaR[X,p]+VaR[Y,p]$.
<u>Controesempio:</u> consideriamo una partizione con tre eventi elementari $\mathcal{P} = \{\omega_1,\omega_2,\omega_3\}$ e siano X e Y definiti su $\mathcal{P}$, tale che $\exists Pr(\omega_i)$ per ogni i. Poniamo $p = 0.6$, dai dati in tabella vediamo che $\overleftarrow{F}_X(0.6) = 20 = \overleftarrow{F}_Y(0.6) = VaR[X,0.6] = VaR[Y,0.6]$. Mentre $\overleftarrow{F}_{X+Y}(0.6) = 60 = VaR[X+Y,0.6] > VaR[X,0.6]+VaR[Y,0.6] = 40$.
La mancanza di questa proprietà è un problema, per cui è stato a lungo discusso se tenere questa misura di rischio in *Solvency II*, in quanto disincentiva l'impresa ad investire in capitale allocato ed incentiva a rivolgersi a terzi.

Tabella Controesempio

	X	Y	$X+Y$	$Pr(\omega_i)$
ω_1	10	10	20	0.2
ω_2	20	40	60	0.4
ω_3	40	20	60	0.4

Osservazione: sia $F_X(x) \geq F_Y(x)\ \forall x \in \mathbb{R}$, allora $VaR[X,p] < VaR[Y,p]$, con $p \in]0,1[$ fissato in maniera arbitraria. Si dimostra che $VaR[X,p] < VaR[Y,p]$ vale anche con $p \in [0,1]$.
∘ $\boxed{\text{p=0}}$ $\overleftarrow{F}_X(0) = -\infty = \overleftarrow{F}_Y(0)$
∘ $\boxed{\text{p=1}}$ Si ha che $A_1^{(X)} = \{x \in \mathbb{R} : F_X(x) = 1\}$ e $A_1^{(Y)} = \{x \in \mathbb{R} : F_Y(x) = 1\}$, con $A_1^{(Y)} \subset A_1^{(X)}$, poiché $F_X(\cdot) \geq F_Y(\cdot)$ e quindi abbiamo che:
- $A_1^{(X)} = \emptyset$ e $A_1^{(Y)} = \emptyset \quad \Rightarrow \quad \overleftarrow{F}_X(1) = +\infty = \overleftarrow{F}_Y(1)$
- $A_1^{(X)} \neq \emptyset$ e $A_1^{(Y)} \neq \emptyset$, allora entrambi sono inferiormente limitati e $A_1^{(Y)} \subset A_1^{(X)} \quad \Rightarrow \quad \overleftarrow{F}_Y(1) \geq \overleftarrow{F}_X(1)$
- $A_1^{(X)} \neq \emptyset$ e $A_1^{(Y)} = \emptyset$, allora $\overleftarrow{F}_Y(1) = +\infty \geq \overleftarrow{F}_X(1) \in \mathbb{R}$

Quindi l'ordinamento vale su $\mathbb{R}^*$, $\forall p \in [0,1]$.
Inoltre vale il viceversa, ovvero se $\overleftarrow{F}_X(p) \leq \overleftarrow{F}_Y(p)$ per ogni $p \in [0,1]$, allora $F_X(x) \geq F_Y(x)$ per ogni $x \in \mathbb{R}$.
Fissato $x \in \mathbb{R}$, consideriamo $F_Y(x) \in [0,1]$, possiamo calcolare $\overleftarrow{F}_X(F_Y(x)) \leq \overleftarrow{F}_Y(F_Y(x))$, per ogni p. Allora $\overleftarrow{F}_X(F_Y(x)) \leq \overleftarrow{F}_Y(F_Y(x)) \leq x \Rightarrow F_Y(x) \leq F_X(x)$, e per l'arbitrarietà di x vale $\forall x \in \mathbb{R}$.

Dunque Y risulta più rischioso di X poiché si ha una maggiore massa sulle code e la relazione è detta di dominanza stocastica del primo ordine, cioè $X \preceq_{ST} Y$.

Capitale di rischio basato sul VaR

L'assicuratore è esposto al rischio che i premi incassati non siano sufficienti per fronteggiare i risarcimenti e deve perciò attuare un sistema di *risk management* tramite la tariffazione, i caricamenti di sicurezza e disponendo di adeguate riserve tecniche, oppure trasferendo parte del rischio ad un riassicuratore o ad *alternative risk transfer*.

Sia fissato un portafoglio su un arco di tempo mono annuale e siano X il risarcimento totale del portafoglio e R il capitale allocato al portafoglio. Consideriamo P il montepremi dei premi puri, tale che $P = \mathbb{E}(X) + m = \sum_i P_i^{(e)} + \sum_i m_i$. Si definisce l'evento rovina $(X > R + P)$ e si parla di *solvibilità in senso assoluto*, se l'evento rovina risulta impossibile $\forall X$. Il concetto risulta insensato, in quanto la quantità a destra dovrebbe essere molto grande, in modo da coprire qualunque possibile scenario. Perciò si parla di *solvibilità in senso probabilistico* e si definisce il requisito di solvibilità $R + P$ tale che $Pr(X > R + P) \leq 1 - p \Leftrightarrow F_X(R + P) \geq p$, con $p \in]0,1[$. Se $R + P = \min\{x \in \mathbb{R} : F_X(x) \geq p\} = VaR[X,p]$, allora il VaR rappresenta il requisito minimo di solvibilità. Dunque se P è dato, si ha che $R = VaR[X - P, p] = VaR[X,p] - (\mathbb{E}(X) + m)$ ed R è detto *capitale di rischio basato sul* VaR. E questo, sommato ai premi, è il minimo valore che consente di rendere soddisfatto il requisito di solvibilità.
Dunque se X è dato, l'equazione $R + \mathbb{E}(X) + m = VaR[X,p]$ ha come incognite R e m.
Studiamo l'effetto della riassicurazione sul capitale allocato: abbiamo X, P, $VaR[X,p] = P + R \Leftrightarrow R = VaR[X,p] - P$, $X = \Gamma + X_{CED} = \Gamma + X_R$, $P - P_R$ e $R' = VaR[\Gamma,p] - (P - P_R)$.
Essendo $\Gamma \leq X$, allora $VaR[\Gamma,p] \leq VaR[X,p]$, tuttavia però anche $(P - P_R) < P$, perciò dobbiamo analizzare nelle diverse forme riassicurative quanto diminuiscono i vari componenti.
∘ **Quota-Share:** $\Gamma = aX$ con $0 < a < 1$, $X_R = (1-a)X$ e $P_R = (1-a)P$, proporzionale al rischio ceduto. Il requisito di solvibilità è $VaR[\Gamma,p] = VaR[aX,p] = aVaR[X,p]$. Mentre il montepremi è $(P - P_R) = P - (1-a)P = aP$. Troviamo quindi che $R' = aVaR[X,p] - aP = a(VaR[X,p] - P) = aR < R$, quindi $R' < R$.
Osservazione: in realtà P_R risulta più piccolo di $(1-a)P$, poiché si tiene conto della commissione riconosciuta dal riassicuratore, quindi a maggior ragione $R' < R$.

4.3 15/11/2019

Altre misure di rischio

Il $VaR[\cdot, p]$, oltre a dare il minimo valore che soddisfa il requisito di solvibilità, viene utilizzato al fine di introdurre un caricamento per il rischio nella valutazione delle passività a bilancio. *Solvency II* e *IFRS17* sanciscono che le passività, quali Riserva Premi e Riserva Sinistri, debbano essere scritte a bilancio, indicando una $BestEstimate$ e un margine o caricamento per il rischio.

Esempio: Sia X il risarcimento totale o anche il valore attuale dei flussi dei pagamenti futuri per sinistri già verificati, ma non completamente risarciti. A fronte di X nel bilancio si ha la riserva $BestEstimate$ ($\mathbb{E}(X)$), il caricamento per il rischio (*Risk Margin* o *Market Value Margin*, oppure *Risk Adjustment* in $IFRS17$). Uno dei metodi per calcolare il *Risk Adjustment* è $R = VaR[X, p] - \mathbb{E}(X)$, con $Pr(X > \mathbb{E}(X) + R) = Pr(X > VaR[X, p]) \leq 1 - p$.

Il VaR è dunque una misura molto utile per controllare la probabilità di rovina, tuttavia è stata criticata poiché non fornisce alcun controllo nel caso in cui $X > VaR[X, p]$, motivo per cui sono stati introdotti i seguenti indicatori:

◦ ***Conditional Tail Expectation (Tail-VaR):*** data la probabilità $p \in]0, 1[$, $CTE[X, p] = \mathbb{E}(X|X > VaR[X, p])$ è il valore atteso di X, condizionato all'ipotesi che superi il VaR.

◦ ***Mean Excess Function:*** è l'eccedenza attesa rispetto al VaR, nell'ipotesi che lo superi, ovvero $ex(VaR[X, p]) = \mathbb{E}(X - VaR[X, p]|X > VaR[X, p])$.

◦ ***Expected Shortfall (Stop-Loss Premium):*** $ES[X, p] = \mathbb{E}((X - VaR[X, p])_+)$, cioè il valore atteso della parte positiva, che è definita con $X - VaR[X, p]$ se maggiore di zero, 0 altrimenti.

É detta anche *Stop-Loss Premium*, poiché: sia X il risarcimento totale di portafoglio, sia $L > 0$ la priorità, $X = \Gamma + X_R$ dove $\Gamma = \min\{X, L\}$, $X_R = \max\{X - L, 0\} = (X - L)_+$. Quindi $\mathbb{E}(X_R) = \mathbb{E}((X - L)_+)$ è il premio equo *Stop-Loss*.

◦ ***Tail-VaR:*** $TVaR[X, p] = \frac{1}{1-p}\int_p^1 VaR[X, u]du$, cioè la media integrale dei valori del VaR per valori di p elevati. Vi sono analogie con CTE, perrciò vengono chiamati entrambi $Tail - VaR$.

Alcuni legami tra le misure di rischio:

◦ $ES(X, p) = \mathbb{E}((X - v)_+) = Pr(X \leq v)\mathbb{E}((X - v)_+|X \leq v) + Pr(X > v)\mathbb{E}((X - v)_+|X > v) = Pr(X \leq v)\mathbb{E}(0|X \leq v) + Pr(X > v)\mathbb{E}(X - v|X \geq v) = Pr(x > VaR[X, p])\mathbb{E}(X - VaR[X, p]|X > VaR[X, p]) = (1 - F_X(VaR[X, p]))ex(VaR[X, p])$. Quindi abbiamo trovato che $ex(VaR[X, p]) = \frac{ES(VaR[X,p])}{\overline{F}_X(VaR[X,p])}$.

◦ $CTE[X, p] = \mathbb{E}(X|X > v) + v - v = \mathbb{E}(X - v|X > v) + v = ex(VaR[X, p]) + VaR[X, p]$. Quindi abbiamo che $CTE[X, p] = \frac{ES(X,p)}{\overline{F}_X(VaR[X,p])} + VaR[X, p]$

Osservazione: sia Y n.a. con F_Y come FdR, se $U \sim Unif(0, 1)$, allora $\overleftarrow{F}_Y(U) \overset{d}{=} Y$. Quindi questo equivale a dire che $\mathbb{E}(Y) = \mathbb{E}(\overleftarrow{F}_Y(U))$. Quindi $\mathbb{E}(\overleftarrow{F}_Y(U)) = \int_0^1 \overleftarrow{F}_Y(u)f_U(u)du = \int_0^1 \overleftarrow{F}_Y(u) \cdot 1du = \int_0^1 \overleftarrow{F}_Y(u)du$.

◦ $ES[X, p] = \mathbb{E}((X - v)_+) = \mathbb{E}((\overleftarrow{F}_X(U) - v)_+) = \int_0^1 (\overleftarrow{F}_X(u) - v)_+du = \int_0^1 (\overleftarrow{F}_X(u) - \overleftarrow{F}_X(p))_+du$ e quindi possiamo avere i seguenti casi:

- $u \leq p$, allora per monotonia $\overleftarrow{F}_X(u) \leq \overleftarrow{F}_X(p)$, quindi segue che $(\overleftarrow{F}_X(u) - \overleftarrow{F}_X(p))_+ = 0$
- $u > p$, allora per monotonia $\overleftarrow{F}_X(u) > \overleftarrow{F}_X(p)$, quindi segue che $(\overleftarrow{F}_X(u) - \overleftarrow{F}_X(p))_+ = \overleftarrow{F}_X(u) - \overleftarrow{F}_X(p)$.

Quindi $ES[X, p] = \int_0^p (\overleftarrow{F}_X(u) - \overleftarrow{F}_X(p))_+du + \int_p^1 (\overleftarrow{F}_X(u) - \overleftarrow{F}_X(p))_+du = \int_0^p 0du + \int_p^1 (\overleftarrow{F}_X(u) - \overleftarrow{F}_X(p))du = \int_p^1 \overleftarrow{F}_X(u)du - \overleftarrow{F}_X(p)(1-p)$. Troviamo quindi che $\int_p^1 \frac{\overleftarrow{F}_X(u)du}{1-p} = \frac{ES[X,p]}{1-p} + VaR[X, p]$, ne segue che $TVaR[X, p] = \frac{ES[X,p]}{1-p} + VaR[X, p]$.

Si nota che il $TVaR$ differisce dal CTE per il denominatore, ovvero $\overline{F}_X(VaR[X,p]) = 1 - F_X(\overleftarrow{F}(p)) \leq 1-p$. Le due misure quindi coinciderebbero, se si prendesse p non in corrispondenza dei punti di salto e quindi $CTE[X,p] \geq TVaR[X,p] \geq VaR[X,p]$ in generale. Si ha l'uguaglianza nei punti di continuità: sia $p \in]0,1[$ e poniamo $\lambda \stackrel{\Delta}{=} \overleftarrow{F}_X(p) \in \mathbb{R}$. Sia F_X continua in λ, allora si dimostra che $F_X(\overleftarrow{F}_X(p)) = F_X(\lambda) = p$. Si ipotizza per assurdo che $F_X(\lambda) > p$, per definizione di continuità, $\forall I_{F_X(\lambda)}, \exists I_\lambda : \forall x \in I_\lambda, F_X(x) \in I_{F_X(\lambda)}$.
Scegliamo un opportuno intervallo $I_{F_X(\lambda)} =]F_X(\lambda) - \epsilon, F_X(\lambda) + \epsilon[$ tale che $F_X(\lambda) > p$, a cui è associato il corrispondente intorno completo per la continuità I_λ, tale che $\forall x \in I_\lambda$, $F_X(x) > p$. Sia $\overline{x} \in I_\lambda$, tale che $\overline{x} < \lambda$, allora $F_X(\overline{x}) > p$, quindi $\overline{x} \in A_p$, ma assurdo, poiché $\overline{x} < \lambda = \inf\{A_p\}$. Quindi si è dimostrato che se la funzione $F_X(\cdot)$ è continua, le due misure coincidono in quel punto.
In generale $CTE[X,p] \geq TVaR[X,p]$, se F_X è continua $CTE[X,p] = TVaR[X,p]$, $\forall p \in]0,1[$, altrimenti $CTE[X,p] > TVaR[X,p]$, se $p \in]F(\overline{x}_0), F_X(x_0)[$, con $F(\overline{x}_0) \neq F_X(x_0)$.

Proprietà del $TVaR$

Sia $\mathcal{X}$ l'insieme dei rischi, sia $p \in]0,1[$ e sia $tVaR[\cdot,p]$ la misura di rischio definita su $\mathcal{X}$, allora:
- É una misura coerente di rischio, quindi verifica anche la subadditività, che non vale per $VaR[\cdot,p]$ e non è detto che valga anche per $CTE[X,p]$
- Verifica il *no Rip-Off*
- Non introduce caricamenti ingiustificati
- Introduce il caricamento di sicurezza, a differenza del $VaR[\cdot,p]$

Capitale di rischio basato sul $TVaR$

Consideriamo un fissato portafoglio su un arco di tempo mono annuale, sia X il risarcimento totale, sia P il montepremi netto, con $P = \mathbb{E}(X) + m$, sia R la dotazione di capitale allocato al portafoglio e il requisito di solvibilità è pari a $Pr(X > R + P) \leq 1-p$ con p "grande" e $R = TVaR[X,p] - P$. Il $TVaR$ permette di far fronte ai risarcimenti superiori al VaR entro il valore atteso dei risarcimenti, condizionatamente all'ipotesi che superino il VaR.

Osservazione

Finora abbiamo considerato solo rischi puri $X \geq 0$, ma se si considerano i n.a. G come grandezze del tipo Profit-Loss, cioè saldo tra entrate e uscite, gli eventi possibili sono $G > 0$ o $G \leq 0$. Possiamo avere due tipologie di problemi, per cui cambia il calcolo di requisito di capitale basto sul VaR:
1 $G < 0$, allora R è tale che $Pr(R + G < 0) \leq 1-p \Leftrightarrow Pr(G < -R) \leq 1-p \Leftrightarrow Pr(-G > R) \leq 1-p \Leftrightarrow 1 - Pr(-G \leq R) \leq 1-p \Leftrightarrow Pr(-G \leq R) \geq p \Leftrightarrow F_{-G}(R) \geq p \Leftrightarrow F_L(R) \geq p$, con $L = -G$ e quindi $R = \min\{x \in \mathbb{R} : F_L(R) \geq p\} = \overleftarrow{F}_L(p)$.
2 $G < \mathbb{E}(G)$, allora R è tale che $Pr(R + G < \mathbb{E}(G) \leq 1-p \Leftrightarrow Pr(G < \mathbb{E}(G) - R) \leq 1-p \Leftrightarrow Pr(-G > -\mathbb{E}(G) + R) \leq 1-p \Leftrightarrow 1 - Pr(-G \leq -\mathbb{E}(G) + R) \leq 1-p \Leftrightarrow Pr(-G \leq -\mathbb{E}(G) + R) \geq p \Leftrightarrow F_{-G}(R - \mathbb{E}(G)) \geq p \Leftrightarrow F_L(R - \mathbb{E}(G)) \geq p$, con $L = -G$ e quindi $R - \mathbb{E}(G) \geq \overleftarrow{F}_L(p)$, quindi $R = \overleftarrow{F}_L(p) + \mathbb{E}(G)$.
Quindi come si interpreta il $VaR[G,p]$, qualora sia negativo?
- Se $VaR[G,p] = \overleftarrow{F}_L(p)$ sulla perdita, risulta un valore positivo ed è interpretabile come capitale
- Se $VaR[G,p] = \overleftarrow{F}_G(p)$ sul guadagno, risulta un valore negativo, ed allora lo interpreto come quantile.

4.4 18/11/2019

Misura di rischio di Wang

Sia $\mathcal{X}$ l'insieme dei rischi, sia $\xi > 0 \in \mathbb{R}$, è detta *misura di rischio di Wang*, il funzionale $\rho_\xi : \mathcal{X} \to \mathbb{R}$, tale che $\rho_\xi(X) = \int_0^{+\infty} (\overline{F}_X(x))^{\frac{1}{\xi}} dx$, con $\overline{F}_X(x) = F_X(x)$.

Proprietà

Dati $X \geq 0$, dotato di F_X, e $\xi > 0 \in \mathbb{R}$, posto $F_X^{(\xi)}(x) = 1 - (\overline{F}_X(x))^{\frac{1}{\xi}} = 1 - (1 - F_X(x))^{\frac{1}{\xi}}$, cioè il complemento a 1 della funzione integranda, si ha che:
(i) $F_X^{(\xi)}(\cdot)$ è un Funzione di Ripartizione
(ii) $F_X^{(\xi)}(\cdot)$ è una valutazione coerente di X

Inciso

Sia $f^*(x) = x^\alpha$, con $\alpha > 0 \in \mathbb{R}$, cioè una potenza ad esponente reale. Consideriamo la funzione nell'intervallo $x \in [0, 1]$, la f^* è definita in $\mathbb{R}_+/\{0\}$, poiché non definita in 0, ma se consideriamo $\lim_{x\to 0} f^*(x) = 0$, quindi si estende per continuità la funzione:

$$f^*(x) = \begin{cases} x^\alpha & x \in]0, 1] \\ 0 & x = 0 \end{cases}$$

Dimostrazioni

(i) $F_X^{(\xi)}$ è un FdR se e solo se rispetta le proprietà di una FdR:
∘ $F_X^{(\xi)} : \mathbb{R} \to [0, 1]$, quindi F_X è definita in $\mathbb{R}$, tale che $\forall x \in \mathbb{R}\ F_X(x) \in [0, 1]$, quindi anche $1 - F_X(x) \in [0, 1]$. Per l'osservazione precedente anche la potenza è inclusa in $[0, 1]$. Quindi $F_X^{(\xi)}(\cdot) : \mathbb{R} \to [0, 1]$.
∘ $F_X^{(\xi)}$ è monotona non decrescente. Se F_X è monotona non decrescente, allora $1 - F_X$ è monotona non crescente. La potenza a esponente reale è crescente, quindi la composizione $(1 - F_X(\cdot))^{\frac{1}{\xi}}$ è monotona non crescente. Ma allora $1 - (1 - F_X(\cdot))^{\frac{1}{\xi}}$ è monotona non decrescente.
∘ $\lim_{x\to -\infty} F_X^{(\xi)}(x) = 0$. Si ha che $F_X \to 0$, quindi $1 - F_X \to 1$ e quando la base tende a 1, la potenza reale, tende a 1. Quindi per il teorema del limite della funzione composta, si ha che $(1 - F_X(x))^{\frac{1}{\xi}} \to 1$ e quindi $1 - (1 - F_X(x))^{\frac{1}{\xi}} \to 0$.
∘ $\lim_{x\to +\infty} F_X^{(\xi)}(x) = 1$. Si ha che $F_X \to 1$, quindi $1 - F_X \to 0$ e quando la base tende a 0, la potenza reale, tende a 0. Quindi per il teorema del limite della funzione composta, si ha che $(1 - F_X(x))^{\frac{1}{\xi}} \to 0$ e quindi $1 - (1 - F_X(x))^{\frac{1}{\xi}} \to 1$.
∘ $F_X^{(\xi)}$ è continua a destra. F_X è continua a destra, quindi $1 - F_X$ è continua a destra. La potenza reale è una funzione continua, quindi la composta è continua a destra, quindi $1-(1-F_X(x))^{\frac{1}{\xi}}$ è continua a destra.

(ii) $F_X^{(\xi)}$ è una valutazione coerente per X, ovvero può essere scelta per valutare X coerentemente.
Affinché lo sia devono essere vere le seguenti condizioni:
∘ Se $(X \leq x) = \emptyset$, allora $F_X^{(\xi)}(x) = 0$, poiché $F_X(x) = 0 \Rightarrow (1 - F_X(x))^{\frac{1}{\xi}} = 1 \Rightarrow F_X^{(\xi)}(x) = 1 - 1 = 0$
∘ Se $(X \leq x) = \Omega$, allora $F_X^{(\xi)}(x) = 1$, poiché $F_X(x) = 1 \Rightarrow (1 - F_X(x))^{\frac{1}{\xi}} = 0 \Rightarrow F_X^{(\xi)}(x) = 1 - 0 = 1$
∘ Se $(X \leq x) = (X \leq y)$, con $x \neq y$, allora $F_X(x) = F_X(y) \Rightarrow F_X^{(\xi)}(x) = F_X^{(\xi)}(y)$
Inoltre notiamo che la funzione integranda $(\overline{F}_X(x))^{\frac{1}{\xi}} = 1 - F_X^{(\xi)}(x) = \overline{F}_X^{(\xi)}(x)$, cioè è pari alla funzione delle code di $F_X^{(\xi)}$.

Osservazioni

(1) $\rho_\xi(X) = \int_0^{+\infty} (\overline{F}_X(x))^{\frac{1}{\xi}} dx = \int_0^{+\infty} \overline{F}_X^{(\xi)}(x) dx = \mathbb{E}_\xi(X)$

② Sia $\xi > 1$, si confrontano le due valutazioni $F_X \Rightarrow \mathbb{E}(X) = \int_0^{+\infty} \overline{F}_X(x)dx$ e $F_X^{(\xi)} \Rightarrow \mathbb{E}_\xi(X) = \int_0^{+\infty} \overline{F}_X^{(\xi)}(x)dx$.
Si consideri arbitrariamente un x fissato e si ha $1 - F_X(x) \leq (1 - F_X(x))^{\frac{1}{\xi}}$, con se $F_X(x) \notin \{0, 1\}$ allora $1 - F_X(x) < (1 - F_X(x))^{\frac{1}{\xi}}$. Troviamo quindi che $\forall x \in \mathbb{R}$ $\int_0^{+\infty} \overline{F}_X(x)dx \leq \int_0^{+\infty} \overline{F}_X^{(\xi)}(x)dx$.
Troviamo che $\mathbb{E}(X) < \rho_\xi(X)$, nel caso in cui X non sia n.a. certo o quasi certo, e quindi ρ_ξ è una misura di rischio che introduce un caricamento di sicurezza.
Inoltre troviamo che $F_X^{(\xi)}$ è una trasformata di F_X che dà più peso alle code , poiché $\overline{F}_X(x) \leq \overline{F}_X^{(\xi)}(x)$, $\forall x \in \mathbb{R}$, ovvero a parità di confidenza, F_X risulta una valutazione più rischiosa.
③ Se $\mathcal{X}$ è l'insieme dei risarcimenti totali per polizza, allora potremmo utilizzare la misura di rischio come un principio di calcolo del premio, in quanto assegna un caricamento di sicurezza e $\Pi(X) = \rho_\xi(X)$, con $\xi > 1$, ed è detto principio di calcolo del premio di Wang o *Risk-Adjusted Principle*.
Si riscontra che $\Pi(X) = \rho_\xi(X) = \mathbb{E}_\xi(X)$ segue il principio della speranza matematica, o in particolare il principio di equità, considerando F_X la valutazione realistica e $F_X^{(\xi)}$ la valutazione prudenziale.

Misura di Wang Generalizzata

Sia $\mathcal{X}$ l'insieme dei rischi, sia $g(\cdot)$ definita in $[0, 1]$, monotona non decrescente, tale che $g(0) = 0$ e $g(1) = 1$. Si dice *misura di Wang Generalizzata* con funzione di distorsione $g(\cdot)$, il funzionale in $\mathcal{X}$ tale che $\rho_g(X) = \int_0^{+\infty} g(\overline{F}_X(x))dx$.
Se $g(t) = t^{\frac{1}{\xi}}$, con $\xi > 0$ si ottiene la misura di Wang ed in generale non si ha che $1 - g(\overline{F}_X)$ sia una Funzione di Ripartizione.

Proprietà

Sia ρ_g, definita in $\mathcal{X}$ a valori in $\mathbb{R}$, una misura di rischio di Wang generalizzata, allora verifica il *No Rip-Off*, la *traslatività*, l'*omogeneità positiva*, la *monotonia*, non introduce caricamenti di sicurezza ingiustificati ed è *subadditiva*, se e solo se $g(\cdot)$ è concava. Quindi se $g(\cdot)$ è concava, allora ρ_g è una misura coerente di rischio. Introduce un caricamento di sicurezza se e solo se $g(t) > t$, $\forall t \in [0, 1]$.
Essendo la misura di Wang un elemento di questa classe, verifica la medesime proprietà e in particolare:
∘ ρ_ξ è subadditiva $\Leftrightarrow \xi > 1$
∘ $\rho_x i$ introduce un caricamento di sicurezza $\Leftrightarrow \xi > 1$.

5 Valutazione Mediante Simulazione Stocastica

Supponiamo di avere la valutazione probabilistica di un ente aleatorio e siamo interessati a quella di un suo trasformato, ci ritroviamo di fronte ad un'operazione formalmente possibile, ma operativamente spesso complessa.
Sia X un n.a. con FdR F_X e $Y = \psi(X)$ una sua trasformata, possiamo calcolare $\mathbb{E}(Y)$, $\mathbb{E}(Y^k)$ e F_Y. Queste stime però si complicano, nel caso in cui $\underline{X} = (X_1, \dots, X_n)$ sia un vettore aleatorio e $\underline{Y} = \psi(\underline{X})$, oppure in cui siamo interessati alla trasformata del processo stocastico $\{X_1, X_2, \dots\} \sim \mathcal{F}$, $Y = \psi(X_1, X_2, \dots)$.
Ad esempio considerato $X = \sum_i^N Y_i = \psi(N, Y_1, \dots)$, se assegniamo la base tecnica $(Pr(N = n)_{n \geq 0}, F_Y)$ e l'ipotesi di distribuzione composta, allora possiamo calcolare $\mathbb{E}(X)$, $V(X)$ e $F_X(x) = \sum Pr(N = n)F_Y^{*(n)}(x)$, con quest'ultima che però in generale è incognita. Per questo motivo si ricorre alla valutazione tramite simulazione stocastica.
La simulazione di un ente aleatorio di legge assegnata, corrisponde alla sostituzione dell'ente aleatorio originale, con un nuovo ente, che riproduca lo stesso stato di incertezza dell'ente originale. Ovvero un

ente al quale riteniamo di poter attribuire la medesima valutazione probabilistica.
Quindi abbiamo X un n.a. con FdR F_X e $X' \stackrel{d}{=} X$, quindi $X' \sim F_X$, oppure $\underline{X} \sim F_{\underline{X}}$ e $\underline{X}' \stackrel{d}{=} \underline{X}$, quindi $\underline{X}' \sim F_{\underline{X}}$.
Se X' simula X e X' è osservabile con determinazione x', allora si dice che x' è una determinazione simulata di X. Sia $Y = \psi(\underline{X})$ e $\underline{X} \sim F_{\underline{X}}$, se $\underline{X}'$ simula $\underline{X}$, allora $Y' = \psi(\underline{X}')$ simula Y. Se $\underline{X}'$ è osservabile con determinazioni $\underline{x}' = (x'_1, \dots, x'_n)$, allora $y' = \psi(\underline{x}')$ è determinazione simulata di Y.
Siano $\underline{X}'_1, \dots, \underline{X}'_m$ IID, con ogni $\underline{X}'_i$ che simula $\underline{X}$, allora $Y'_1 = \psi(\underline{X}'_1), \dots, Y'_m = \psi(\underline{X}'_m)$ sono IID e Y'_i simula Y, ovvero $Y'_i \stackrel{d}{=} Y$. In particolare possiamo vedere $Y'_1, \dots, Y'_m$ come un campione di Y, dal quale possiamo costruire i momenti tramite gli stimatori campionari $\overline{Y}'$ e S^2.
Se $\underline{X}'_1, \dots, \underline{X}'_m$ sono osservabili con determinazioni $\underline{x}'_1, \dots, \underline{x}'_m$, allora $y'_1 = \psi(\underline{x}'_1), \dots, y'_m = \psi(\underline{x}'_m)$ è il campione osservato, da cui otteniamo $\widehat{\mathbb{E}}(Y) = \overline{y}$, $\widehat{V}(Y)$ e $\widehat{F}_Y(y) = \frac{1}{m} card\{y'_i : y'_i \leq y\}$.

<u>**Simulazione di un numero aleatorio $X \sim F_X$**</u>

Sia U un n.a. con determinazioni in $]0,1[$ e $U \sim Unif(0,1)$, allora $X' = \overset{\leftarrow}{F}_X(U) \stackrel{d}{=} X$. Se U è osservabile con determinazione u, allora $x' = \overset{\leftarrow}{F}_X(u)$ è determinazione simulata di X. In particolare se F_X è continua in $\mathbb{R}$ e strettamente crescente dove $F_X \neq \{0,1\}$, allora $\overset{\leftarrow}{F}_X = F_X^{-1}$ in $]0,1[$ e $X' = F_X^{-1}(U)$.
Ad esempio, se X è un n.a. discreto, con determinazioni $x_1 < x_2 < \dots$ e $p_i = Pr(X = x_i)$, allora $F_X(x) = \sum_{i:x_i<x} p_i$ ed abbiamo visto che $\overset{\leftarrow}{F}_X(p) = x_i \Leftrightarrow F_X(x_{i-1}) < p \leq F_X(x_i)$. Quindi dato $U \sim Unif(0,1)$, allora $X' = \overset{\leftarrow}{F}_X(U) = x_i \Leftrightarrow F_X(x'_{i-1}) < u < F_X(x'_i)$.
Quindi la simulazione di un n.a. si basa su un n.a. con distribuzione uniforme in $]0,1[$.
Se abbiamo invece bisogno di più simulazioni simulate, ci servono delle determinazioni che siano realizzazioni di variabili aleatorie IID, ovvero di una sequenza $U_1, \dots, U_m$ IID ed otterremo un campione di n.a. IID $X'_1 = \overset{\leftarrow}{F}(U_1), \dots, X'_m = \overset{\leftarrow}{F}(U_m)$, con X'_i che simula X.

In generale per costruire i n.a. simulati si sfrutta un generatore di numeri pseudo-casuali, ovvero un algoritmo deterministico, inizializzato con un *seme*, che produce sequenze di numeri $u_1, u_2, \dots$, con determinate proprietà statistiche, che consentano di interpretarli come determinazioni di una sequenza di n.a. $U_1, U_2, \dots$ IID.

5.1 21/11/2019

<u>**Simulazione di un vettore aleatorio o di un processo stocastico a parametro discreto**</u>

Sia $\underline{X} = (X_1, \dots, X_n) \sim F_{\underline{X}}$ oppure un processo stocastico $\{X_1, X_2, \dots\} \sim \mathcal{F}$ e siano $U_1, U_2, \dots$ numeri aleatori IID, con $U_i \sim Unif(0,1)$. Si hanno i seguenti due casi:
1 $X_1, X_2, \dots$ sono n.a. stocasticamente indipendenti.
Avendo assegnato la legge del processo, si hanno anche le distribuzioni marginali, di conseguenza possiamo ottenere la $\overset{\leftarrow}{F}_{X_i}$. Consideriamo i n.a. $\overset{\leftarrow}{F}_{X_1}(U_1)$, $\overset{\leftarrow}{F}_{X_1}(U_2), \dots$, questi risultano stocasticamente indipendenti, in quanto trasformati uno ad uno di n.a. stocasticamente indipendenti e $\overset{\leftarrow}{F}_{X_i}(U_i) \stackrel{d}{=} X_i$.
In caso di processo stocastico si simula solo il segmento iniziale, essendo il processo una successione infinita.
2 $X_1, X_2, \dots$ non sono n.a. stocasticamente indipendenti.
Siano $u_1, u_2, \dots$ le determinazioni di $U_1, U_2, \dots$, supposti osservabili, allora essendo assegnata $\mathcal{F}$, possiamo calcolare F_{X_1}, quindi $x'_1 = \overset{\leftarrow}{F}_{X_1}(u_1)$ e dato x'_1 troviamo $F_{X_2|X_1=x'_1}$ e $x'_2 = \overset{\leftarrow}{F}_{X_2|X_1=x'_1}(u_2)$, e così via. Si prova che $(x'_1, x'_2, \dots)$ è determinazione simulata di $\{X_1, X_2, \dots\}$, cioè un ente aleatorio con la medesima distribuzione dell'ente originale.

Simulazione di $W \sim N(0,1)$

F_W è strettamente crescente e continua, allora $\overleftarrow{F}_W(U) = F_W^{-1}(U) = W'$: Tuttavia non disponiamo della funzione inversa in forma chiusa, però si può applicare la trasformazione di *Box-Müller*. Si prova che se U_1, U_2 sono IID con $U_i \sim Unif(0,1)$, allora $W_1 = \sqrt{-2\ln U_i}\cos(2\pi U_2)$ e $W_2 = \sqrt{-2\ln U_i}\sin(2\pi U_2)$ sono IID e $W_i \sim N(0,1)$.

Simulazione di $X \sim N(\mu, \sigma^2)$

Consideriamo $W = \frac{X-\mu}{\sigma} \sim N(0,1)$, quindi $X = \sigma W + \mu$. Quindi generando w' come visto al paragrafo precedente, possiamo trasformarlo in una simulazione generata di X.

Simulazione di $Y \sim LN(\mu, \sigma)$

Consideriamo $Y = \log(X) \sim N(\mu, \sigma^2)$, quindi troviamo che $X = e^Y$. Quindi per simulare Y, è necessario simulare $X \sim N(\mu, \sigma^2)$ e poi trasformarlo in un $LN(\mu, \sigma)$.

Simulazione di $X \sim Gamma$

(A) Sia $X \sim \Gamma(1, \rho) \sim exp(\rho)$, allora $F_X(x) = 1 - e^{-\rho x}$, con $x > 0$ è strettamente monotona in $]0, +\infty[$, quindi possiamo scrivere in forma chiusa $\overleftarrow{F}_X(u) = F_X^{-1}(u) = -\frac{1}{\rho}\ln(1-u)$, con $u \in]0,1[$. Allora considerati $U \sim Unif(0,1)$ e $X' = \overleftarrow{F}_X(u) = -\frac{1}{\rho}\ln(1-u) \overset{d}{=} X$, quest'ultimo simula X. Inoltre essendo $U \sim Unif(0,1)$, si prova che $1 - U \overset{d}{=} U$, ma allora anche $X'' = -\frac{1}{\rho}\ln(U)$ simula X.
(B) Sia $X \sim \Gamma(n, \rho)$, con $n \in \mathbb{N}$, cioè una *distribuzione Erlangiana* con $f_X(x) = \frac{\rho^n}{(n-1)!}x^{n-1}e^{-\rho x}$, si ha la densità in forma chiusa, ma non la FdR.
Inciso: una proprietà della famiglia Γ dice che dati $X_1, \dots, X_n$ n.a. IID, con $X_i \sim \Gamma(\alpha, \rho)$, se consideriamo $\sum_i^n X_i$, questo si distribuisce come una $\Gamma(n\alpha, \rho)$
Quindi possiamo vedere $X = X_1 + \dots + X_n \sim \Gamma(n, \rho)$, come una somma di $X_i \sim \Gamma(1, \rho) \sim exp(\rho)$, simulando quindi n X_i' e sommandoli per ottenere una simulazione di X.
(C) Sia $X \sim \Gamma(\alpha, \rho)$, con $\alpha > 0 \in \mathbb{R}$. Per simularlo ci affidiamo al pacchetto R.

Simulazione di $N \sim Poi(\lambda)$

Sia $N \sim Poi(\lambda)$, allora $Pr(N = n) = \frac{e^{-\lambda}\lambda^n}{n!} = p_n$ ed è possibile calcolare $F_N(\cdot)$. Si può quindi simulare direttamente, considerando $U \sim Unif(0,1)$ e $N' = \overleftarrow{F}_N(U) \overset{d}{=} N$. Il problema della simulazione diretta, sta nell'algoritmo di ricerca, poiché dobbiamo trovare n^* tale che $F_N(n*-1) < u \leq F_N(N*)$ ed essendo $n \in \mathbb{N}$, potrebbe essere che il software ci impieghi molto tempo.

Quando consideriamo il processo di Poisson di intensità λ $\{N(t), t \geq 0\}$, possiamo considerare anche il processo dei tempi di inter-arrivo $\{W_1, W_2, \dots\}$ con n.a. $W_i \sim exp(\lambda)$ IID e il processo dei tempi di attesa $\{T_0, T_1, \dots\}$, con $T_m = W_1 + \dots + W_m$ e consideriamo i seguenti eventi:
∘ $N(t) = n \Leftrightarrow T_n \leq t \wedge T_{n+1} > t \Leftrightarrow T_n < t < T_{n+1} \Leftrightarrow W_1 + \dots + W_n \leq t < W_1 + \dots + W_n + W_{n+1}$
∘ $N(t) = 0 \Leftrightarrow T_1 > t \Leftrightarrow t < W_1$
Allora possiamo descrivere il n.a. $N(t) = \sum_{n=0}^{+\infty} n \cdot |N(t) = n| = \sum_n n|W_1 + \dots + W_n \leq t < W_1 + \dots + W_n + W_{n+1}|$ e se il processo $\{W_1', W_2', \dots\}$ simula $\{W_1, W_2, \dots\}$ con determinazioni $\{w_1', w_2', \dots\}$, allora si ha che $N'(t) = \sum_{n\geq 0} n|w_1' + \dots + w_n' \leq t < w_1' + \dots + w_{n+1}'|$ simula $N(t) \sim Poi(\lambda t)$.

Simulazione di $N \sim BN(\alpha, p)$

Sia $N \sim BN(\alpha, p)$, con $\alpha > 0$ e $p \in]0,1[$, consideriamo la relazione della binomiale negativa con la mistura di Poisson-Gamma, ovvero siano $N'|\Theta = \theta \sim Poi(\lambda\theta)$ e $\Theta \sim \Gamma(\alpha, \alpha)$, allora $N' \sim BN(\alpha, p)$, con $p = \frac{\alpha}{\alpha+\lambda}$. Di conseguenza consideriamo il vettore aleatorio (Θ, N'), le cui componenti non sono stocasticamente indipendenti.

Simulazione di $X \sim PoissonComposto(\lambda, F_Y)$

Sia $X = \sum_i^N Y_i$, con ipotesi che $\forall n \geq 0$ $Y_1|N = n, \dots, Y_n|N = n$ sono IID, che $Y_i|N = n \sim F_Y$, con $i \leq n$, ha legge che non dipende da n e che $N \sim Poi(\lambda)$. Si assegna la base tecnica e si simula il processo. Generiamo u_0 e consideriamo $\overleftarrow{F}_N(u_0) = n'$. Se $n' = 0$, allora $x' = 0$, se $n' > 0$, allora si genera u_1, da cui ricaviamo $y'_1 = \overleftarrow{F}_{Y|N=n'}(u_1)$, si genera u_2 e calcoliamo $y'_2 = \overleftarrow{F}_{Y_2|N=n, Y_1=y'_1}(u_2) = \overleftarrow{F}_Y(u_2)$, fino a trovare y'_n. Infine troviamo $x' = y'_1 + \cdots + y'_n$, determinazione simulata di X.
Il problema sorge se X si riferisce al risarcimento totale per il portafoglio, poiché si devono effettuare molte simulazioni, essendo N molto grande.

5.2 25/11/2019

Metodi approssimativi per la simulazione di $X \sim PoissonComposta$

1 **Approssimazione di Wilson-Hilferty**

Sia $X = \sum_i^N Y_i$ con $X \sim PoiComp(\lambda, F_Y)$, $\mathbb{E}(X) \triangleq \mu_X = \mathbb{E}(N)\mathbb{E}(Y)$, $\sigma_X^2 = \mathbb{E}(N)\mathbb{E}(Y^2)$ e $\gamma_X = \mathbb{E}[(x - \mu_X^3)]\sigma_x^{-3} = \frac{\lambda\mathbb{E}(Y^3)}{\sigma^3}$ il coefficiente di asimmetria. Ma F_X non può essere scritta in forma chiusa e quindi viene approssimata con la seguente formula, dove $\Phi(\cdot)$ è la FdR di una $N(0, 1)$:

$$F_X(x) \simeq \Phi\left[h\left(\frac{x - \mu_X}{\sigma_X}\right)\right]$$

$$h(t) = c_1 + c_2(t + c_3)^{\frac{1}{3}} \quad \text{dove} \quad c_1 = \frac{\gamma_X}{6} - \frac{6}{\gamma_X}\ , c_2 = 3\left(\frac{2}{\gamma_X}\right)^{\frac{2}{3}}\ , c_3 = \frac{2}{\gamma_X}$$

Il procedimento si basa sul rendere, tramite $h(\cdot)$, simmetrica una distribuzione che non lo è , simile all'approssimazione di Taylor, per poi calcolarla tramite la distribuzione $N(0, 1)$. Affinché ci sia una buona approssimazione, γ_X deve essere sufficientemente piccolo.
In forza dell'approssimazione si ha che $F_X(\cdot) \simeq F_{X'}(\cdot)$ e quindi $X' \overset{d}{\simeq} X$.

2 **Approssimazione Normal-Power**

Si dimostra che per $x > 0$ e γ_X piccolo, vale:

$$F_X(x) \simeq \Phi\left(-\frac{3}{\gamma_X} + \sqrt{1 + \frac{9}{\gamma_X^2} + \frac{6}{\gamma_X}\left(\frac{x - \mu_X}{\sigma_X}\right)}\right)$$

E si può provare che $X' = \mu_X + \sigma_X W + \frac{1}{6}\sigma_X\gamma_X(W^2 - 1)$, con $W \sim N(0, 1)$ e si ha xhe $X' \overset{d}{=} X$.
Questa approssimazione è comoda, perché per simulare possiamo basarci sui soli primi tre momenti della distribuzione composta.

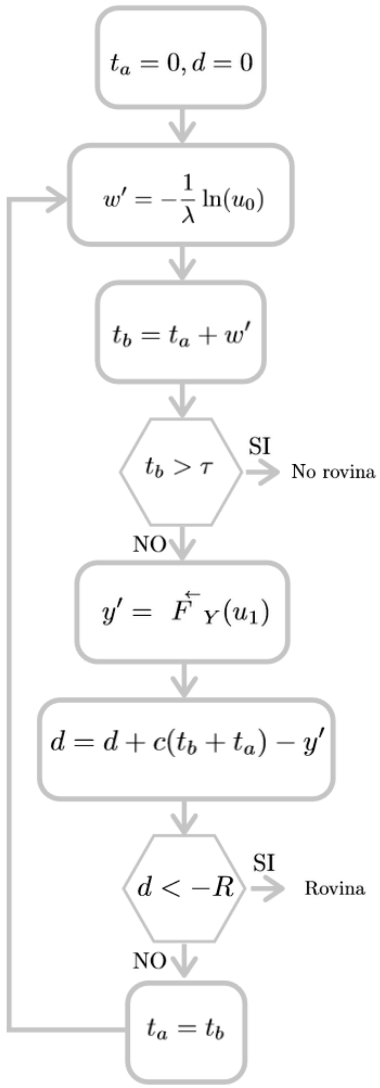
Problema: sia X il risarcimento totale per un portafoglio in un fissato arco di tempo monoannuale, per le normative attuali, $R = VaR[X, p] - (\mathbb{E}(X) + m)$, con $p = 0.995$. Assumiamo $X \sim PoiComp(\lambda, F_Y)$ e per calcolare il $VaR[X, p] = \overleftarrow{F}_X(p)$, abbiamo bisogno di F_X. Mediante l'approccio simulativo, generiamo $x'_1, \dots, x'_m$ determinazioni di X e costruiamo la FdR empirica, mediante la cardinalità. Poiché $F_m \simeq F_x$, allora $\widehat{VaR}[X, p] = \overleftarrow{F}_m(p)$.

Simulazione di un processo di Poisson

Sia $\{N(t), t \geq 0\}$ di intensità λ, può essere simulato nonostante sia a tempo continuo, grazie alla relazione con il processo degli intertempi $\{W_1, W_2, \dots\}$ IID e con $W_i \sim exp(\lambda)$. Per simulare una traiettoria in $[0, \tau]$ si inizia generando u_1, determinazione di $U_1 \sim Unif(0,1)$, e calcoliamo il primo intertempo $w_1' = -\frac{1}{\lambda}\ln(u_1) = t_1'$. Se $t_1' > \tau$, allora $N(t) = 0$, $\forall t \in [0, \tau]$, mentre se $t_1' \leq \tau$, allora $N(t) = 0$, $\forall t \in [0, t_1'[$ e generiamo u_2. Calcoliamo w_2' e troviamo $t_2' = w_1' + w_2'$. Se $t_2 > \tau$, allora $N(t) = 1$, $\forall t \in [t_1', \tau]$, oppure se $t_2' \leq \tau$, si abbiamo che $N(t) = 1$, $\forall t \in [t_1', t_2'[$. Il procedimento continua fino a trovare $t_{n*}' > \tau$.

Valutazione della probabilità di rovina per il modello classico della TCR

Consideriamo il processo della riserva $\{R(t), t \geq 0\}$, dove $R(t) = R + C_t + S(t)$, con $S(t) = \sum_{h=1}^{N(t)} Y_h$. L'evoluzione del processo della riserva dipende dall'evoluzione di processo dei risarcimenti cumulati $\{S(t), t \geq 0\}$, che a sua volta dipende dai processi $\{Y_1, Y_2, \dots\}$ e $\{N(t), t \geq 0\}$. Mediante le simulazioni possiamo valutare la probabilità di rovina su un orizzonte temporale limitato:$\Psi(R, \tau) = Pr(\bigvee_{0 \leq t \leq \tau} R(t) < 0)$. Di seguito è riportato il diagramma a blocchi della simulazione:



6 Riserve Tecniche

Le riserve tecniche rappresentano i debiti nei confronti degli assicurati e compaiono nelle passività del bilancio, al contrario delle riserva patrimoniali, che sono costituite dall'utile accantonato nel tempo e perciò sono libere da impegni, infatti compaiono nel patrimonio netto.
Le riserve principali nel ramo danni sono:
1 **Riserva Premi:** si forma alla chiusura di esercizio per far fronte a sinistri futuri, con le relative spese, per contratti non ancora scaduti, a fronte dei premi già incassati.
2 **Riserva Sinistri:** si forma alla chiusura di esercizio per far fronte alle passività già sorte, anche in esercizi precedenti, ma non ancora completamente risarcite, a causa del fatto che il processo di liquidazione risulta molto articolato.
3 **Riserva di Senescenza:** nasce dai contratti di tipo malattia con durata pluriennale e a premio livellato, si forma infatti dal fatto che i primi premi pagati sono maggiori del premio naturale.
4 **Riserva di Perequazione:** la normativa civilistica la prevede per alcuni rami, per coperture di elevata variabilità nel segno della gestione tecnica $G(t)$, mentre per altri rami è integrata nella riserva premi. In generale però non ha un significato di passività, ma di fondo per il rischio, motivo per cui con le nuove normative (*IFRS17*) verrà sostituita dai capitali a copertura per il rischio.

Decreto 173/97 (Normativa Civilistica)
1 Prevede che la riserva premi sia formata da due componenti: la *riserva per frazioni di premio* e la *riserva per rischi in corso.*
∘ **Riserva per Frazioni di Premio:** nasce al momento della quietanza di pagamento, indipendentemente che il premio sia stato incassato, ed è proporzionale al tempo di durata del contratto alla chiusura di esercizio. Se un contratto di copertura monoannuale viene stipulato in t_h, con scadenza quindi in $1+t_h$, allora la riserva per frazioni di premio sarà $P_h^T(1-\alpha')t_h$, dove P_h^T sono i premi lordi dedotti dalle imposte e dalle tasse e α' sono l'aliquota delle spese di acquisizione.
Giustificazione metodo Pro-Rata Temporis: sia $S(t) = \sum_{h=1}^{N(t)} Y_h$ il risarcimento per una fissata polizza in $[0,t]$, consideriamo il processo $\{S(t), t \geq 0\}$ e nelle ipotesi del modello classico della TCR si ha che $\{S(t), t \geq 0\} \sim PoiComp(\lambda, F_Y)$.
Considerando il risarcimento su un intervallo di ampiezza Δt, $S(t+\delta) - S(t)$ è a incrementi stazionari ed indipendenti ed è uguale a $S(t+\Delta t) - S(t) = \lambda \Delta t \mathbb{E}(Y) = \mathbb{E}(S(t+1) - S(t))\Delta t$. E quindi risulta proporzionale in base alla durata e al risarcimento annuo atteso, nel caso del modello della TCR, ed inoltre in termini di premio equo si ha $\Delta t P_{annuo}$ che corrisponde a $P_h^T(1-\alpha')\Delta t$.
Di conseguenza l'ipotesi sottostante al metodo della *Pro-Rata Temporis* è quella della TCR. I fenomeni stagionali vengono annullati per le compensazioni tra polizze, nel caso di portafogli bilanciati.

6.1 28/11/2019

Un altro metodo per la stima della riserva per frazioni di premio è il *metodo forfettario*, che risulta un'approssimazione del metodo *Pro-Rata*. Si applica un'aliquota prefissata ai premi lordi contabilizzati dell'esercizio. Ad esempio per RCA è 0.4, per assicurazione standard 0.35 e 0.15 per assicurazioni brevi. Questo metodo viene utilizzato poiché se assumiamo che tutti i contratti vengano stipulati a metà anno, con il metodo *Pro-Rata* abbiamo:

$$R_P(1) = \sum_{i=1}^{n} P_h^T(1-\alpha')t_h = \sum_{i=1}^{n} \frac{1}{2} P_h^T(1-\alpha') = \frac{1-\alpha'}{2} \sum_{i=1}^{n} P_h^T = \frac{1-\alpha'}{2} P^T(1)$$

Se $\alpha' = 0.3$ otteniamo il metodo forfettario. Perciò assumendo che i contratti vengano stipulati in maniera uniforme, il metodo forfettario approssima quello *Pro-Rata Temporis*.
Prima del 1997 la R_P veniva calcolata con il metodo forfettario, con aliquote diverse, poiché non era semplice controllare polizza per polizza la durata residua. Attualmente si applica se $R_{ProRata} < R_{Forfait} \leq 1.02 R_{ProRata}$, ovvero purché sia prudenziale, ma non eccessivamente.
◦ **Riserva per Rischi in Corso:** nasce nel momento in cui vi è insufficienza di accantonamento di R_p, cioè i sinistri e le spese future superano l'importo, è quindi un'integrazione.
Questa riserva deve essere accantonata, poiché l'ipotesi sottostante la *Pro-Rata Temporis* è la stazionarietà, ma possono esserci dei fenomeni stagionali e se il portafoglio non è sufficientemente bilanciato, si ha carenza. Perciò risulta un'integrazione dovuta alla carenza nella valutazione della riserva per frazioni di premio. La normativa non fornisce indicazioni sui metodi di calcolo, mentre nella circolare dell'ISVAP sono proposti due approcci:
- Criterio Analitico: viene valutata la riserva iniziale $P_h^T(1-\alpha')t_h$ e la riserva con il nuovo stato di informazione $P_h^{T\prime}(1-\alpha'')t_h$, se $P_h^T(1-\alpha')t_h < P_h^{T\prime}(1-\alpha'')t_h$, allora c'è carenza e la differenza costituisce la riserva per rischi in corso.
- Criterio Semplificato: si considera come indicatore di sinistrosità il *Combined Ratio*, ovvero il rapporto tra la competenza sinistri e la competenza premi dell'esercizio, al netto delle provvigioni, e può essere interpretato come costo per sinistro per unità di premio.
Consideriamo il volume dei premi VP, composto dalla riserva per frazioni di premio e dalle rate di premio future, per contratti dell'esercizio, al netto delle provvigioni. Quindi $CR \cdot VP$ è una stima dei costi per risarcimento e spese future per contratti dell'esercizio. Se $CR > 1$, allora $CR \cdot VP > VP$ e quindi $VP(1 - CR)$ è la riserva per rischi in corso.
2 Sia per un sinistro t_1 il momento di accadimento, t_2 il momento di denuncia, t_3 il momento di liquidazione e t_4 il momento di chiusura. Solitamente la distanza tra t_1 e t_2 è breve, altrimenti si parla di sinistri IBNR, e dal momento t_2 inizia il processo di liquidazione. In t_3 viene determinato l'importo e questo verrà risarcito tra t_3 e t_4, momento nel quale il sinistro viene chiuso. La passività nasce dal momento in cui avviene il sinistro t_1 e rimane in sospeso fino a t_4, questo periodo è chiamato *Outstanding Loss Liability* e a fronte di ciò viene accantonato R_S.
Le tipologie di spese vengono suddivise in 5 categorie:
◦ *Allocated Loss Adjusted Expenses:* sono spese direttamente imputabili (es. avvocati, liquidazioni,...)
◦ *Unallocated Loss Adjusted Expenses:* sono spese non direttamente imputabili (es. spese amministrative,...)
Nel Comma 2 si fa riferimento al costo ultimo, cioè al costo del futuro risarcimento, ma la valutazione avviene in 1. Nel Comma 3 è specificato il metodo di calcolo, ovvero il metodo dell'inventario: i liquidatori fanno una valutazione di quanto verrà risarcito e del costo ultimo sinistro per sinistro.
Inoltre la riserva sinistri viene utilizzata per valutare la redditività del portafoglio, per valutare la possibilità di ridurre i premi e per il controllo interno. Perciò la normativa permette di applicare anche al metodo del costo medio, che è una tecnica statistico-attuariale, per quanto concerne i sinistri di maggiore incertezza e quelli con caratteristiche simili, invece che valutare sinistro per sinistro.
Attualmente la valutazione della R_S è multifase e coinvolge gli uffici liquidativi e gli uffici direzionali dell'impresa, che può applicare metodi statistico-attuariali, ed è perciò complessa.
Inoltre si deve tener conto dei fattori inflattivi, in particolare del *claim inflaction*, che si suddivide in: inflazione monetaria di settore, determinata dalla variazione dei costi specifici del settore, e la *super imposed tarifation*, per cui l'effetto sui costi non è imputabile ad aspetti monetari, ma a disposizioni legislative o decisioni giurisprudenziali.

La valutazione deve essere accurata, avendo conseguenza sulle gestioni future, tuttavia vi sono limiti nell'accantonare dal punto di vista del fisco, dell'immagine che si dà agli investitori e per le analisi di controllo interno.
Non tutti i sinistri vengono riservati, infatti l'azienda cerca di risarcire rapidamente, si cerca di riservare solo per i sinistri per cui si hanno valutazioni complesse.

Processo di smontamento della Riserva

Fissata l'attenzione su un esercizio θ, una parte dei sinistri avvenuti in θ viene risarcita, una parte no, per questo si forma $R_{S,\theta}^{(\theta)}$. In $\theta+1$ ci sono elementi che fanno aumentare o ridurre la riserva, come i pagamenti, le rivalutazioni o riaperture di sinistri e otteniamo $R_{S,\theta}^{(\theta+1)}$, fino ad ottenere $R_{S,\theta}^{(\theta+t)}$, dove t è la durata del processo di smontamento. Solitamente t è stabile all'interno del ramo e rappresenta la durata entro la quale la maggior parte dei sinistri dell'esercizio di origine sono chiusi.
In RCA il processo di smontamento della riserva, viene in gran parte chiusa nell'anno di origine con differimento di 1 anno, mentre in termini di costo medio per sinistro, un sinistro chiuso con differimento 0 è molto più basso di uno chiuso negli anni successivi.

Metodi Statistico-Attuariali per la valutazione della Riserva Sinistri

Si distinguono i metodi deterministici da quelli stocastici. Sono moltissime le tecniche deterministiche, che in sono degli algoritmi che producono una stima puntuale *best-estimate*, tecnicamente su basi empiriche e successivamente giustificati con metodologie. Per quanto riguarda i metodi stocastici, essendo le quantità da valutare dei n.a., vengono introdotte delle ipotesi probabilistiche sulla base dei dati e si stimano i parametri. Si possono avere anche indicatori sulla qualità delle stime e talvolta la distribuzione di probabilità.
Fino a pochi anni fa si utilizzavano solo metodi deterministici, ma con *Solvency II* e *IFRS17* nasce l'esigenza dell'approccio stocastico. É comunque permesso applicare qualche algoritmo deterministico, ma riportando comunque i risultati in un quadro stocastico.

Schema di riferimento generale per dati aggregati

Consideriamo una tipologia di sinistri con caratteristiche di analogia, vogliamo valutare la R_S alla chiusura dell'esercizio il $31/12/\theta$, in particolare i risarcimenti e spese future per sinistri avvenuti in $\theta, \theta+1, \dots$.
Si indica con i il periodo di origine del sinistro, dipende dall'unità di tempo, e con origine si può intendere la stipulazione della polizza, l'accadimento del sinistro o la denuncia. Consideriamo tutti i sinistri con medesimo periodo di origine i e i pagamenti possono avvenire in $i, i+1, \dots, i+t$, con t durata del processo di smontamento e sono detti periodi di pagamento o di calendario. I sinistri pagati con un differimento maggiore di t sono considerati code e per questo si fa una valutazione separata e si corregge R_S.
Indichiamo che $T = 0, \dots, t$ il differimento o sviluppo di un pagamento, rispetto al periodo di origine, e ipotizziamo che:
◦ l'unità di misura di i e j sia la medesima e che sia l'anno
◦ alla chiusura i sinistri più vecchi saranno quelli di $\theta - t + 1$ e si codificano gli anni di origine con $i = j = 0, \dots, t$ e in questo modo t ha un significato sia di durata del processo di smontamento, che di anno della valutazione.
Indichiamo con X_{ij} una grandezza connessa con la valutazione di R_S, relativa all'anno di origine i, chiusi con differimento j.
Gli X_{ij} possono essere i pagamenti incrementali P_{ij}, ovvero l'importo per sinistri con anno di accadimento i, pagati con differimento j, e quindi pagati in $i+j$. Oppure possono essere i pagamenti cumulati C_{ij},

ovvero l'importo totale pagato per sinistri con anno di origine i e differimento minore o uguale a j, cioè $C_{ij} = \sum_{h=0}^{j} P_{ij}$. O ancora il rapporto tra P_{ij} e ω_{ij} o tra C_{ij} e ω_{ij}, dove gli ω_{ij} sono una grandezza nota, che solitamente dipendono solo dall'anno di origine (ω_i) e sono una misura di esposizione, come ad esempio i premi di competenza dell'anno i.

6.2 29/11/2019

Possiamo anche considerare i P^*_{ij} e i C^*_{ij} che sono i pagamenti incrementali e cumulati espressi a valore nominale dell'epoca di pagamento $i + j$, perciò i P_{ij} non sono omogenei in termini monetari. Per questo siamo interessati ad esprimerli a valore dell'anno t, ma per fare questo abbiamo bisogno dei tassi di inflazione $\phi_1, \phi_2, \dots, \phi_t$, così che $P^*_{ij} = P_{ij}(1 + \phi_{i+j+1}) \dots (1 + \phi_t)$.
Inoltre possono essere considerati gli *Incurred Loss* I_{ij}, che sono la stima alla fine del periodo $i + j$ del costo complessivo dei sinistri origine in i. Siccome la valutazione del costo ultimo può essere fatta via via nel tempo, si può rilevarne l'evoluzione, per quanto concerne i sinistri con origine i all'aumentare dell'informazione in $i + j$. Inoltre si ha che I_{ij} è formato da C_{ij} e dalla Riserva di inventario, cioè la componente già pagata, più le somme ancora da pagare.
Infine le altre grandezze che possono essere usate sono il numero di sinistri che sono avvenuti in i e che sono stati chiusi con differimento j N_{ij} e il loro numero cumulato N^C_{ij}, oppure il numero di sinistri che sono stati denunciati in i e che sono stati chiusi con differimento j D_{ij} e il loro numero cumulato D^C_{ij}.

Queste grandezze vengono rappresentate nella tabella di *Run-Off*, in cui l'ipotesi fondamentale è che la riserva venga completamente smontata in t. La tabella si suddivide in due triangoli, il triangolo superiore, che contiene i dati con le grandezze date, e il triangolo inferiore, che contiene le grandezze incognite, perché riferite ad anni di calendario futuri. In basso nella Tabella 1 vediamo come è strutturata la classica tabella di X_{ij}, con sulle righe gli anni di origine, mentre sulle colonne gli anni di differimento. Sulla diagonale troviamo gli anni di calendario ($i + j = t$). In tabella sono raffigurati con $*$ i valori disponibili al tempo t, mentre nel triangolo inferiore i valori che vogliamo stimare.

Nei modelli stocastici vengono introdotti il processo stocastico $\{X_{ij}; i = 0, \dots, t; j = 0, \dots, t\}$, delle ipotesi probabilistiche e si usano i dati come campione osservato degli X_{ij} con $i + j \le t$, per stimare i parametri del modello probabilistico, per poi ottenere le stime di X_{ij}, con $i + j > t$.
Nei modelli deterministici invece gli X_{ij} con $i + j \le t$ sono visti come noti ed attraverso opportuni algoritmi, si stimano gli X_{ij}, con $i + j > t$.

$i \vert j$	0	1		j		t
0	$*$	$*$	$*$	$*$	$*$	$*$
1	$*$	$*$	$*$	$*$	$*$	$\circ$
	$*$	$*$	$*$	$*$	$\circ$	$\circ$
i	$*$	$*$	$*$	$\circ$	$\circ$	$\circ$
	$*$	$*$	$\circ$	$\circ$	$\circ$	$\circ$
t	$*$	$\circ$	$\circ$	$\circ$	$\circ$	$\circ$

Table 1: Esempio X_{ij}

Metodo della Catena (*Chain Ladder*)

É il metodo più utilizzato, si prendono come grandezze di riferimento i C_{ij} e si formulano le seguenti ipotesi:
∘ $C_{ij} > 0$
∘ $C_{ij} \ge C_{i,j-1}$, ovvero $P_{ij} = C_{ij} - C_{i,j-1} \ge 0$
Se $P_{ij} < 0$, allora si hanno degli introiti, come recuperi dai riassicuratori, o rivalse nei confronti

dell'assicurato, ma questi vengono valutati, per le normative, separatamente, quindi rimaniamo sull'ipotesi fatta.
I dati disponibili sono i C_{ij}, con $i + j \leq t$, ovvero il triangolo superiore della Tabella 1, di cui $C_{i,t-i}$ corrisponde all'importo pagato ad oggi per sinistri con origine in i. Invece C_{it} è l'importo pagato per sinistri con origine i, entro t anni, ovvero il costo complessivo dei sinistri con origine in i ed è detto *Costo Ultimo*.
$C_{it} - C_{i,t-i} \triangleq OLL_i$ è l'*Outstanding Loss Liability* per l'anno i, cioè quanto si deve ancora pagare. Preso $\widehat{C}_{it}$ come stima del costo ultimo, si pone $R_i = \widehat{C}_{it} - C_{i,t-i}$ come stima *Best Estimate* della riserva sinistri.
Il nostro obbiettivo è dunque stimare le quantità $\{C_{it}\}_{1 \leq i \leq t}$, per ottenere le stime di R_i.
Ipotesi base del modello:
∘ $\forall i, j$ tale che $i + j \leq t$ il rapporto $\frac{C_{ij}}{C_{i,j-1}}$ non dipende da i, ovvero se $\exists f_0, \ldots, f_{t-1} > 0$, tali che $\frac{C_{ij}}{C_{i,j-1}} = f_{j-1}$.
La condizione $\frac{C_{ij}}{C_{i,j-1}} = f_{j-1}$ può essere riscritta equivalentemente come:
- $C_{ij} = f_{j-1} C_{i,j-1}$
- $\exists f_0, \ldots, f_{t-1}$ tale che $\frac{C_{ij}}{C_{i,j-1}} = f_{j-1}$, se e solo se $\exists \alpha_0, \ldots, \alpha_t > 0$ e $\exists r_0, \ldots, r_t \geq 0$, con $r_0 > 0$, tale che $P_{ij} = \alpha_i r_j$.

Dimostrazione
$\Rightarrow$ Se $\exists \alpha_0, \ldots, \alpha_t > 0, r_0, \ldots, r_t \geq 0$ tali che $P_{ij} = \alpha_i r_j$, allora $\exists f_0, \ldots f_{t-1}$ tali che $C_{ij} = f_{j-1} C_{i,j-1}$. Infatti $C_{ij} = \sum_j P_{ij} = \sum_j \alpha_i r_j = \alpha_i \sum_j r_j$, quindi troviamo che $f_{j-1} \triangleq \frac{\alpha_i \sum_h^j r_h}{\alpha_i \sum_h^{j-1} r_h} = \frac{\sum_h^j r_h}{\sum_h^{j-1} r_h}$, che non dipende da i.
$\Leftarrow$ Se $\exists f_0, \ldots, f_{t-1} > 0$, tali che $C_{ij} = f_{j-1} C_{i,j-1}$, allora $\exists \alpha_0, \ldots, \alpha_t > 0, r_0, \ldots, r_t \geq 0$ tali che $P_{ij} = \alpha_i r_j$. Infatti, dato che per $j = 0$ $P_{i0} = C_{i0}$ e che per $j \geq 1$ $P_{ij} = C_{ij} - C_{i,j-1}$, si ha che $C_{i,j+1} = C_{ij} f_j$, $C_{i,j+2} = C_{i,j+1} f_{j+1}, \ldots$, quindi $C_{it} = C_{ij} \prod_{h=j}^{t-1} f_h$ e di conseguenza $C_{ij} = C_{it} (\prod_{h=j}^{t-1} f_h)^{-1}$, Quindi $P_{ij} = C_{ij} - C_{i,j-1} = C_{it} [(\prod_{h=j}^{t-1} f_h)^{-1} - (\prod_{h=j-1}^{t-1} f_h)^{-1}] = \alpha_i r_j$, dato che C_{it} dipende solo da i e la quadra dipende solo da j.

Quindi l'ipotesi base del modello implica che i pagamenti incrementali si possano rappresentare tramite un modello moltiplicativo $P_{ij} = \alpha_i r_j$, in cui α_i dipende solo dal periodo di origine e r_j solo dal periodo di differimento. poiché per i modelli moltiplicativi i fattori sono determinati a meno di un fattore moltiplicativo non nullo, non è restrittivo assumere che $\sum_h r_h = 1$. Infatti se $\sum_h r_h \neq 1$, allora potremmo considerare $P_{ij} = \alpha_i r' \frac{r_j}{r'} = \alpha_i' r_j'$, così che $\sum_h r_h' = 1$. In questo modo è possibile interpretare r_j come delle aliquote ed in particolare è la quota del costo ultimo pagato con differimento j, infatti $\sum_j P_{ij} = C_{it} = \alpha_i \sum_j r_j = \alpha_i$, che quindi è interpretabile come il costo ultimo per l'anno di origine i.
Da questo emerge un'altra ipotesi molto forte: r_j non dipende dall'anno di origine i e questo implica la stabilità nel processo di smontamento della riserva, poiché la sequenza degli $(r_j)_{j \geq 0}$ resta costante nel tempo per i diversi periodi di sviluppo del costo ultimo, variando solo gli α_i.
Osservazione: $P_{ij} = \alpha_i r_j \Leftrightarrow \frac{C_{ij}}{C_{i,j-1}}$, ma la prima condizione non è soddisfacente, lo sarebbe in termine attesi, ovvero $\mathbb{E}(P_{ij}) = \alpha_i r_j$, ma ciò comporterebbe un modello stocastico.

Quindi a partire da $f_{j-1} = \frac{C_{ij}}{C_{i,j-1}}$, si stimale k componenti: $C_{i,t-i+1} = C_{i,t-i} f_{t-i}$, $C_{i,t-i+2} = C_{i,t-i} f_{t-i} f_{t-i+1}$. Quindi stimiamo, per $i + j > t$, $C_{ij} = C_{i,t-i} \prod_{h=t-i}^{j-1} f_h$. Troviamo infine che $OLL_i = C_{it} - C_{i,t-i} = C_{i,t-i} (\prod_{h=t-i}^{t-1} f_h - 1)$.
Di conseguenza ci basta stimare i fattori di sviluppo, per ottenere il costo ultimo e la riserva.

Stima dei fattori di sviluppo

Sia $C_{ij} = C_{i,j-1}f_{j-1}$, $\forall i,j$, allora $\sum_i C_{ij} = \sum_i C_{i,j-1}f_{j-1} \Rightarrow f_{j-1} = \frac{\sum_i C_{ij}}{\sum_i C_{i,j-1}}$. In particolare si considera i tale che è il massimo indice per cui abbiamo i dati, così da ottenere una stima il più corretto possibile, quindi prendiamo i tale che $i+j=t$ ed otteniamo $\widehat{f}_{j-1} = \frac{\sum_{i=0}^{t-j} C_{ij}}{\sum_{i=0}^{t-j} C_{i,j-1}}$. Da questi valori otteniamo tutti i $\widehat{C}_{ij}$ con $i+j>t$ e $\widehat{OLL}_i = R_i = \widehat{C}_{it} - C_{i,t-i}$, che dipende dall'ultimo pagamento cumulato e dal *link-ratio*.

Osservazioni sul metodo della catena

1 L'algoritmo si basa sull'ipotesi $\frac{C_{ij}}{C_{i,j-1}} = f_{j-1} \Leftrightarrow P_{ij} = \alpha_i r_j$, quindi i rapporti fra i pagamenti cumulati consecutivi non vengono a dipendere dall'anno di origine, ma solo dal differimento. E questo equivale a dire che i pagamenti incrementali seguono un modello moltiplicativo con un fattore che dipende solo dall'anno di origine e uno che dipende solo dal differimento, ovvero $\frac{P_{ij}}{P_{i,j-1}} = \frac{\alpha_i r_j}{\alpha_i r_{j-1}} \triangleq \overline{f}_{j-1}$. Perciò osserviamo che l'ipotesi vale anche per i pagamenti incrementali ed è quindi possibile applicare l'algoritmo anche su di essi, anche se nella pratica si applica sui cumulati, così da compensare tramite le k somme.

2 $C_{it} = C_{ij}\prod_{h=j}^{t-1} f_h \Leftrightarrow C_{ij} = C_{it}(\prod_{h=j}^{t-1} f_h)^{-1} = C_{it}b_j$, dove i b_j sono fattori che dipendono solo dal differimento. Quindi i pagamenti cumulati con differimento minore o uguale a j, possono essere espressi come il prodotto tra il costo ultimo per la quota del costo ultimo pagata con differimento minore o uguale a j. Si verifica che $b_j \geq b_{j-1} \Leftrightarrow \frac{b_j}{b_{j-1}} \geq 1 \Leftrightarrow (\prod_{h=j}^{t-1} f_h)^{-1}(\prod_{h=j-1}^{t-1} f_h) \geq 1 \Leftrightarrow f_{j-1} \geq 1$, quindi sempre essendo C_{ij} sempre maggiore o uguale di $C_{i,j-1}$. La sequenza $b_0 \leq \cdots \leq b_t = 1$ è detta *Claims Development Pattern*, con i singoli fattori detti *grossing-up* e rappresentano l'evoluzione delle aliquote dei pagamenti cumulati nel tempo.

3 Sono detti metodi di *link-ratio* i metodi che hanno alla base il metodo della catena sui costi cumulati, che però variano nel procedimento di stima degli f_{j-1}. Sono ancora una media ponderata dei *link ratio* osservati il più a lungo possibile, ma con pesi pari a $\omega_i^{(j)} \geq 0$ e $\sum_i \omega_i^{(j)} = 1$, ovvero $\widehat{f}_{j-1} = \sum_i \frac{C_{ij}}{C_{i,j-1}}\omega_i^{(j)}$.

4 Siano $\widehat{f}_0, \dots, \widehat{f}_{t-1}$ i *link-ratio* stimati. Può succedere che sulla coda le stime siano poco affidabili, poiché le ultime sono ottenute su pochi dati, perciò nella pratica si utilizzano delle tecniche di *smoothing*, ossia delle perequazioni per gli ultimi *link-ratio*. Fissato $\widehat{f}_{j-1}$, sia $h(j) = a \cdot b^{j-k}$, dove a e b sono stimati dai minimi quadrati e otteniamo $\widehat{f}_k^s = \widehat{a}\widehat{b}^{j-k}$, con $j \geq k+1$ e ottenendo così $\widehat{f}_0^s, \dots, \widehat{f}_{t-1}^s$.

5 É possibile ritrovare l'ipotesi base del modello anche per altre grandezze, quindi si può applicare il metodo anche a P_{ij}, N_{ij}, N_{ij}^c e spesso per I_{ij} si applica sul rapporto $\frac{I_{ij}}{I_{i,j-1}}$, dove l'ultima colonna della tabella è una stima del costo ultimo $\widehat{I}_{it}$, infatti risulta $R_i = \widehat{I}_{it} - C_{i,t-1}$.
Dunque si può considerare sia la Tabella dei C_{ij}, che dei I_{ij}, però spesso nella pratica accade che $\widehat{I}_{it} \neq \widehat{C}_{it}$ e quindi spesso, essendo importanti entrambe le informazioni, accade che:
∘ si applichino dei metodi stocastici, basati sulla logica della catena, al fine di ottenere stime della riserva prossime tra le due tabelle
∘ si applichino metodi che sfruttino entrambe le informazioni, per ottenere un'unica stima di R_i

6 L'ipotesi $P_{ij} = \alpha_i r_j$ risulta un'ipotesi molto forte per quanto concerne la stabilità espressa dalle aliquote r_j. Infatti possono esserci casi in cui si hanno delle modifiche al processo di smontamento della riserva o cambi nel portafoglio oppure ancora diverse cadenze nel pagamento per sinistri di grande entità, che rendono inaccettabile l'ipotesi. Un altro fattore sono gli effetti inflattivi, di cui si tiene conto nel metodo della catena modificato, ovvero a partire dalla matrice dei C_{ij}, attraverso le differenze prime,

si ricava la matrice dei P_{ij}, attraverso dei tassi inflattivi, si ottengono i P^*_{ij} a valori correnti dell'anno t, quindi si ricrea la matrice dei cumulati C^*_{ij} e, poiché i rapporti non dipendono da i, si può applica il metodo della catena per trovare i $\widehat{C}^*_{ij}$ con $i+j>t$. Però per le normative dicono che i valori devono essere espressi a valore nominale, quindi si trovano i $\widehat{P}^*_{ij}$ con $i+j>t$ e si moltiplicano per i coefficienti di inflazione, ovvero $\widehat{P}_{ij} = \widehat{P}^*_{ij}(1+\phi_{t+1})\dots(1+\phi_{i+j})$. Infine troviamo $\widehat{R}_i = \sum_{j=t-i+1}^{t} \widehat{P}_{ij}$ e $R = \sum_{i=1}^{t} R_i$.

Metodo di separazione di Taylor

Le grandezze di interesse sono i P_{ij} e nel modello deterministico si assume che $P_{ij} = \alpha_i r_j \lambda_{i+j}$, cioè un modello moltiplicativo con α_i un fattore che dipende dall'anno di origine, r_j che dipende solo dal differimento e λ_{i+j} che dipende dall'anno di calendario. Se $\lambda_{i+j} = \lambda \ \forall i,j$, allora si hanno le ipotesi del modello precedente.
L'idea di fondo è che la stabilità rappresentata da r_j venga in parte turbata dall'anno di calendario, ad esempio da effetti inflattivi, che vogliamo appunto catturare con λ_{i+j}.
Si parte quindi dalla tabella dei P_{ij} ed abbiamo una grandezza legata all'anno di origine, che fornisce una misura di esposizione w_i, come ad esempio il numero di sinistri denunciati nell'anno i n_i o i premi di copertura PC_i. Si pone $\alpha_i = w_i$, che di conseguenza non è un parametro, come base useremo gli n_i, e il modello diventa $P_{ij} = n_i r_j \lambda_{i+j}$, con r_j e λ_{i+j} parametri da stimare e $\sum_j r_j = 1$.
Si suppone che in termini reali il costo medio per sinistro μ risulti costante al variare dell'anno di origine. Sia λ_{i+j} il costo medio per sinistro a valori correnti dell'anno $i+j$, quindi $\lambda_{i+j} = \mu\rho_{i+j}$, dove ρ_{i+j} è il coefficiente di adeguamento dei costi, in termini inflattivi può essere visto come $\lambda_{i+j} = \mu(1+\phi_1)\dots(1+\phi_{i+j})$. Quindi troviamo che $P_{ij} = n_i\mu\rho_{i+j}r_j$.

Stima dei parametri

Dai dati otteniamo $r_0,\dots,r_t$ e $\lambda_0,\dots,\lambda_t$ e vogliamo le stime di $\lambda_{t+1},\dots,\lambda_{2t}$. Calcoliamo la matrice degli $S_{ij} = \frac{P_{ij}}{n_i}$, con $i+j\le t$ sono dati, e questi corrispondono alla matrice dei parametri del modello $r_j\lambda_{j+i}$. Dalla tabella dei dati emerge una certa regolarità, ovvero sulla diagonale esterna si ha sempre λ_t costante, mentre sulle colonne sono sempre costanti gli r_j. Di conseguenza grazie a queste regolarità e all'ipotesi del modello moltiplicativo con $\sum_j r_j = 1$, allora troviamo che $\sum_{i+j=t} S_{ij} = \sum_{j=0}^{t} \lambda_t r_j = \lambda_t \sum_{j=0}^{t} r_j = \widehat{\lambda}_t$ e di conseguenza $\widehat{r}_t = \frac{S_{0t}}{\lambda_t}$. A questo punto stimiamo λ_{t-1} e r_{t-1} attraverso l'equazione $\sum_{i+j=t-1} S_{ij} = \sum_{j=0}^{t-1} r_j\lambda_{t-1} = \lambda_{t-1}(1-\widehat{r}_t)$, e quindi troviamo $\widehat{\lambda}_{t-1}$, come rapporto dei due fattori. Invece $\widehat{r}_{t-1}$ lo troviamo da $S_{0,t-1} + S_{0t} = r_{t-1}\lambda_{t-1} + r_{t-1}\lambda_t$. Continuiamo questo algoritmo, fino a stimare $\widehat{r}_0,\dots,\widehat{r}_t$ e $\widehat{\lambda}_0,\dots,\widehat{\lambda}_t$. I generici $\widehat{r}_h$ e $\widehat{\lambda}_h$ saranno:

$$\widehat{\lambda}_h = \frac{\sum_{i+j=h} S_{ij}}{1-\sum_{k=1}^{t-h}\widehat{r}_{h+k}} \quad \text{e} \quad \widehat{r}_h = \frac{\sum_{i=0}^{t-h} S_{ih}}{\sum_{k=0}^{t-h}\widehat{\lambda}_{h+k}}$$

6.3 05/12/2019

Per stimare il triangolo inferiore dei P_{ij} è necessario stimare i $\lambda_{t+1},\dots,\lambda_{2t}$ e per farlo ci riferiamo all'interpretazione che i $\lambda_k = \mu\rho_k$, dove ρ_k è il coefficiente di adeguamento dei costi all'anno k. Se ρ_k sono i tassi di inflazione, possiamo vederlo come $\lambda_k = \mu(1+\phi_1)\dots(1+\phi_k)$ e si osserva che $\frac{\lambda_k}{\lambda_{k-1}} = \frac{\rho_k}{\rho_{k-1}} = 1+\phi_k$. I tassi futuri di inflazione si ottengono o attraverso stime esterne e quindi stimiamo direttamente i $\widehat{\lambda}_{i+j}$, oppure si fanno delle perequazioni: si sceglie una funzione che rispecchi l'andamento dei dati e si applicano i minimi quadrati $\min_{a,b}(f(k) - \frac{\widehat{\lambda}_k}{\widehat{\lambda}_{k-1}})^2$, da cui ricaviamo $\widehat{a}$ e $\widehat{b}$. Possibili $f(k)$ possono essere la funzione lineare $f(k) = b + ak$ o quella esponenziale $f(k) = ab^k$. In questo modo troviamo $\frac{\widehat{\lambda}_{t+1}}{\widehat{\lambda}_t} = \widehat{f}(t+1)$, da cui ricaviamo $\widehat{\lambda}_{t+1}$ e poi i successivi proseguendo.

Adesso è quindi possibile costruire il triangolo inferiore e calcolare di conseguenza $\widehat{OLL_i}$.
Osservazione: il metodo presente un grande vantaggio rispetto a quello della catena, poiché permette di ottenere stime dell'inflazione passata dai dati. Di fatti depurati i dati dall'inflazione monetaria si può estrarre l'inflazione *super-imposed* passata. I problemi si hanno nelle estrapolazione dell'inflazione futura.

Metodo del Loss Ratio Semplice

Le grandezze di interesse sono i C_{ij} e si ipotizza di conoscere i premi di competenza PC_i dell'anno di origine i. Poniamo il *Loss Ration Ultimate* come $\frac{C_{it}}{PC_i} \triangleq L_i$, ovvero il rapporto tra il costo complessivo dei sinistri con anno di origine i, cioè il costo ultimo, e i premi che competono a coprire quei sinistri. Allora posso interpretare il costo ultimo come $C_{it} = PC_i L_i$, tuttavia gli L_i andrebbero stimati, ma assumiamo di averli come dati esogeni l_i dati dalla compagni o stimati sul mercato.
Poniamo quindi $\widehat{L}_i = l_i$ e quindi $\widehat{C}_{it} = l_i PC_i$ e quindi la riserva $R_i = \widehat{C}_{it} - C_{i,t-1} = l_i PC_i - C_{i,t-i}$.
Osservazione: il metodo risulta molto semplice, ma non sfrutta l'evoluzione dei sinistri, basandosi solamente sui valori finali. Ciò può essere visto come un vantaggio in situazioni di carenza di dati o di poca affidabilità.

Metodo di BornHuetter-Ferguson

Le grandezze di interesse sono i C_{ij} e si fanno delle ipotesi analoghe al modello della catena:
∘ $C_{ij} > 0 \ \forall i, j$
∘ $C_{i,j+1} \geq C_{ij}$
Si suppone che i C_{ij} siano descritti da un modello moltiplicativo $C_{ij} = \mu_i b_j$, con $b_t = 1$. Quindi si ha che $C_{it} = \mu_i \cdot 1$, ovvero μ_i può essere interpretato come il costo ultimo per i sinistri di origine i. Mentre dalla seconda ipotesi ricaviamo che $C_{i,j} \leq C_{i,j+1} \Leftrightarrow b_j \leq b_{j+1}$ e si interpretano come le aliquote del costo ultimo pagato con differimento minore o uguale a j. In realtà si nota che il modello sottostante è equivalente a quello *Chain Ladder*:

$$\text{BF: } \frac{C_{ij}}{C_{i,j-1}} = \frac{\mu_i b_j}{\mu_i b_{j-1}} = \frac{b_j}{b_{j-1}} \triangleq f_{j-1} \qquad \text{CL: } \frac{C_{ij}}{C_{i,j-1}} = f_{j-1} \Rightarrow C_{ij} = C_{it}\left(\prod_{h=j}^{t-1} f_h\right)^{-1} \triangleq C_{it} b_j$$

La differenza risiede nell'approccio di calcolo dei parametri, infatti $OLL_i = C_{it} - C_{i,t-i} = \mu_i - \mu_i b_{t-i} = \mu_i(1 - b_{t-i})$ è la quota del costo ultimo ancora da pagare.
Stime dei parametri
Il parametro μ_i si ottiene da stime esterne, o da $\widehat{\mu}_i$ dati oppure tramite il metodo *Loss Ratio Semplice*, ovvero $\widehat{\mu}_i = l_i PC_i$, mentre per i $\widehat{b}_j$, si stimano prima i $\widehat{f}_0 \ldots, \widehat{f}_{t-1}$ con il metodo CL, poi stimiamo $\widehat{b}_j^{CL} = (\prod_{h=j}^{t-1} f_h)^{-1}$.
Osservazioni
1 Il costo ultimo stimato $\widehat{C}_{it}$ si ottiene dai dati di portafoglio $C_{i,t-i}$ e dai dati esogeni, per la stima di μ_i, pesati però per $(1 - \widehat{b}_{t-i}^{CL})$ ed infatti $\widehat{C}_{it} = C_{i,t-i} + \mu_i(1 - \widehat{b}_{t-i}^{CL})$.
2 Anche $\widehat{b}_j^{CL} \leq \widehat{b}_{j+1}^{CL}$, quindi $1 - \widehat{b}_j^{CL} \geq 1 - \widehat{b}_{j+1}^{CL}$, cioè il peso conferito alla stima esterna risulta minore per generazioni lontane da 0.
3 Se confrontiamo i costi ultimi:

$$\widehat{C}_{it}^{BF} = C_{i,t-i} + \mu_i(1 - \widehat{b}_{t-i}^{CL}) = C_{i,t-i} \frac{\prod_{h=t-i}^{t-1} \widehat{f}_h}{1} \frac{1}{\prod_{h=t-i}^{t-1} \widehat{f}_h} + \mu_i(1 - \widehat{b}_{t-i}^{CL}) = \widehat{C}_{it}^{CL} \cdot \widehat{b}_{t-i}^{CL} + \widehat{\mu}_i(1 - \widehat{b}_{t-i}^{CL})$$

Quindi si riscontra che attraverso questo approccio il costo ultimo viene espresso tramite una media

ponderata di due stime del costo ultimo: la stima CL basata sui dati della matrice C_{ij} e la stima dai dati esterni μ_i. Si ha dunque una mistura come nelle formule di credibilità, perciò è un metodo molto utilizzato.
4 Nel modello CL $R_i = C_{i,t-i}(\prod_{t-i}^{t-1} f_h - 1)$, quindi per costruzione si dà molto peso alle diagonali e questo comporta che nei rami in cui sono pochi i risarcimenti iniziali, si abbia una sottostima della riserva. Il metodo FB permette di compensare, mediante le informazioni esterne.

Metodo Fisher-Lange

Le grandezze di interesse sono i pagamenti incrementali a valori correnti dell'anno t P^*_{ij} e il numero di sinistri con origine in i e chiusi con differimento j n_{ij}. Si suppone che sussista un effetto moltiplicativo del tipo $P^*_{ij} = n_{ij}\sigma^*_j$. Possiamo interpretare $\sigma^*_j = \frac{P^*_{ij}}{n_{ij}}$ come la quota dei pagamenti mediamente attribuibili a ciascun sinistro. Qui l'ipotesi forte è che σ^*_j dipenda solo da j, cioè che il costo medio per sinistro aggiustato rispetto all'inflazione dipenda solo dal differimento j, indipendentemente dall'anno di origine. Inoltre si riscontra che non vi è coerenza tra numeratore e denominatore di σ^*_j, poiché nel numeratore potrebbero non apparire i possibili pagamenti parziali tra $]i, i+t[$, mentre al denominatore non compaiono i sinistri chiusi dopo $i + j$, che però potrebbero essere stati considerati come pagamenti nel numeratore. Tuttavia se si riscontra dai dati la regolarità che permette di accogliere l'ipotesi, si può applicare il modello. Quindi dati i P_{ij}, stimiamo la matrice dei P^*_{ij} tramite i tassi di inflazione o i coefficienti di adeguamento stimati tramite Taylor. Per creare la matrice dei n_{ij} a partire dai n_i abbiamo due metodi:
1 Si assume un modello moltiplicativo del tipo $n_{ij} = n_i\psi_j$. Si ipotizza quindi che ψ_j è indipendente da i e possiamo interpretarlo come $\psi_j = \frac{n_{ij}}{n_i}$, ovvero la quota dei sinistri chiusi con differimento j. Quindi lo stimiamo da $\sum_i n_{ij} = \sum_i n_i\psi_j$, quindi $\widehat{\psi}_j = \frac{\sum_i n_i}{\sum_i n_{ij}}$, con $i = 0, \ldots, t-j$, ovvero il numero massimo per ogni colonna. A questo punto stimiamo $\widehat{n}_{ij} = n_i\widehat{\psi}_j$, per ogni $i + j > t$.
Se consideriamo la riga i-esima può capitare che $\sum_{j=0}^{t-i} n_{ij} + \sum_{j=t-i+1}^{t} \widehat{n}_{ij} \neq n_i$, ovvero che il totale di riga stimato sia diverso da quello effettivo. Notiamo però che l'ipotesi fatta per il modello di n_{ij}, ci riconduce al modello *Chain Ladder*, infatti $\frac{n_{ij}}{n_{i,j-1}} = \frac{n_i\psi_j}{n_i\psi_{j-1}} = \frac{\psi_j}{\psi_{j-1}} = f_{j-1}$, quindi possiamo stimarli direttamente con il CL e ottenere quindi i risultati di riga coerenti.
Osservazione: $n_{ij} = n_i\psi_j \Leftrightarrow \sum_{h=j}^{t} n_{ih} = \sum_{h=j}^{t} n_i\psi_h = n_i \sum_h \psi_h$ quindi se rapportiamo il metodo 1 a questo otteniamo $\frac{n_{ij}}{\sum_h n_{ih}} = \frac{n_i\psi_j}{n_i\sum_h \psi_h} = \frac{\psi_j}{\sum_h \psi_h} \triangleq \phi_j$, quindi $n_{ij} = \phi_j \sum_{h=j}^{t} n_{ih}$, quindi ϕ_j corrisponde ad una quota dei sinistri con origine in i e chiusi con differimento $\geq j$. Quindi l'ipotesi $n_{ij} = n_i\psi_j$ implica $n_{ij} = \phi_j[n_i - \sum_{h=0}^{j-1} n_{ih}]$, ma non vale il viceversa.
2 Assumiamo $n_{ij} = \phi_j[n_i - \sum_{h=0}^{j-1} n_{ih}]$, stimiamo i parametri ϕ_j:

$$\sum_{i\in I} n_{ij} = \sum_{i\in I} \phi_j[n_i - \sum_{h=0}^{j-1} n_{ih}] \Leftrightarrow \sum_{i\in I} n_{ij} = \phi_j \sum_{i\in I}[n_i - \sum_{h=0}^{j-1} n_{ih}] \Leftrightarrow \widehat{\phi}_j = \frac{\sum_{i\in I} n_{ij}}{\sum_{i\in I}[n_i - \sum_{h=0}^{j-1} n_{ih}]}$$

Quindi stimato $\widehat{n}_{i,t-i+1}$ direttamente dai dati, stimiamo il passo dopo come $\widehat{n}_{i,t-i+2} = \widehat{\phi}_{t-i+2}[n_i - \sum_{h=0}^{t-i} n_{ih} - \widehat{n}_{i,t-i+1}]$ e quello generale come $\widehat{n}_{ij} = \widehat{\phi}_j[n_i - \sum_{h=0}^{t-i} n_{ih} - \sum_{h=t-i+1}^{j-1} \widehat{n}_{ih}]$.
Stimiamo adesso $\widehat{\sigma}^*_j = \sum_{i=0}^{t-j} \frac{P^*_{ij}}{n_{ij}}$ e di conseguenza i $\widehat{P}^*_{ij}$ con $i + j > t$, che poi riportiamo a valori nominali per la normativa.

Il modello di FL ha quindi il vantaggio di esprimere non solo gli importi, ma anche la numerosità del processo di smontamento.

Sinistri IBNR

Si definisce un sinistro tardivo, o IBNR, alla chiusura dell'esercizio θ un sinistro avvenuto nell'anno θ o in anni precedenti, che saranno denunciati in seguito. Questo fenomeno si sviluppa a causa di sinistri che si verificano in prossimità della chiusura dell'esercizio oppure se l'assicurato non è ancora a conoscenza del sinistro (casa vacanze) o anche sinistri che richiedono tempo per manifestarsi.
Tendenzialmente non è un fenomeno rilevante per gli assicuratori diretti, infatti in RCA si ha ogni anno circa il 4% di sinistri IBNR, di cui il 90% viene denunciato entro i primi 3 mesi dalla chiusura. Risulta però un fenomeno significativo per i riassicuratori, può essere infatti che in una copertura XL il risarcimento per un sinistro si valutato in θ $y < L$, ma se poi questo viene rivalutato come $y' > L$ allora il riassicuratore dovrà risarcire.
Osservazione: presa una tabella di *Run-Off*, prendendo come anno d'origine l'anno di accadimento, si ottiene come stima della riserva $R_i = \widehat{C}_{it} - C_{i,t-i}$, per i sinistri avvenuti nell'anno i, e questa comprende la componente degli $IBNR$, mentre se consideriamo C_{it} come il costo ultimo dei sinistri denunciati in i, allora non si tiene conto dei sinistri IBNR. La normativa civilistica sottolinea l'esigenza dell'accantonamento di riserva per i sinistri IBNR, richiede però che siano tenute separate.

Metodo del numero medio per costo medio IBNR

Il metodo si applica quando il fenomeno non è troppo rilevante. Abbiamo dati dall'anno $\theta - T$ e si suppone che il differimento della denuncia sia al massimo di un anno. Vogliamo stimare alla chiusura dell'esercizio θ la riserva, valutando il costo complessivo dei sinistri avvenuti in θ e che saranno denunciati in $\theta + 1$. Siano n_i^{IBNR} il numero di sinistri IBNR avvenuti alla chiusura dell'anno i e n_i il numero di sinistri avvenuti in i. Si ipotizza che $\frac{n_i^{IBNR}}{n_i} \triangleq p$, ovvero la quota di sinistri IBNR, e che questa sia costante al variare di i. Sia C_i^{*IBNR} il costo medio per sinistro IBNR espresso a valori correnti dell'anno θ e lo consideriamo costante C^* per ogni i. Si stima $\hat{p} = \sum_{\theta-T}^{\theta-1} n_i^{IBNR}\omega_i$, con $\omega_i \geq 0$ e $\sum_i \omega_i = 1$ ed anche $\widehat{C}^* = \sum_{i=\theta-T}^{\theta-1} \omega_i' C_i^{*IBNR}$. Noi non conosciamo il totale dei sinistri in θ, però $n_\theta^{IBNR} = n_\theta p = (n_\theta^{IBNR} + n_\theta^D)p$, quindi $n_\theta^{IBNR} = n_\theta^D \frac{\widehat{p}}{1-\widehat{p}}$. A questo punto aggiustiamo opportunamente $\widehat{C}^*$ ed otteniamo $\widehat{C}_\theta^{IBNR}$, con cui stimiamo la riserva $R_i^{IBNR} = \widehat{n}_\theta^{IBNR}\widehat{C}_\theta^{IBNR}$.

Consideriamo adesso di essere a chiusura dell'esercizio θ, ma che ci possano essere sinistri tardivi con massimo differimento d della denuncia dall'accaduto. Si indica con I_{ij}^{IBNR} la stima del costo ultimo per sinistri avvenuti nell'anno i denunciati con differimento j, quindi denunciati in $i+j$, e questa è effettuata all'anno di valutazione d. Per valutare la riserva IBNR bisogna stimare $I_{i,d-i+1}^{IBNR}, \dots, I_{i,d}^{IBNR}$. Si possono usare i metodi visti per la riserva sinistri anche per questa grandezza, se sono ammissibili le ipotesi.
C'è anche una classe di metodi che introduce una misura di esposizione M_i, ovvero una grandezza collegata al fenomeno IBNR, per cui si ipotizza che $\frac{I_{ij}^{IBNR}}{M_i} \triangleq r_j$, e quindi indipendenza dall'anno i. Una possibile grandezza è la competenza premi per l'anno di origine i $M_i = PC_i$ e quindi si stima $\widehat{r}_j$ come una media ponderata $\widehat{r}_j = \sum_{i=0}^{d-j} \omega_i^{(j)} \frac{I_{ij}}{PC_i}$. Calcolati quindi $\widehat{r}_j$, possiamo proiettare gli $I_{ij}^{IBNR} = PC_i\widehat{r}_j$ e di conseguenza stimare la riserva come $R_i^{IBNR} = \sum_{j=d-i+1}^{d} I_{ij}^{IBNR} = PC_i \sum_{j=d-i+1}^{d} \widehat{r}_j$.

Metodo di Tarbel

Questo metodo viene utilizzato quando il massimo differimento è 1. É possibile stimare alla chiusura dell'anno θ in questo modo la riserva: $R_\theta^{IBNR} = I_{\theta-1}^{IBNR}\underbrace{n_{10,11,12}^{(\theta)}C_{10,11,12}^{(\theta)}}_{M_\theta}$.

Ovvero considerare come misura di esposizione $n_{10,11,12}^{(\theta)}C_{10,11,12}^{(\theta)}$, ovvero le grandezze degli ultimi tre mesi dell'anno.