

PERSONALIZZAZIONE DEL PREMIO BASATA SULL'ESPERIENZA

La **personalizzazione a priori** nella determinazione dei premi assicurativi ha l'obiettivo di ripartire un portafoglio in classi di rischi con forti caratteristiche di analogia e di determinare i premi da attribuire ai rischi di ciascuna classe.

I premi sono differenziati in funzione di caratteristiche specifiche dei rischi ritenute influenti sulla sinistrosità, **fattori di rischio** o **variabili tariffarie**, osservabili *a priori*: prima di disporre di informazioni sulla sinistrosità.

I rischi sono ripartiti in **classi tariffarie**, si ottiene così la **struttura della tariffa**, e si valutano i premi da richiedere agli assicurati di ciascuna classe, **premi a priori** o **premi collettivi**. Si ottiene la **tariffa**.

In alcune coperture, nell'ambito di ogni classe tariffaria rimane una notevole eterogeneità residua nella sinistrosità, causata da fattori non osservabili. Ad esempio in RCA: abilità nella guida, temperamento del guidatore, prontezza di riflessi, conoscenza e rispetto delle norme del codice della strada, abitudini di vita,

Inoltre, ci sono fattori di rischio che non possono essere utilizzati come variabili tariffarie, ad esempio, il sesso dell'assicurato.

Al fine di formulare previsioni sulla sinistrosità di un individuo può essere molto utile tenere conto della sua personale storia di sinistrosità. L'esperienza individuale di sinistrosità viene vista come rivelatrice di caratteristiche individuali, non osservabili *a priori*, che influenzano la sinistrosità e determinano l'eterogeneità residua nelle classi tariffarie.

Si attua una ulteriore differenziazione dei premi. I premi *a priori* sono modificati tenendo conto dell'esperienza individuale di ciascun assicurato: si riconoscono sconti agli assicurati virtuosi, si penalizzano con aggravi gli altri.

Si passa dunque dal premio *a priori* (o premio collettivo: comune ai rischi di una classe tariffaria) ad un **premio basato sull'esperienza individuale**, con l'obiettivo di valutare in modo migliore i rischi e dunque di fare pagare a ciascun assicurato, nel tempo, un premio che corrisponda maggiormente al profilo di rischio dell'assicurato.

Premio *a priori* o collettivo



Storia individuale di sinistrosità

Premio a posteriori o basato sull'esperienza

Si parla di **personalizzazione basata sull'esperienza** (*experience rating* o *merit rating*). Può essere ottenuta seguendo diversi approcci

- approccio Bayesiano,
- approccio della Teoria della Credibilità,
- sistemi Bonus-Malus; *No-claim discount*.

Richiami. In uno spazio di probabilità,

$$(\mathbb{P}, \mathcal{A}, Pr(\cdot)),$$

(i) sia Y un numero aleatorio (misurabile),

$$Y(\cdot): \mathbb{P} \rightarrow \mathbb{R}, \quad Y^{-1}(B) \in \mathcal{A}, \text{ per ogni } B \in \mathcal{B} \text{ boreiano di } \mathbb{R},$$

\Rightarrow def. di misurabilità

tale che $E(|Y|) < +\infty$, (in particolare, $Y = |A|$, indicatore di evento);
 $\hookrightarrow Y \in \mathcal{L}^1$

(ii) sia U un altro numero aleatorio (misurabile),

$$U(\cdot): \mathbb{P} \rightarrow \mathcal{X} \subset \mathbb{R}, \quad \mathcal{X} \text{ insieme immagine di } U(\cdot).$$

Se U ha un numero finito di determinazioni e $Pr(U = u) > 0$, per ogni determinazione di U , si pone

Altrimenti $Pr(U=u)=0$

$$E(Y|U = u) = \frac{E(Y|U = u)}{Pr(U = u)}.$$

Segue

$$E(Y) = \sum_u E(Y|U = u) Pr(U = u).$$

In tale caso si ha dunque la proprietà di disintegrità del valore atteso rispetto alla partizione associata a U .

Funzione di regressione. Si prova che, nelle ipotesi (i) e (ii), esiste una funzione, indicata con $E(Y|U = \cdot)$, detta **funzione di regressione** di Y su U ,

$$E(Y|U = \cdot) : \mathcal{X} \subset \mathbb{R} \rightarrow \mathbb{R},$$

tale che

- $E(Y|U = \cdot)$ è P_U integrabile, dove P_U è la misura immagine: $P_U(]-\infty, u]) = Pr(U \leq u)$, ed è unica a meno di insiemi P_U -trascurabili,
- $E(Y|U \in B) = \int_{\{U \in B\}} Y dPr = \int_B E(Y|U = u) P_U(du) = \int_B E(Y|U = u) dF_U(u)$, per ogni $B \in \mathcal{B}$,
dove l'ultima uguaglianza sussiste se $E(Y|U = \cdot)$ è Riemann-Stieltjes integrabile. In tale caso, in particolare,

Proprietà di disintegrità generale

$$E(Y) = \int_{\mathbb{R}} E(Y|U = u) dP_U(du) = \int_{\mathbb{R}} E(Y|U = u) dF_U(u). \quad (3.1)$$

Collegamento della (3.1) con la proprietà di disintegrità della speranza matematica.

La funzione di regressione verifica le seguenti proprietà

$$E(\alpha|U = \cdot) = \alpha, \text{ se } \alpha \text{ è certo},$$

$$E(Y_1|U = \cdot) \leq E(Y_2|U = \cdot), \text{ se } Y_1 \leq Y_2,$$

$$E(Y_1|U = \cdot) = E(Y_2|U = \cdot), \text{ se } Y_1 = Y_2 \text{ qc},$$

$$a \leq E(Y|U = \cdot) \leq b, \text{ se } a \leq Y \leq b,$$

$$E(\sum_{i=1}^n \alpha_i Y_i | U = \cdot) = \sum_{i=1}^n \alpha_i E(Y_i | U = \cdot),$$

$$E(h(U)Y|U = \cdot) = h(U)E(Y|U = \cdot),$$

$$E(Y|U = \cdot) = E(Y), \text{ se } Y, U \text{ sono stocasticamente indipendenti.}$$

Proprietà speranze
matematiche subordinate

Consideriamo ora la seguente composizione

$$\begin{array}{ccccc} & & = E[Y|U] & & \\ \mathbb{P} & \xrightarrow{U(\cdot)} & \mathcal{X} & \xrightarrow{E(Y|U=\cdot)} & \mathbb{R} \\ \omega & \longrightarrow & U(\omega) = u & \longrightarrow & E(Y|U=u) \end{array}$$

Scrivo così per sottolineare che U è una v.a.

Rimane definito su \mathbb{P} un numero aleatorio, che indichiamo con $E(Y|U)$, tale che

$$\omega \in \mathbb{P} \longrightarrow E(Y|U = U(\omega)). \quad (3.2)$$

La speranza matematica di $E(Y|U)$

$$E[E(Y|U)] = \int_{\mathbb{R}} E(Y|U=u) dF_U(u) = E(Y). \quad (3.3)$$

Verifico le proprietà di disintegrazibilità

↑
per (3.1)

Speranza matematica condizionata ad una σ -algebra. Sia \mathcal{H} una sotto- σ -algebra di \mathcal{A} . Si prova che, nelle ipotesi (i) e (ii), esiste un numero aleatorio, indicato con $E(Y|\mathcal{H})$, definito a meno di eventi di \mathcal{H} di probabilità nulla, tale che

- $E(Y|\mathcal{H})$ è \mathcal{H} -Borel misurabile,
- $E(Y|H) = \int_H Y dPr = \int_H E(Y|\mathcal{H}) dPr$, per ogni $H \in \mathcal{H}$. In particolare,

$$E(Y) = \int_{\Omega} Y dPr = \int_{\Omega} E(Y|\mathcal{H}) dPr = E[E(Y|\mathcal{H})]. \quad (3.4)$$

Dato il numero aleatorio U , misurabile, si può considerare la σ -algebra indotta da U , $\mathcal{U} = U^{-1}(\mathcal{B})$, per $B \in \mathcal{B}$, e la speranza matematica di Y condizionata a \mathcal{U} , $E(Y|\mathcal{U})$ che viene anche indicata con $E(Y|U)$.

~~✓~~

$$U: (\Omega, \mathcal{A}, P_r) \rightarrow (\mathbb{R}, \mathcal{B}, \mathbb{P})$$

$$\text{Andiamo a considerare } (U^{-1}(\mathcal{B}), \mathcal{B} \in \mathcal{B}) = U^{-1}(\mathcal{B}) \subset \mathcal{A}$$

Per la (3.4), si ha

$$E(Y) = E[E(Y|\mathcal{U})] = E[E(Y|U)].$$

Si veda la (3.3). Si prova che il numero aleatorio $E(Y|U)$ definito dalla (3.2) è una versione di $E(Y|\mathcal{U})$.

Nel seguito il numero aleatorio indicato con $E(Y|U)$, potrà essere inteso come il numero aleatorio definito dalla (3.2) o come $E(Y|\mathcal{U})$,

$$\begin{array}{c}
 \omega \rightarrow U(\omega) = u \rightarrow \overbrace{E[Y|U=u]}^{\substack{\text{valore in } u \text{ della funzione} \\ \text{di regressione} \\ \text{di } Y \text{ su } U}} \\
 \nearrow \\
 E(Y|U) \\
 \searrow \\
 E(Y|\mathcal{U}) \text{ speranza matematica di } Y \text{ condizionata} \\
 \text{alla } \sigma\text{-algebra indotta da } U
 \end{array}$$

In tutti i casi si ha

$$E(Y) = E[E(Y|U)] = \int_{\mathbb{R}} E(Y|U=u) dF_U(u).$$

Le precedenti nozioni si estendono:

- $E(Y|U_1 = \cdot, \dots, U_n = \cdot)$ funzione di regressione di Y su U_1, \dots, U_n ,
- $E(Y|U_1, \dots, U_n)$, speranza matematica di Y condizionata alla σ -algebra indotta da U_1, \dots, U_n ,

con

$$E(Y) = E[E(Y|U_1, \dots, U_n)].$$

o decomposizione

Proprietà di scomposizione della varianza. Nelle ipotesi (i) e (ii), supponiamo che il numero aleatorio Y sia tale che $E(Y^2)$ sia finita. Allora esiste $E(Y^2|U)$ e sussiste la $E(Y^2) = E[E(Y^2|U)]$.

Si pone

$$var(Y|U) = E(Y^2|U) - E(Y|U)^2.$$

Segue

$$E[var(Y|U)] = E[E(Y^2|U)] - E[E(Y|U)^2] = E(Y^2) - E[E(Y|U)^2]$$

e

$$var[E(Y|U)] = E[E(Y|U)^2] - (E[E(Y|U)])^2 = E[E(Y|U)^2] - E(Y)^2$$

Sommando membro a membro, si ottiene

$$E[var(Y|U)] + var[E(Y|U)] = E(Y^2) - E(Y)^2 = var(Y).$$

■

APPROCCIO BAYESIANO ALLA PERSONALIZZAZIONE BASATA SULL'ESPERIENZA

Consideriamo una collettività di rischi “analoghi”, omogenei rispetto a caratteristiche osservabili *a priori*, ed “indipendenti”. Nella collettività vi sia eterogeneità residua.

Per un fissato assicurato nella collettività, l' i -esimo, sia Y_{it} un numero aleatorio di interesse legato alla sinistrosità nell’anno t , ad esempio

- $Y_{it} = N_{it}$ il numero di sinistri nell’anno t ,
- $Y_{it} = X_{it}$ il risarcimento totale nell’anno t .

Al variare di t , si ottiene un processo stocastico a parametro discreto,

$$\{Y_{i1}, Y_{i2}, \dots\}.$$

Per tenere conto dell’eterogeneità residua dovuta a fattori **non osservabili**, si introduce un parametro **aleatorio**, U_i , specifico per il fissato assicurato, che è detto **parametro di rischio o componente di eterogeneità**.

U_i può essere per esempio

{
 • capacità riflessi
 • attenzione alle guida
 • sobrio o meno

Dato che U_i è legato al **profilo di rischio** dell'assicurato, è naturale supporre che la valutazione probabilistica sulla sinistrosità dell'individuo dipenda dal valore del parametro di rischio. Si assume che, se il valore di U_i fosse noto, attribuiremmo una specifica legge al processo dei numeri aleatori d'interesse.

Introduciamo le seguenti ipotesi:

Ipotesi: d'indipendenze subordinata

- $Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots$ siano stocasticamente indipendenti e di legge assegnata, per ogni u determinazione di U_i ,
- sia assegnata la legge di U_i , (la funzione di ripartizione F_{U_i} o la funzione di densità f_{U_i} o la funzione di probabilità nel caso discreto).

La funzione di ripartizione di U_i è detta **funzione di struttura**.

Con le ipotesi dette, rimane assegnata la legge del processo $\{Y_{i1}, Y_{i2}, \dots\}$, ed è la legge mistura: dato A evento logicamente dipendente dal processo,

$$Pr(A) = \int Pr(A|U_i = u)dF_{U_i}(u).$$

Abbiamo fissato l'attenzione su un assicurato della collettività, ma per ciascuno degli assicurati si ha un processo di numeri aleatori di interesse ed un parametro di rischio.

In forza dei giudizi qualitativi di “analogia” ed “indipendenza”, si assume che i processi $\{U_i, Y_{i1}, Y_{i2}, \dots\}$, $i = 1, 2, \dots$, siano

- identicamente distribuiti, più in dettaglio che
 - per ogni u , i processi $\{Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots\}$ siano identicamente distribuiti al variare di i ,
 - i numeri aleatori U_i siano identicamente distribuiti al variare di i ,
- stocasticamente indipendenti: informazioni sulla storia di sinistrosità degli altri assicurati non portino a modificare la valutazione probabilistica del processo per l'assicurato i -esimo.

Si può allora fissare l'attenzione su un fissato assicurato della collettività: "individuo campione".

Per semplificare la notazione, nel seguito, omettiamo l'indice dell'individuo. Sia

$$\{U, \underbrace{Y_1, Y_2, \dots}_{\substack{\text{Var. d'interesse} \\ \hookrightarrow \text{Var che descrive l'eterogeneità}}}\}$$

il processo del parametro aleatorio di rischio e dei numeri aleatori di interesse, con le ipotesi

- $Y_1|U = u, Y_2|U = u, \dots$ stocasticamente indipendenti, di legge assegnata, per ogni u determinazione possibile di U , \rightarrow Subordinatamente indipendenti ma non sono stocasticamente indipendenti
- U di legge assegnata.

Osserviamo che i numeri aleatori del processo $\{Y_1, Y_2, \dots\}$ **non sono stocasticamente indipendenti**. Ad esempio, fissato n ,

$$F_{Y_1, \dots, Y_n}(y_1, \dots, y_n) = \int Pr(Y_1 \leq y_1, \dots, Y_n \leq y_n | U = u) dF_U(u)$$

$$= \int \prod_{t=1}^n Pr(Y_t \leq y_t | U = u) dF_U(u)$$

↑ per l'indipendenza condizionata

$$\neq \prod_{t=1}^n \int Pr(Y_t \leq y_t | U = u) dF_U(u) = \prod_{t=1}^n Pr(Y_t \leq y_t).$$

← indipendenti subordinatamente a U

Pertanto, le ipotesi introducono una particolare struttura di dipendenza stocastica tra i numeri aleatori del processo $\{Y_1, Y_2, \dots\}$: la dipendenza è dovuta unicamente al fatto che non si conosce il valore di U . Se fosse noto il valore del parametro di rischio, sarebbe noto il profilo di rischio dell'assicurato, e, in tale caso, i numeri aleatori di interesse sarebbero ritenuti stocasticamente indipendenti.

Trattiamo il problema di valutare il “premio equo” in $T + 1$, inteso come la valutazione per la grandezza di interesse nell’anno $T + 1$:

$\nearrow N_{T+1}$ numero dei sinistri: “premio equo per la frequenza”
 $\downarrow Y_{T+1}$
 $\searrow X_{T+1}$ risarcimento totale: premio equo

Notiamo che

$$E(Y_{T+1})$$

è il **premio equo a priori**, il **premio collettivo**, comune agli assicurati della collettività. *Dei rischi analoghi:*

Ma questo premio non corrispondente al profilo di rischio dell’assicurato. Infatti, idealmente, se potessimo conoscere il valore del parametro di rischio e questo fosse u , il premio dovrebbe essere

$$E(Y_{T+1}|U = u) = \mu_{T+1}(u).$$

Tale premio è detto **premio individuale** in $T + 1$.

Il valore di U non è però osservabile, allora il **premio individuale in $T + 1$ è aleatorio**

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

Osserviamo che il premio collettivo è il valore atteso del premio individuale

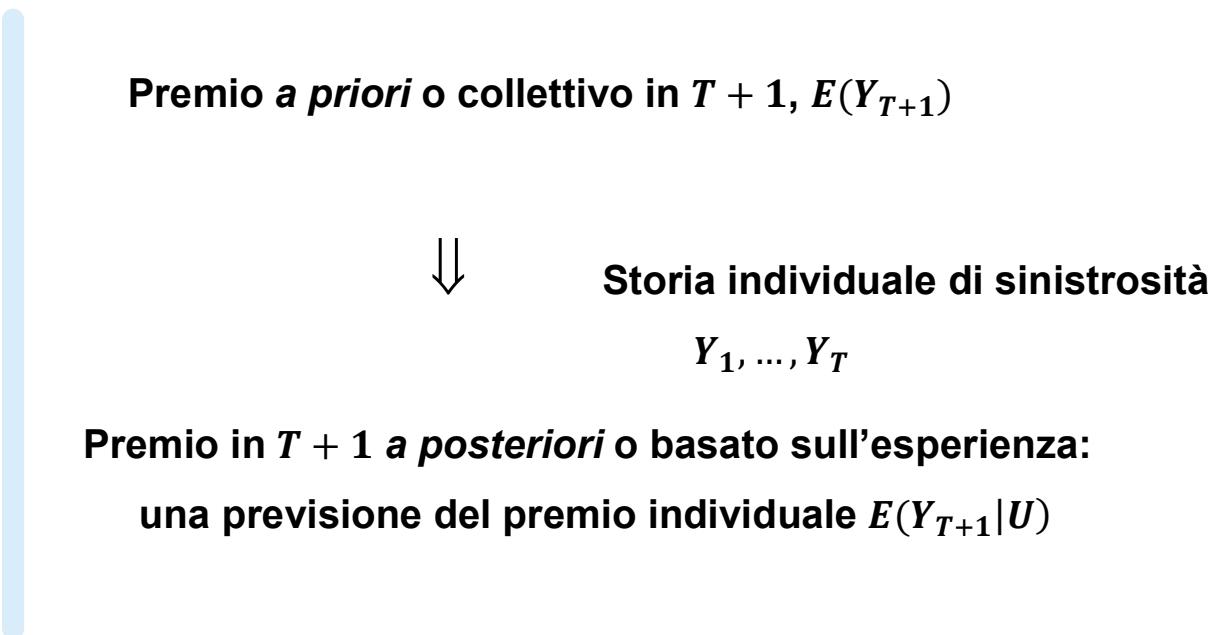
$$E(Y_{T+1}) = E[E(Y_{T+1}|U)] = E[\mu_{T+1}(U)].$$

Al fine di stimare il premio individuale $E(Y_{T+1}|U)$, si potrebbe utilizzare il premio *a priori* $E(Y_{T+1})$.

Ma poiché i numeri aleatori del processo di interesse non sono stocasticamente indipendenti, informazioni sulla storia passata, su Y_1, \dots, Y_T , possono aiutare a migliorare la valutazione su Y_{T+1} , rispetto alla valutazione *a priori*.

Le serie storiche sono correlate seriamente → I dati del passato possono aiutare a prevedere quelli del futuro

Obiettivo della personalizzazione basata sull'esperienza: utilizzare le osservazioni sul processo dei numeri aleatori di interesse per passare dal premio collettivo ad un premio che meglio rispecchi il profilo di rischio dell'assicurato.



Se Y_1, \dots, Y_T sono numeri aleatori discreti e y_1, \dots, y_T i valori osservati, posto $H_T = (Y_1 = y_1, \dots, Y_T = y_T)$, con $\Pr(H_T) > 0$, come stima del premio individuale possiamo considerare il premio **a posteriori** o **premio bayesiano**

$E(Y_{T+1}|H_T)$. *Non è una variabile aleatoria
E' un evento, non ha una σ-algebra propria*

Si ha

$$\begin{aligned} E(Y_{T+1}|H_T) &= \int \underbrace{E(Y_{T+1}|H_T, U=u)}_{\substack{\uparrow \text{indipendenza condizionata} \\ \rightarrow \text{togliendo } H_T}} dF_{U|H_T}(u) \\ &= E(Y_{T+1}|U=u) \\ &= \int \mu_{T+1}(u) dF_{U|H_T}(u) = E[\mu_{T+1}(U)|H_T]. \end{aligned}$$

Notiamo che se U ha distribuzione dotata di densità f_U (**densità a priori**)

$$\int \mu_{T+1}(u) dF_{U|H_T}(u) = \int \mu_{T+1}(u) f_{U|H_T}(u) du,$$

dove, per il teorema di Bayes, la **densità a posteriori** è

$$f_{U|H_T}(u) = \frac{f_U(u) \Pr(H_T|U=u)}{\Pr(H_T)} = \frac{f_U(u) \Pr(H_T|U=u)}{\int \Pr(H_T|U=u) f_U(u) du}.$$

*Densità a priori verso in giù
↓ Pr(H_T|U=u) sottostante dell'ipotesi*

Premio collettivo $E(Y_{T+1}) = E[\mu_{T+1}(U)]$ valore atteso di $\mu_{T+1}(U)$ calcolato con la distribuzione *a priori*

$$\downarrow H_T$$

Premio bayesiano $E(Y_{T+1}|H_T) = E[\mu_{T+1}(U)|H_T]$ valore atteso di $\mu_{T+1}(U)$ calcolato con la distribuzione *a posteriori*

La valutazione del profilo di rischio dell'assicurato, modellato tramite il parametro U , è aggiornata tramite il meccanismo bayesiano

$$\begin{aligned} U &\rightarrow U|H_T \\ F_U(f_U) &\rightarrow F_{U|H_T}(f_{U|H_T}) \end{aligned}$$

ciò porta a modificare il premio.

Si noti che, considerati gli assicurati i, j , si ha $U_i =^d U_j$, ma, anche a parità di durata dell'osservazione, in generale, $U_i|H_{iT} \neq^d U_j|H_{jT}$.

Più in generale, si considera lo **stimatore bayesiano del premio individuale in $T + 1$**

St. matore var. elettrice
 $\omega \rightarrow (Y_1, \dots, Y_T)(\omega) \rightarrow \overbrace{E[\mu_{T+1}(U)|Y_1 = Y_1(\omega), \dots, Y_T = Y_T(\omega)]}^{\substack{\text{valore della funzione di regressione} \\ \text{di } \mu_{T+1}(U) \text{ su } Y_1, \dots, Y_T \\ \text{o valore atteso condizionato}}}$
 $\widetilde{\mu_{T+1}(U)} = E[\mu_{T+1}(U)|Y_1, \dots, Y_T]$

\nearrow
 \downarrow
 speranza matematica di $\mu_{T+1}(U)$
 condizionata alla σ -algebra indotta da Y_1, \dots, Y_T

Nelle ipotesi del modello, si ha

$$\widetilde{\mu_{T+1}(U)} = E[\mu_{T+1}(U)|Y_1, \dots, Y_T] = E[Y_{T+1}|Y_1, \dots, Y_T].$$

(Lo si è visto nel caso particolare. In generale segue dalle proprietà della speranza matematica condizionata: se \mathcal{K} è sotto σ -algebra di \mathcal{H} , $E(Y|\mathcal{K}) = E[E(Y|\mathcal{H})|\mathcal{K}]$).

Richiami. Dati X un numero aleatorio non osservabile e Y_1, \dots, Y_T una sequenza di numeri aleatori osservabili, si pone il problema di “stimare X a partire da Y_1, \dots, Y_T , commettendo un errore più piccolo possibile”. Spesso come misura dell’errore si considera la **perdita quadratica attesa** e si cerca

$$\operatorname{argmin}_{g \in \mathcal{G}} E[(X - g(Y_1, \dots, Y_T))^2].$$

Se $\operatorname{var}(X) < +\infty$ e \mathcal{G} è l’insieme delle funzioni a quadrato $P_{(Y_1, \dots, Y_T)}$ -integrabile, il precedente problema ammette soluzione, unica a meno di insiemi trascurabili. La soluzione è la funzione di regressione di X su Y_1, \dots, Y_T

$$g^*(y_1, \dots, y_T) = E[X|Y_1 = y_1, \dots, Y_T = y_T].$$

Allora, lo stimatore di X che consente di rendere minima la perdita quadratica attesa è

$$E[X|Y_1, \dots, Y_T],$$

che è anche detto **stimatore dei minimi quadrati**. ■

Nel nostro problema, il numero aleatorio non osservabile è il premio individuale $\mu_{T+1}(U)$ e i numeri osservabili sono Y_1, \dots, Y_T , la sequenza dei numeri aleatori del processo di interesse.

Allora, lo **stimatore bayesiano del premio individuale** $\widetilde{\mu_{T+1}(U)} = E[\mu_{T+1}(U)|Y_1, \dots, Y_T]$ è la **funzione del processo di osservazione che approssima al meglio il premio individuale $\mu_{T+1}(U)$ nel senso che rende minima la perdita quadratica attesa**.

Si noti che possiamo limitarci a tenere conto della storia di ciascun assicurato singolarmente e non dell'esperienze del collettivo, per l'ipotesi di indipendenza stocastica dei processi relativi ai diversi assicurati.

Modello Poisson-gamma per il processo dei numeri di sinistri

Consideriamo una collettività di rischi “analoghi”, omogenei rispetto a caratteristiche osservabili *a priori*, ed “indipendenti”. Nella collettività vi sia eterogeneità residua.

Per un fissato assicurato nella collettività, l’ i -esimo, sia N_{it} il numero di sinistri nell’anno t e U_i il parametro aleatorio di rischio. Assumiamo

- a) $\{U_i, N_{i1}, N_{i2}, \dots\}$ processi iid al variare di i , *Non le variabili all'interno del processo*
 - b) fissato i ,
 - b1) $N_{i1}|U_i = u, N_{i2}|U_i = u, \dots$ stocasticamente indipendenti per ogni u determinazione di U_i ,
 - b2) $N_{it}|U_i = u \sim P(\lambda u)$, per ogni t , $\lambda > 0$,
 - b3) $U_i \sim \text{gamma}(\alpha, \alpha)$. *→ ha una potenzialità di semplificazione*
↳ E' una gemme per motivi di calcolo, ha media 1
- Il processo $\{N_{i1}, N_{i2}, \dots\}$, con le ipotesi b1), b2), b3), ha legge che dipende dai due parametri λ, α , che coerentemente con l’ipotesi a), non dipendono da i . E’ detto **processo Poisson-gamma (λ, α)**.

Si ha

- $E(U_i) = 1, \text{var}(U_i) = 1/\alpha$
- $E(N_{it}|U_i = u) = \lambda u, E(N_{it}|U_i) = \lambda U_i$ **premio individuale**
- $E(N_{it}) = E [E(N_{it}|U_i)] = E(\lambda U_i) = \lambda$ **premio a priori o collettivo**
- $\text{var}(N_{it}) = E[\text{var}(N_{it}|U_i)] + \text{var}[E(N_{it}|U_i)] = E(\lambda U_i) + \text{var}(\lambda U_i) = \lambda + \alpha^{-1}\lambda^2$

Notiamo che, fissato λ , minore è α , maggiore è la varianza di U_i , maggiore la dispersione del premio individuale: maggiore l'eterogeneità nell'ambito della collettività.

Il processo $\{N_{i1}, N_{i2}, \dots\}$ può essere visto come il processo degli incrementi annui di un processo mistura di poissoniani con misturante $\text{gamma}(\alpha, \alpha)$, allora sappiamo che

$$N_{it} \sim BN\left(\alpha, \frac{\alpha}{\alpha + \lambda}\right)$$

↗ Distribuzione non subordinata

e che i numeri aleatori del processo non sono stocasticamente indipendenti.

Il processo Poisson-gamma è un **processo scambiabile**.

Per verificare la scambiabilità, fissiamo l'attenzione su un assicurato e, per semplicità di notazione, omettiamo l'indice.

Il processo $\{N_1, N_2, \dots\}$ è **scambiabile** se, per ogni $T > 0$, per ogni $(\pi(1), \dots, \pi(T))$ permutazione di $(1, \dots, T)$, riesce $(N_1, \dots, N_T) =^d (N_{\pi(1)}, \dots, N_{\pi(T)}) \Leftrightarrow$ per ogni (n_1, \dots, n_T) , es. $(N_1, N_2, N_3) \stackrel{d}{=} (N_2, N_3, N_1)$

$$Pr(N_1 = n_1, \dots, N_T = n_T) = Pr(N_{\pi(1)} = n_1, \dots, N_{\pi(T)} = n_T).$$

Fissiamo T , (n_1, \dots, n_T) e consideriamo

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

$$= \int_0^{+\infty} \underbrace{\prod_{t=1}^T Pr(N_t = n_t | U = u)}_{\substack{\uparrow \\ \text{indipendenza condizionata}}} \frac{\alpha^\alpha}{\Gamma(\alpha)} u^{\alpha-1} e^{-\alpha u} du$$

$$\prod_{t=1}^T e^{-\lambda u} \frac{(\lambda u)^{n_t}}{n_t!} = e^{-\lambda u T} \frac{(\lambda u)^{\sum_{t=1}^T n_t}}{\prod_{t=1}^T n_t!}$$

per le ip. di Poisson - gamma

$$\begin{aligned}
Pr(N_1 = n_1, \dots, N_T = n_T) &= \int_0^{+\infty} e^{-\lambda u T} \frac{(\lambda u)^{\sum_{t=1}^T n_t}}{\prod_{t=1}^T n_t!} \frac{\alpha^\alpha}{\Gamma(\alpha)} u^{\alpha-1} e^{-\alpha u} du \\
&= \frac{\alpha^\alpha \lambda^{\sum_{t=1}^T n_t}}{\Gamma(\alpha) \prod_{t=1}^T n_t!} \int_0^{+\infty} \underbrace{u^{\alpha + \sum_{t=1}^T n_t - 1} e^{-(\alpha + \lambda T)u}}_{\text{nucleo della } gamma(\alpha + \sum_{t=1}^T n_t, \alpha + \lambda T)} du \\
&\quad \uparrow
\end{aligned}$$

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

$$= \frac{\Gamma(\alpha + \sum_{t=1}^T n_t)}{\Gamma(\alpha) \prod_{t=1}^T n_t!} \left(\frac{\alpha}{\alpha + \lambda T} \right)^\alpha \left(\frac{\lambda}{\alpha + \lambda T} \right)^{\sum_{t=1}^T n_t}$$

Nei precedenti passaggi, sostituendo (N_1, \dots, N_T) con $(N_{\pi(1)}, \dots, N_{\pi(T)})$ si arriva al medesimo risultato.

Processo Poisson-gamma, aggiornamento bayesiano,

Ancora per il fissato assicurato, determiniamo il premio *a posteriori* in $T + 1$, supponendo di avere osservato $H_T = (N_1 = n_1, \dots, N_T = n_T)$.

Ricordiamo che

- il premio *a priori* è $E(N_{T+1}) = \lambda$,
- il premio individuale è $E(N_{T+1}|U) = \mu_{T+1}(U) = \lambda U$,
- il premio *a posteriori* è

$$E(N_{T+1}|H_T) = E(\mu_{T+1}(U)|H_T) = E(\lambda U|H_T) = \lambda E(U|H_T),$$

La densità *a posteriori* del parametro di rischio è

$$\begin{aligned} f_{U|H_T}(u) &= \frac{f_U(u)Pr(H_T|U=u)}{Pr(H_T)} \propto f_U(u)Pr(H_T|U=u) \\ &= \frac{\alpha^\alpha}{\Gamma(\alpha)} u^{\alpha-1} e^{-\alpha u} e^{-\lambda u T} \frac{(\lambda u)^{\sum_{t=1}^T n_t}}{\prod_{t=1}^T n_t!} \propto u^{\alpha-1} e^{-\alpha u} e^{-\lambda u T} u^{\sum_{t=1}^T n_t} = u^{\alpha+\sum_{t=1}^T n_t-1} e^{-(\alpha+\lambda T)u} \end{aligned}$$

Pertanto,

$$f_{U|H_T} \sim \text{gamma}(\alpha + \sum_{t=1}^T n_t, \alpha + \lambda T)$$

Si ha che la distribuzione *a posteriori* è ancora una gamma, ma con parametri aggiornati:

$$U \sim \text{gamma}(\alpha, \alpha)$$

con

$$E(U|H_T) = \frac{\alpha + \sum_{t=1}^T n_t}{\alpha + \lambda T}$$

$$U|H_T \sim \text{gamma}(\alpha + \sum_{t=1}^T n_t, \alpha + \lambda T)$$

↗ $E(N_{T+1})$
 ↳ numero atteso di sx. in $t=1, \dots, T$
 ↳ Numero effettivo di sx.
 Se $\sum_{t=1}^T n_t = \lambda T$ allora $\alpha' = \alpha''$
 $E(U|H_T) = U$

Il premio *a posteriori* o bayesiano in $T + 1$ è

$$E(N_{T+1}|H_T) = \lambda E(U|H_T) = \lambda \frac{\alpha + \sum_{t=1}^T n_t}{\alpha + \lambda T}$$

Notiamo che dipende dalla valutazione *a priori* tramite i parametri λ, α e dalla storia individuale di sinistrosità tramite $T, \sum_{t=1}^T n_t$.

Abbiamo isolato la storia della sinistrosità

Riesce

$$\begin{aligned} E(N_{T+1}|H_T) &= \lambda \frac{\alpha + \sum_{t=1}^T n_t}{\alpha + \lambda T} \\ &= \lambda \frac{\alpha}{\alpha + \lambda T} + \frac{\sum_{t=1}^T n_t}{T} \frac{\lambda T}{\alpha + \lambda T}. \end{aligned}$$

Pesi

Il premio *a posteriori* è mistura di

$$\lambda = E(N_{T+1})$$

premio *a priori*

$$\frac{\sum_{t=1}^T n_t}{T}$$

numero medio annuo di sinistri osservati,
sintesi della storia di sinistrosità

I pesi della mistura sono

$$1 - z_T = \frac{\alpha}{\alpha + \lambda T} \quad z_T = \frac{\lambda T}{\alpha + \lambda T}$$

Si noti che,

- fissati λ, α, z_T è crescente al crescere di T e $z_T \rightarrow 1$, per $T \rightarrow +\infty$,
- fissati λ, T, z_T è decrescente al crescere di α : minore è l'eterogeneità residua, minore il peso attribuito alla storia individuale di sinistrosità, maggiore è l'eterogeneità residua, maggiore il peso attribuito alla storia individuale di sinistrosità.

Nel premio *a posteriori*,

$$E(N_{T+1}|H_T) = \lambda \frac{\alpha + \sum_{t=1}^T n_t}{\alpha + \lambda T},$$

il fattore

$$c_{T+1}(H_T) = \frac{\alpha + \sum_{t=1}^T n_t}{\alpha + \lambda T} = \frac{E(N_{T+1}|H_T)}{E(N_{T+1})}$$

è detto **coefficiente di aggiornamento o revisione bayesiana**: è il fattore che, moltiplicato per il premio *a priori*, fornisce il premio *a posteriori*. E' anche detto **coefficiente bonus-malus bayesiano**.

Osservazioni.

- Si ha

$$c_{T+1}(H_T) > 1 \Leftrightarrow \sum_{t=1}^T n_t > \lambda T$$

$$\begin{array}{c} \uparrow \\ E(N_1 + \dots + N_T) \end{array}$$

- La storia di sinistrosità interviene con la sintesi $\sum_{t=1}^T n_t$.
- Assicurati diversi con lo stesso T e con la stessa storia di sinistrosità, anzi, con lo stesso numero totale osservato di sinistri, avranno lo stesso premio a posteriori, altrimenti no.

Esempio. In classe tariffaria 1 si abbia $\lambda = 0.06$, $\alpha = 0.5$

Indice di variabilità

T	Coefficients of Bayesian revision						
	Number of claims						
	0	1	2	3	4	5	6
1	0.89	2.68	4.46	6.25	8.04	9.82	11.61
5	0.63	1.88	3.13	4.38	5.63	6.88	8.13
10	0.45	1.36	2.27	3.18	4.09	5.00	5.91
15	0.36	1.07	1.79	2.50	3.21	3.93	4.64

In classe tariffaria 2 si abbia $\lambda = 0.10$, $\alpha = 0.5$

T	Coefficients of Bayesian revision						
	Number of claims						
	0	1	2	3	4	5	6
1	0.83	2.50	4.17	5.83	7.50	9.17	10.83
5	0.50	1.50	2.50	3.50	4.50	5.50	6.50
10	0.33	1.00	1.67	2.33	3.00	3.67	4.33
15	0.25	0.75	1.25	1.75	2.25	2.75	3.25

I coefficienti di revisione della classe tariffaria 1 sono sempre più elevati di quelli della classe 2:

- se due rischi non provocano sinistri, il rischio in classe 1 ottiene una riduzione di premio inferiore del rischio in classe 2,
- se entrambi i rischi provocano lo stesso numero di sinistri il rischio in classe 1 riceve una penalizzazione maggiore del rischio in classe 2.

Possiamo ottenere la distribuzione *a posteriori* di $N_{T+1}|H_T$:

$$\begin{aligned}
 Pr(N_{T+1} = n|H_T) &= \int_0^{+\infty} Pr(N_{T+1} = n|H_T, U = u) f_{U|H_T}(u) du \\
 &= \int_0^{+\infty} Pr(N_{T+1} = n|U = u) f_{U|H_T}(u) du \\
 &\quad \uparrow \qquad \uparrow \qquad \uparrow \\
 \text{indipendenza condizionata} \quad e^{-\lambda u} \frac{(\lambda u)^n}{n!} \quad & \frac{(\alpha + \lambda T)^{\alpha + \sum_{t=1}^T n_t}}{\Gamma(\alpha + \sum_{t=1}^T n_t)} u^{\alpha + \sum_{t=1}^T n_t - 1} e^{-(\alpha + \lambda T)u} \\
 &= \frac{\Gamma(\alpha + \sum_{t=1}^T n_t + n)}{\Gamma(\alpha + \sum_{t=1}^T n_t)n!} \left(\frac{\alpha + \lambda T}{\alpha + \lambda(T + 1)} \right)^{\alpha + \sum_{t=1}^T n_t} \left(\frac{\lambda}{\alpha + \lambda(T + 1)} \right)^n
 \end{aligned}$$

Si ha

$$N_{T+1} \sim BN \left(\alpha, \frac{\alpha}{\alpha + \lambda} \right) \quad N_{T+1}|H_T \sim BN \left(\alpha + \sum_{t=1}^T n_t, \frac{\alpha + \lambda T}{\alpha + \lambda(T + 1)} \right)$$

Modello Poisson-gamma, stima dei parametri

Per calcolare il premio *a posteriori* per un individuo della collettività si devono conoscere i valori dei parametri α, λ , legati alla valutazione *a priori*, e la storia di sinistrosità.

Nell'approccio bayesiano, se α, λ non sono noti, dovrebbero essere trattati come numeri aleatori, assegnando una distribuzione di probabilità e le ipotesi del modello dovrebbero essere intese come condizionate ai valori dei due parametri.

Nell'approccio bayesiano "empirico", i parametri α, λ sono considerati certi, ma non noti, e sono stimati dai dati.

Supponiamo di disporre dei dati per r assicurati della collettività, osservati ciascuno per T anni.

Per l'assicurato i -esimo, siano N_{it} il numero aleatorio di sinistri nell'anno t , U_i il parametro aleatorio di rischio, n_{it} valore osservato di N_{it} .

Per i processi dei diversi assicurati si assumono le ipotesi a), b).

Consideriamo la verosimiglianza dell'osservazione,

$$L(\lambda, \alpha; \mathbf{n}) = \prod_{i=1}^r Pr(N_{i1} = n_{i1}, \dots, N_{iT} = n_{iT}; \lambda, \alpha)$$

$$= \prod_{i=1}^r \frac{\Gamma(\alpha + \sum_{t=1}^T n_{it})}{\Gamma(\alpha) \prod_{t=1}^T n_{it}!} \left(\frac{\alpha}{\alpha + \lambda T} \right)^\alpha \left(\frac{\lambda}{\alpha + \lambda T} \right)^{\sum_{t=1}^T n_{it}}$$

Posto $n_i = \sum_{t=1}^T n_{it}$, $n = \sum_{i=1}^r n_i$ e considerata la log-verosimiglianza, si ha

$$l(\lambda, \alpha; \mathbf{n}) = \log L(\lambda, \alpha; \mathbf{n})$$

$$= \sum_{i=1}^r \left[\log \frac{\Gamma(\alpha + n_i)}{\Gamma(\alpha)} - \log(\prod_{t=1}^T n_{i,t} !) + \alpha(\log \alpha - \log(\alpha + \lambda T)) + n_i(\log \lambda - \log(\alpha + \lambda T)) \right]$$

$$\equiv \sum_{i=1}^r \left[\sum_{j=0}^{n_i-1} \log(\alpha + j) + \alpha(\log \alpha - \log(\alpha + \lambda T)) + n_i(\log \lambda - \log(\alpha + \lambda T)) \right]$$

Il sistema delle equazioni di log-verosimiglianza è

$$\begin{cases} \frac{\partial l(\lambda, \alpha; \mathbf{n})}{\partial \lambda} = 0 \\ \frac{\partial l(\lambda, \alpha; \mathbf{n})}{\partial \alpha} = 0 \end{cases}$$

Consideriamo, in particolare la prima equazione

$$\begin{aligned} \frac{\partial l(\lambda, \alpha; \mathbf{n})}{\partial \lambda} &= \sum_{i=1}^r \left[-\alpha \frac{1}{\alpha + \lambda T} T + n_i \left(\frac{1}{\lambda} - \frac{1}{\alpha + \lambda T} T \right) \right] \\ &= -\frac{\alpha T}{\alpha + \lambda T} r + \frac{n}{\lambda} - \frac{T}{\alpha + \lambda T} n = 0 \\ &\Updownarrow \\ \hat{\lambda} &= \frac{n}{rT} \end{aligned}$$

La stima di $\lambda = E(N_{it})$ è la frequenza sinistri.

Notiamo che il parametro α deve essere positivo: affinché la stima sia positiva, ci deve essere sufficiente “sovradispersione” nei dati.

Osservazione. Nel modello Poisson-gamma, si ottiene facilmente la distribuzione *a posteriori* per il parametro di rischio e dunque si determina facilmente il premio *a posteriori*.

Ciò dipende dal fatto che, in tale modello, le distribuzioni di U_i e di $N_{it}|U_i = u$, si “combinano” in modo tale che la distribuzione di $U_i|H_{iT}$ è della stessa famiglia della distribuzione di U_i , ma con parametri aggiornati: la famiglia gamma è **coniugata naturale** della Poisson.

In generale, ciò non accade e può essere molto complicato o impossibile determinare la distribuzione a posteriori, se non ricorrendo a tecniche di tipo simulativo.

Alcuni problemi dell’approccio bayesiano

- si devono specificare le distribuzioni di $\{Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots\}$, per ogni u , e la distribuzione di U_i ,
- può accadere che il premio bayesiano sia difficile da calcolare,
- il premio bayesiano può essere difficile da spiegare agli assicurati.

APPROCCIO DELLA TEORIA DELLA CREDIBILITÀ ALLA PERSONALIZZAZIONE BASATA SULL'ESPERIENZA

Consideriamo una collettività di rischi “analoghi”, omogenei rispetto a caratteristiche osservabili *a priori*, ed “indipendenti”. Nella collettività vi sia eterogeneità residua.

Sia P_{T+1} il premio *a priori* (o collettivo) per l’anno $T + 1$.

Supponiamo che l’assicurato i -esimo della classe sia stato osservato nei precedenti T anni.

Siano $y_{i1}, y_{i2}, \dots, y_{iT}$ i valori osservati delle variabili di interesse e $\bar{y}_{iT} = \frac{1}{T} \sum_{t=1}^T y_{it}$ la media: un riassunto della storia di sinistrosità dell’assicurato.

Il premio nell’anno $T + 1$ può essere posto pari a

- P_{T+1} , si tiene conto solo dei dati relativi alla collettività,
- \bar{y}_{iT} , si tiene conto solo dell’esperienza di sinistrosità del particolare assicurato,
- un compromesso tra le due precedenti posizioni estreme.

Una tipica **formula di credibilità** è

$$P_{i,T+1}^{cred}(H_{iT}) = (1 - z_{iT})P_{T+1} + z_{iT}\bar{y}_{iT}$$

dove

- $P_{i,T+1}^{cred}(H_{iT})$ indica il **premio di credibilità** per il dato assicurato, nell'anno $T + 1$, vista la storia di sinistrosità $H_{iT} = (Y_{i1} = y_{i1}, \dots, Y_{iT} = y_{iT})$,
- $0 \leq z_{iT} \leq 1$ è detto **fattore o coefficiente di credibilità**.

Il premio è una **media ponderata** di P_{T+1} , e \bar{y}_{iT} . Il coefficiente di credibilità assegna un “peso” alla sintesi della storia di sinistrosità dell'assicurato.

Si noti che fissando il premio uguale a

- P_{T+1} , il premio potrebbe essere troppo elevato per i rischi virtuosi,
- \bar{y}_{iT} , se $\bar{y}_{iT} = 0$ il premio è nullo; a causa dei pochi dati, \bar{y}_{iT} potrebbe essere una stima non affidabile del risarcimento atteso; non si introduce effetto di solidarietà.

Ragionevolmente, $0 < z_{i1} < z_{i2} < \dots \leq 1$: più è lungo il periodo di osservazione, maggiore è il peso assegnato alla sintesi dell'esperienza individuale. Se $z_{iT} = 1$, si parla di **piena credibilità**.

La valutazione di z_{iT} è determinante per applicare la formula.

Le formule di credibilità sono state introdotte in modo intuitivo e sono state proposte formule empiriche per ottenere z_{iT} . Tuttavia, le formule di credibilità possono essere giustificate in un solido quadro metodologico: la Teoria della Credibilità Bayesiana.

Osservazione. L'approccio della Teoria della Credibilità può anche essere utilizzato per valutare i premi (*a priori*) quando i dati dell'assicuratore per un dato portafoglio siano inadeguati (ad es. poco numerosi, per una copertura assicurativa emessa da poco). Può allora essere ragionevole, ai fini di ottenere la tariffa, combinare i dati del portafoglio con **dati collaterali** (ad es. dati di portafogli con rischi simili o dati di mercato). Il premio di credibilità è una media ponderata di un premio basato su dati collaterali e di un premio basato sui dati di portafoglio.

Credibilità bayesiana

Riprendiamo il quadro dell'approccio bayesiano.

Per un fissato assicurato, indicato con i , in una classe di rischi “analoghi”, omogenei rispetto a caratteristiche osservabili *a priori*, ed “indipendenti”, sia Y_{it} un numero aleatorio di interesse nell’anno t e U_i il parametro di rischio o componente di eterogeneità.

Introduciamo le seguenti ipotesi:

- i processi $\{U_i, Y_{i1}, Y_{i2}, \dots\}$, $i = 1, 2, \dots$, siano stocasticamente indipendenti, informazioni sulla storia di sinistrosità degli altri assicurati non portino a modificare la valutazione probabilistica del processo per l’assicurato i -esimo,
- fissato i , $Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots$ siano stocasticamente indipendenti per ogni u determinazione di U_i .

Al fine di ottenere uno stimatore per il premio individuale in $T + 1$,

$$\mu_{i,T+1}(U_i) = E(Y_{i,T+1}|U_i),$$

invece dello stimatore del premio bayesiano $E(Y_{i,T+1}|Y_{i1}, \dots, Y_{iT}) = E(\mu_{i,T+1}(U_i)|Y_{i1}, \dots, Y_{iT})$, consideriamo una funzione più semplice del processo di osservazione: un polinomio (non omogeneo) di primo grado (funzione lineare affine)

$$\alpha_{i0} + \alpha_{i1}Y_{i1} + \dots + \alpha_{iT}Y_{iT}.$$

Si devono determinare, con qualche criterio, i coefficienti. Come criterio scegliamo di minimizzare **la perdita quadratica attesa**

$$Q_i(\alpha_{i0}, \alpha_{i1}, \dots, \alpha_{iT}) = E \left[(\mu_{i,T+1}(U_i) - \alpha_{i0} - \sum_{t=1}^T \alpha_{it}Y_{it})^2 \right].$$

Come nell'approccio bayesiano, possiamo limitarci a tenere conto della storia di ciascun assicurato singolarmente e non dell'esperienze del collettivo, per l'ipotesi di indipendenza stocastica dei processi relativi ai diversi assicurati.

Per determinare la soluzione, semplifichiamo la notazione omettendo l'indice i . Per la funzione

$$\begin{aligned} Q(\alpha_0, \alpha_1, \dots, \alpha_T) &= E[(\mu_{T+1}(U) - \alpha_0 - \sum_{t=1}^T \alpha_t Y_t)^2] \\ &= \int (\mu_{T+1}(u) - \alpha_0 - \sum_{t=1}^T \alpha_t y_t)^2 dF_{U,Y_1,\dots,Y_T}(u, y_1, \dots, y_T), \end{aligned}$$

consideriamo le condizioni del primo ordine,

$$\begin{cases} \frac{\partial Q(\alpha_0, \alpha_1, \dots, \alpha_T)}{\partial \alpha_0} = 0 \\ \frac{\partial Q(\alpha_0, \alpha_1, \dots, \alpha_T)}{\partial \alpha_h} = 0, \quad h = 1, \dots, T \end{cases}$$

Supponiamo che siano soddisfatte le condizioni che consentono di effettuare la derivazione sotto il segno di integrale.

Si ha

$$\frac{\partial Q(\alpha_0, \alpha_1, \dots, \alpha_T)}{\partial \alpha_0} = \int 2(\mu_{T+1}(u) - \alpha_0 - \sum_{t=1}^T \alpha_t y_t)(-1) dF_{U, Y_1, \dots, Y_T}(u, y_1, \dots, y_T)$$

e

$$\frac{\partial Q(\alpha_0, \alpha_1, \dots, \alpha_T)}{\partial \alpha_0} = 0$$

\Leftrightarrow

$$E[2(\mu_{T+1}(U) - \alpha_0 - \sum_{t=1}^T \alpha_t Y_t)(-1)] = 0$$

\Leftrightarrow

$$E(\mu_{T+1}(U)) = \alpha_0 + \sum_{t=1}^T \alpha_t E(Y_t)$$

\Leftrightarrow

$$E(Y_{T+1}) = \alpha_0 + \sum_{t=1}^T \alpha_t E(Y_t) \quad (*)$$



L'ultima relazione segue dalla

$$E(\mu_{T+1}(U)) = E[E(Y_{T+1}|U)] = E(Y_{T+1}).$$

Per $h = 1, \dots, T$,

$$\frac{\partial Q(\alpha_0, \alpha_1, \dots, \alpha_T)}{\partial \alpha_h} = \int 2(\mu_{T+1}(u) - \alpha_0 - \sum_{t=1}^T \alpha_t y_t)(-y_h) dF_{U, Y_1, \dots, Y_T}(u, y_1, \dots, y_1)$$

$$\frac{\partial Q(\alpha_0, \alpha_1, \dots, \alpha_T)}{\partial \alpha_h} = 0$$

\Leftrightarrow

$$E[2(\mu_{T+1}(U) - \alpha_0 - \sum_{t=1}^T \alpha_t Y_t)(-Y_h)] = 0$$

\Leftrightarrow

$$E[\mu_{T+1}(U)Y_h] = \alpha_0 E(Y_h) + \sum_{t=1}^T \alpha_t E(Y_h Y_t) .$$

Si ha

$$\begin{aligned} E[\mu_{T+1}(U)Y_h] &= E[E(\mu_{T+1}(U)Y_h|U)] = E[\mu_{T+1}(U)E(Y_h|U)] \\ &= E[E(Y_{T+1}|U)E(Y_h|U)] = E[E(Y_{T+1}Y_h|U)] \\ &\quad \uparrow \text{indipendenza condizionata} \\ &= E(Y_{T+1}Y_h) \end{aligned}$$

Pertanto,

$$\frac{\partial Q(\alpha_0, \alpha_1, \dots, \alpha_T)}{\partial \alpha_h} = 0 \Leftrightarrow E(Y_{T+1}Y_h) = \alpha_0 E(Y_h) + \sum_{t=1}^T \alpha_t E(Y_h Y_t).$$

Il sistema delle condizioni del primo ordine

$$\begin{cases} E(Y_{T+1}) = \alpha_0 + \sum_{t=1}^T \alpha_t E(Y_t) \\ E(Y_h Y_{T+1}) = \alpha_0 E(Y_h) + \sum_{t=1}^T \alpha_t E(Y_h Y_t), \quad h = 1, \dots, T \end{cases}$$

\Leftrightarrow

$$\begin{cases} E(Y_{T+1}) = \alpha_0 + \sum_{t=1}^T \alpha_t E(Y_t) \\ cov(Y_h, Y_{T+1}) = \sum_{t=1}^T \alpha_t cov(Y_h, Y_t), \quad h = 1, \dots, T \end{cases}$$

Sistema lineare di $T + 1$ equazioni in $T + 1$ incognite, detto **sistema normale**. Ammette una e una sola soluzione se la matrice varianze-covarianze le vettore di componenti Y_h , $var(Y)$, è invertibile. Si prova che la soluzione è punto di minimo di Q .

Dove:
 $\alpha_0 = E[Y_{T+1}] - \sum_{t=1}^T \alpha_t E[Y_t]$

$$E[Y_{T+1} Y_h] = E[Y_{T+1}] E[Y_h] - \sum_{t=1}^T \alpha_t E[Y_t] E[Y_h] + \sum_{t=1}^T \alpha_t E[Y_t] E[Y_h]$$

$$Cov(Y_{T+1}, Y_h) = \sum_{t=1}^T \alpha_t Cov(Y_t, Y_h)$$

Risolvendo il sistema normale, si trova il punto di minimo di Q : $(\alpha_0^*, \alpha_1^*, \dots, \alpha_T^*)$.

Lo stimatore

$$\widetilde{\mu_{T+1}(U)}^{cred} = \alpha_0^* + \sum_{t=1}^T \alpha_t^* Y_t$$

è detto **stimatore di credibilità lineare** (non omogeneo) **del premio individuale in $T + 1$** .

Vista la storia di sinistrosità $H_T = (Y_1 = y_1, \dots, Y_T = y_T)$, **il premio di credibilità in $T + 1$** è

$$P_{T+1}^{cred}(H_T) = \alpha_0^* + \sum_{t=1}^T \alpha_t^* y_t.$$

Osservazioni.

- Poiché $(\alpha_0^*, \alpha_1^*, \dots, \alpha_T^*)$ è soluzione del sistema normale, dipende solo dai primi due momenti (che si suppone esistano finiti) dei numeri aleatori del processo di osservazione: rispetto all'approccio bayesiano, in questo approccio, non serve specificare completamente le distribuzioni.

$$\Rightarrow * \Leftrightarrow E[Y_{T+1}] = \alpha_0 + \sum_{t=1}^T \alpha_t Y_t$$

- Dalla (*)

$$E(\alpha_0^* + \sum_{t=1}^T \alpha_t^* Y_t) = E(Y_{T+1}) = E(\mu_{T+1}(U)).$$

- Se si considerano le seguenti due funzioni

$$Q_1(\alpha_0, \alpha_1, \dots, \alpha_T) = E[(E(Y_{T+1}|Y_1, \dots, Y_T) - \alpha_0 - \sum_{t=1}^T \alpha_t Y_t)^2],$$

$$Q_2(\alpha_0, \alpha_1, \dots, \alpha_T) = E[(Y_{T+1} - \alpha_0 - \sum_{t=1}^T \alpha_t Y_t)^2],$$

si trova che $(\alpha_0^*, \alpha_1^*, \dots, \alpha_T^*)$, soluzione del sistema normale, è anche il punto di minimo di Q_1 e di Q_2 : lo stimatore del premio di credibilità è anche il migliore stimatore, nel senso della funzione di perdita quadratica attesa, di $E(Y_{T+1}|Y_1, \dots, Y_T)$ e di Y_{T+1} .

Abbiamo considerato un fissato individuo, riprendendo la notazione con l'indice dell'individuo, il sistema normale è

$$\begin{cases} E(Y_{i,T+1}) = \alpha_0 + \sum_{t=1}^T \alpha_t E(Y_{it}) \\ cov(Y_{ih}, Y_{i,T+1}) = \sum_{t=1}^T \alpha_t cov(Y_{ih}, Y_{it}), \quad h = 1, \dots, T \end{cases}$$

La soluzione, in generale, dipende da i , $(\alpha_{i0}^*, \alpha_{i1}^*, \dots, \alpha_{iT}^*)$. Lo stimatore di credibilità del premio in $T + 1$ è

$$\widetilde{\mu_{i,T+1}(U_i)}^{cred} = \alpha_{i0}^* + \sum_{t=1}^T \alpha_{it}^* Y_{it}.$$

Se la condizione di analogia tra gli assicurati consente di accogliere l'ipotesi di uguale distribuzione dei processi $\{U_i, Y_{i1}, Y_{i2}, \dots\}$ al variare di i , o di uguaglianza dei primi due momenti, allora, a parità di T , la soluzione del sistema normale non dipende da i .

Modello di credibilità di Bühlmann

Si assume che

- $\{U_i, Y_{i1}, Y_{i2}, \dots\}$ processi iid al variare di i ,
- fissato i , $Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots$ stocasticamente indipendenti e identicamente distribuiti per ogni u determinazione di U_i , con momento primo e secondo finiti.

Poniamo

$$\mu(u) = E(Y_{it}|U_i = u), \quad v(u) = \text{var}(Y_{it}|U_i = u),$$

e

$$\mu = E[\mu(U_i)] = E[E(Y_{it}|U_i)] = E(Y_{it}),$$

$$v = E[v(U_i)] = E[\text{var}(Y_{it}|U_i)],$$

$$a = \text{var}[\mu(U_i)] = \text{var}[E(Y_{it}|U_i)].$$

Notiamo che $\mu(U_i)$ è il premio individuale nell'anno t , per ogni t .

Il parametro μ è il premio collettivo nell'anno t , per ogni t .

Gli altri due parametri intervengono nella varianza di Y_{it} ,

$$\text{var}(Y_{it}) = E[\text{var}(Y_{it}|U_i)] + \text{var}[E(Y_{it}|U_i)] = v + a.$$

In un portafoglio piuttosto omogeneo, c'è poca dispersione del premio individuale, $a = \text{var}[\mu(U_i)]$ è "piccola", viceversa se vi è notevole eterogeneità residua, a è "grande". La componente a nella scomposizione della varianza tiene conto della parte di "variabilità" dovuta all'eterogeneità residua.

I tre parametri μ, v, a , che non dipendono da i , sono detti **parametri di struttura del modello** e da tali parametri dipende il sistema normale. Si ha infatti

- $E(Y_{it}) = \mu,$
- $\text{var}(Y_{it}) = v + a,$
- per $h \neq t$, $\text{cov}(Y_{ih}, Y_{it}) = E(Y_{ih}Y_{it}) - E(Y_{ih})E(Y_{it})$

$$= E[E(Y_{ih}Y_{it}|U_i)] - E(Y_{it})^2$$

$$= E[E(Y_{ih}|U_i)E(Y_{it}|U_i)] - (E[E(Y_{it}|U_i)])^2$$

↑ per l'indipendenza condizionata

$$= E[\mu(U_i)^2] - (E[\mu(U_i)])^2 = \text{var}(\mu(U_i)) = a.$$

Il sistema normale

$$\begin{cases} E(Y_{i,T+1}) = \alpha_0 + \sum_{t=1}^T \alpha_t E(Y_{it}) \\ cov(Y_{ih}, Y_{i,T+1}) = \sum_{t=1}^T \alpha_t cov(Y_{ih}, Y_{it}), \quad h = 1, \dots, T \end{cases}$$

$$\begin{cases} \mu = \alpha_0 + \mu \sum_{t=1}^T \alpha_t \\ a = \alpha_h (\underbrace{v + a}_{\text{Cov}(Y_{ih}, Y_{i,T+1}) = \text{Var}(Y_{ih})} + \sum_{t \neq h}^T \alpha_t a), \quad h = 1, \dots, T \end{cases} \Leftrightarrow \begin{cases} \sum_{t=1}^T \alpha_t = 1 - \alpha_0 / \mu \\ a = \alpha_h v + a \sum_{t=1}^T \alpha_t, \quad h = 1, \dots, T \end{cases}$$

Posto $k = v/a$,

$$\begin{cases} \sum_{t=1}^T \alpha_t = 1 - \alpha_0 / \mu \\ 1 = \alpha_h k + \sum_{t=1}^T \alpha_t, \quad h = 1, \dots, T \end{cases} \Leftrightarrow \begin{cases} \sum_{t=1}^T \alpha_t = 1 - \alpha_0 / \mu \\ \alpha_h k = 1 - \sum_{t=1}^T \alpha_t, \quad h = 1, \dots, T \end{cases}$$

$$\begin{cases} \sum_{t=1}^T \alpha_t = 1 - \alpha_0 / \mu \\ \alpha_h = \frac{\alpha_0}{\mu k}, \quad h = 1, \dots, T \end{cases} \Leftrightarrow \begin{cases} T \frac{\alpha_0}{\mu k} = 1 - \frac{\alpha_0}{\mu} \\ \alpha_h = \frac{\alpha_0}{\mu k}, \quad h = 1, \dots, T \end{cases}$$

L'ultimo passaggio è ottenuto sommando membro a membro le equazioni del secondo gruppo e sostituendo nella prima equazione. La soluzione del sistema, che non dipende da i , è

$$\alpha_0^* = \mu \frac{k}{T+k}, \quad \alpha_t^* = \frac{1}{T+k}, \quad t = 1, \dots, T.$$

Lo stimatore di credibilità lineare del premio individuale in $T + 1$ è

$$\begin{aligned}\widetilde{\mu_{T+1}(U_i)^{cred}} &= \alpha_0^* + \sum_{t=1}^T \alpha_t^* Y_{it} = \mu \frac{k}{T+k} + \frac{1}{T+k} \sum_{t=1}^T Y_{it} \\ &= (1 - z_T)\mu + z_T \bar{Y}_{iT}\end{aligned}$$

dove

- $z_T = \frac{T}{T+k}$, con $k = \frac{v}{a}$,
- $\bar{Y}_{iT} = \frac{1}{T} \sum_{t=1}^T Y_{it}$, media dei valori della grandezza di interesse.

Vista la storia di sinistrosità $H_{iT} = (Y_{i1} = y_{i1}, \dots, Y_{iT} = y_{iT})$, **il premio di credibilità in $T + 1$ è**

$$P_{i,T+1}^{cred}(H_{iT}) = (1 - z_T)\mu + z_T \frac{\sum_{t=1}^T y_{it}}{T},$$

media ponderata del premio a priori (collettivo) μ e del valor medio osservato della grandezza di interesse $\frac{\sum_{t=1}^T y_{it}}{T}$ (es. numero medio annuo di sinistri, risarcimento medio annuo), con peso o coefficiente di credibilità z_T .

Il coefficiente di credibilità $z_T = \frac{T}{T+k}$, dipende da T, k

- fissato k , z_T è crescente al crescere di T e per $T \rightarrow +\infty$, $z_T \rightarrow 1$: piena credibilità,
- fissato T , z_T è decrescente al crescere di k .

Fissata $\text{var}(Y_{it}) = \nu + a$, se c'è "poca" eterogeneità residua, il premio individuale $E(Y_{it}|U_i) = \mu(U_i)$ ha poca variabilità, a è "piccolo" relativamente a ν e $k = \nu/a$ è "grande" e dunque z_T è "piccolo"; viceversa, se c'è "molta" eterogeneità residua z_T è "grande".

In linea con l'intuizione, in una collettività omogenea il premio collettivo è una stima adeguata del premio individuale; ciò non accade in una collettività eterogenea, in tale caso può essere più utile tenere conto dell'osservazione della storia individuale di sinistrosità per ottenere una migliore valutazione del premio.

Nota. I coefficienti $\alpha_t^* = \frac{1}{T+k}$ dello stimatore del premio di credibilità, per $t = 1, \dots, T$, non dipendono da t . Segue dalla scambiabilità del processo $\{Y_{i1}, Y_{i2}, \dots\}$, in forza della quale

$$E \left[(Y_{i,T+1} - \alpha_{i0} - \sum_{t=1}^T \alpha_{it} Y_{it})^2 \right] = E \left[(Y_{i,T+1} - \alpha_{i0} - \sum_{t=1}^T \alpha_{it} Y_{i\pi(t)})^2 \right],$$

con $(\pi(1), \dots, \pi(T))$ una qualunque permutazione di $(1, \dots, T)$.

Modello di Bühlmann, stima dei parametri

Il premio di credibilità dipende dai parametri di struttura. Nella credibilità bayesiana empirica, tali parametri sono stimati dai dati, tramite stimatori non distorti (stime non parametriche).

Supponiamo di disporre dei dati per r assicurati della collettività, osservati ciascuno per T anni. Per l'assicurato i -esimo, siano Y_{it} la grandezza aleatoria di interesse nell'anno t , U_i il parametro aleatorio di rischio, y_{it} valore osservato di Y_{it} .

Per i processi dei diversi assicurati si assumono le ipotesi del modello, con parametri di struttura μ, ν, a .

Stimatore per μ . Consideriamo le medie campionarie per rischio

$$\bar{Y}_{iT} = \frac{1}{T} \sum_{t=1}^T Y_{it} = \tilde{\mu}_i, \quad i = 1, \dots, r,$$

e la loro media aritmetica

$$\tilde{\mu} = \frac{1}{r} \sum_{i=1}^r \tilde{\mu}_i.$$

$\tilde{\mu}$ è stimatore non distorto per μ .

La stima di μ è

$$\hat{\mu} = \frac{1}{r} \sum_{i=1}^r \hat{\mu}_i = \frac{1}{r} \sum_{i=1}^r \left(\frac{1}{T} \sum_{t=1}^T y_{it} \right).$$

Stimatore per ν . Consideriamo le varianze campionarie per rischio

$$\tilde{\sigma}_i^2 = \frac{1}{T-1} \sum_{t=1}^T (Y_{it} - \tilde{\mu}_i)^2, \quad i = 1, \dots, r,$$

e la loro media aritmetica

$$\tilde{\nu} = \frac{1}{r} \sum_{i=1}^r \tilde{\sigma}_i^2.$$

$\tilde{\nu}$ è stimatore non distorto per ν . Infatti, poiché $Y_{i1}|U_i = u, \dots, Y_{iT}|U_i = u$ sono iid con comune varianza $\text{var}(Y_{it}|U_i = u) = \nu(u)$, si ha

$$E \left[\frac{1}{T-1} \sum_{t=1}^T (Y_{it} - \tilde{\mu}_i)^2 | U_i = u \right] = E(\tilde{\sigma}_i^2 | U_i = u) = \nu(u), \text{ per ogni } u.$$

Segue

$$E(\tilde{\sigma}_i^2) = E[E(\tilde{\sigma}_i^2 | U_i)] = E(\nu(U_i)) = \nu$$

e

$$E(\tilde{\nu}) = E \left(\frac{1}{r} \sum_{i=1}^r \tilde{\sigma}_i^2 \right) = \nu.$$

La stima di ν è

$$\hat{\nu} = \frac{1}{r} \sum_{i=1}^r \hat{\sigma}_i^2 = \frac{1}{r} \sum_{i=1}^r \left(\frac{1}{T-1} \sum_{t=1}^T (y_{it} - \hat{\mu}_i)^2 \right).$$

Stimatore per a . Consideriamo la varianza campionaria delle medie per rischio $\tilde{\mu}_i = \frac{1}{T} \sum_{t=1}^T Y_{it}$

$$\frac{1}{r-1} \sum_{i=1}^r (\tilde{\mu}_i - \tilde{\mu})^2$$

e calcoliamo il valore atteso. I numeri aleatori $\tilde{\mu}_1, \dots, \tilde{\mu}_r$ sono iid. Infatti,

- $\tilde{\mu}_i$ è media aritmetica di Y_{i1}, \dots, Y_{iT} ,
- $(Y_{11}, \dots, Y_{1T}), \dots, (Y_{r1}, \dots, Y_{rT})$ sono iid.

$\Rightarrow \frac{1}{r-1} \sum_{i=1}^r (\tilde{\mu}_i - \tilde{\mu})^2$ è stimatore non distorto della comune varianza dei $\tilde{\mu}_i$

Si ha

$$\begin{aligned} \text{var}(\tilde{\mu}_i) &= E[\text{var}(\tilde{\mu}_i | U_i)] + \text{var}[E(\tilde{\mu}_i | U_i)] \\ &= E\left[\frac{1}{T^2} \sum_{t=1}^T \text{var}(Y_{it} | U_i)\right] + \text{var}\left[\frac{1}{T} \sum_{t=1}^T E(Y_{it} | U_i)\right] \\ &\quad \begin{matrix} \uparrow & \uparrow & \uparrow \\ \text{indipendenza condizionata} & v(U_i) & \mu(U_i) \end{matrix} \\ &= \frac{1}{T} E[v(U_i)] + \text{var}[\mu(U_i)] \\ &= \frac{1}{T} v + a \end{aligned}$$

Pertanto,

$$E \left(\frac{1}{r-1} \sum_{i=1}^r (\tilde{\mu}_i - \tilde{\mu})^2 \right) = \frac{1}{T} v + a.$$

Allora uno **stimatore** non distorto per a è

$$\tilde{a} = \frac{1}{r-1} \sum_{i=1}^r (\tilde{\mu}_i - \tilde{\mu})^2 - \frac{1}{T} \tilde{v}.$$

La **stima** di a è

$$\hat{a} = \frac{1}{r-1} \sum_{i=1}^r (\hat{\mu}_i - \hat{\mu})^2 - \frac{1}{T} \hat{v}.$$

Notiamo che, in una collettività non omogenea, il parametro a è positivo. Tale dovrà risultare la stima.

Esempio. $Y_{it} = N_{it}$ numero annuo di sinistri, $r = 20, T = 10$.

$i t$	1	2	3	4	5	6	7	8	9	10	$\hat{\mu}_i$	$\hat{\sigma}_i^2$	Premio
1	1	0	1	0	0	0	1	0	0	0	0.30	0.233	0.222
2	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.052
3	1	0	0	0	0	0	0	0	0	0	0.10	0.100	0.109
4	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.052
5	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.052
6	0	0	0	0	0	1	0	0	0	0	0.10	0.100	0.109
7	0	1	0	0	0	0	0	0	0	0	0.10	0.100	0.109
8	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.052
9	0	1	0	0	0	0	0	0	0	1	0.20	0.178	0.165
10	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.052
11	1	1	0	0	1	0	0	0	1	0	0.40	0.267	0.278
12	0	0	0	0	0	0	0	1	0	1	0.20	0.178	0.165
13	0	0	0	0	0	0	0	0	0	1	0.10	0.100	0.109
14	0	1	0	0	0	0	1	0	0	0	0.20	0.178	0.165
15	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.052
16	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.052
17	0	0	0	0	0	0	1	0	0	1	0.20	0.178	0.165
18	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.052
19	1	0	1	0	0	1	0	1	0	1	0.50	0.278	0.334
20	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.052

$$\hat{\mu}_i = \frac{1}{T} \sum_{t=1}^T n_{it}, \quad \hat{\mu} = \frac{1}{r} \sum_{i=1}^r \hat{\mu}_i = 0.12, \quad \hat{\sigma}_i^2 = \frac{1}{T-1} \sum_{t=1}^T (n_{it} - \hat{\mu}_i)^2, \quad \hat{v} = \frac{1}{r} \sum_{i=1}^r \hat{\sigma}_i^2 = 0.094444,$$

$$\hat{a} = \frac{1}{r-1} \sum_{i=1}^r (\hat{\mu}_i - \hat{\mu})^2 - \frac{1}{T} \hat{v} = 0.012240, \quad \hat{k} = \hat{v}/\hat{a} = 7.716197,$$

$$z_T = \frac{T}{T+k} = 0.564455, \quad 1 - z_T = 0.435545, \quad P_{i,T+1}^{cred}(H_{iT}) = (1 - z_T)\hat{\mu} + z_T\hat{\mu}_i$$

Esempio. Modello Poisson-mistura semiparametrico per i numeri di sinistri

Per l'assicurato i -esimo, siano N_{it} il numero di sinistri nell'anno t e U_i il parametro aleatorio di rischio. Assumiamo

- $\{U_i, N_{i1}, N_{i2}, \dots\}$ processi iid al variare di i ,
- fissato i ,

$N_{i1}|U_i = u, N_{i2}|U_i = u, \dots$ stocasticamente indipendenti per ogni u determinazione di U_i ,

$N_{it}|U_i = u \sim P(\lambda u)$, per ogni t , $\lambda > 0$,

U_i tale che $E(U_i) = 1$, $var(U_i) = \sigma^2$.

Sono soddisfatte le ipotesi del modello di Bühlmann, in aggiunta è specificata la distribuzione condizionata. Si ha

$$\mu(u) = E(N_{it}|U_i = u) = \lambda u, \quad v(u) = var(N_{it}|U_i = u) = \lambda u.$$

I parametri di struttura sono

$$\mu = E(N_{it}) = E[E(N_{it}|U_i)] = E(\lambda U_i) = \lambda E(U_i) = \lambda,$$

$$v = E[var(N_{it}|U_i)] = E(\lambda U_i) = \lambda,$$

$$a = var[E(N_{it}|U_i)] = var(\lambda U_i) = \lambda^2 var(U_i) = \lambda^2 \sigma^2.$$

Si può ottenere lo stimatore di credibilità del premio per la frequenza in $T + 1$,

$$\widetilde{\mu_{T+1}(U_i)}^{cred} = (1 - z_T)\mu + z_T \frac{1}{T} \sum_{t=1}^T N_{it},$$

dove $z_T = \frac{T}{T+k}$, con $k = \frac{v}{a}$.

Modello di credibilità di Bühlmann-Straub

Nel modello di Bühlmann i numeri aleatori del processo di interesse sono giudicati identicamente distribuiti. In molti casi tale ipotesi non è adeguata. Ad esempio,

- M_{it} numero di sinistri per la polizza i , nell'anno t , con tempo di esposizione esp_{it} ,
- X_{it} risarcimento totale per la polizza i , nell'anno t , con tempo di esposizione esp_{it} oppure con esposizione monetaria w_{it} ,
- X_{it} risarcimento totale per il portafoglio i , nell'anno t , con montepremi P_{it} .

Se, negli esempi, poniamo

$$Y_{it} = M_{it}/esp_{it}, \quad Y_{it} = X_{it}/esp_{it}, \quad Y_{it} = X_{it}/w_{it}, \quad Y_{it} = X_{it}/P_{it},$$

esprimiamo le tre grandezze per unità di “esposizione”. Per tali numeri aleatori si può pensare di accogliere un’ipotesi di uguale valore atteso, al variare di t , non di uguale distribuzione: è piuttosto naturale valutare di maggiore varianza una grandezza con minore “esposizione”.

Il modello di Bühlmann-Straub consente di tenere conto di tali aspetti.

Si assume

- $\{U_i, Y_{i1}, Y_{i2}, \dots\}$ processi stocasticamente indipendenti al variare di i ,
- U_1, U_2, \dots identicamente distribuiti, *Non più anche Y_{i1}, Y_{i2}, \dots*
- fissato i , $Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots$ stocasticamente indipendenti per ogni u determinazione di U_i , con momento primo e secondo finiti tali che

$$E(Y_{it}|U_i = u) = \mu(u), \quad var(Y_{it}|U_i = u) = \frac{\nu(u)}{m_{it}},$$

$\mu(\cdot), \nu(\cdot)$ non dipendono da i, t , m_{it} è una misura di “volume”, nota, che dipende da i, t .

Poniamo

$$\begin{aligned} \mu &= E[\mu(U_i)] = E[E(Y_{it}|U_i)] = E(Y_{it}), \\ \nu &= E[\nu(U_i)] = E[m_{it} var(Y_{it}|U_i)] = m_{it} E[var(Y_{it}|U_i)], \\ a &= var[\mu(U_i)] = var[E(Y_{it}|U_i)]. \end{aligned}$$

Analogamente al modello di Bühlmann, $\mu(U_i)$ è il premio individuale nell'anno t , per ogni t . Il parametro μ è il premio collettivo nell'anno t , per ogni t , a è una “misura” di eterogeneità, ν, a intervengono nella varianza di Y_{it} ,

$$var(Y_{it}) = E[var(Y_{it}|U_i)] + var[E(Y_{it}|U_i)] = \frac{\nu}{m_{it}} + a.$$

I tre parametri μ, v, a , che non dipendono da i , sono i **parametri di struttura del modello** e da tali parametri dipende il sistema normale. Si ha infatti

- $E(Y_{it}) = \mu$,
- $var(Y_{it}) = \frac{v}{m_{it}} + a$,
- per $h \neq t$, $cov(Y_{ih}, Y_{it}) = E(Y_{ih}Y_{it}) - E(Y_{ih})E(Y_{it})$

$$= E[E(Y_{ih}Y_{it}|U_i)] - E(Y_{it})^2$$

$$= E[E(Y_{ih}|U_i)E(Y_{it}|U_i)] - (E[E(Y_{it}|U_i)])^2$$

$$\quad \uparrow \text{per l'indipendenza condizionata}$$

$$= E[\mu(U_i)^2] - (E[\mu(U_i)])^2 = var(\mu(U_i)) = a$$

Il sistema normale

$$\begin{cases} E(Y_{i,T+1}) = \alpha_0 + \sum_{t=1}^T \alpha_t E(Y_{it}) \\ cov(Y_{ih}, Y_{i,T+1}) = \sum_{t=1}^T \alpha_t cov(Y_{ih}, Y_{it}), \quad h = 1, \dots, T \end{cases}$$

$$\begin{cases} \mu = \alpha_0 + \mu \sum_{t=1}^T \alpha_t \\ a = \alpha_h \left(\frac{v}{m_{ih}} + a \right) + \sum_{\substack{t=1 \\ t \neq h}}^T \alpha_t a, \quad h = 1, \dots, T \end{cases} \Leftrightarrow \begin{cases} \sum_{t=1}^T \alpha_t = 1 - \alpha_0 / \mu \\ a = \alpha_h \frac{v}{m_{ih}} + a \sum_{t=1}^T \alpha_t, \quad h = 1, \dots, T \end{cases}$$

Posto $k = v/a$ e $m_{i\bullet} = \sum_{t=1}^T m_{it}$,

$$\begin{cases} \sum_{t=1}^T \alpha_t = 1 - \alpha_0 / \mu \\ 1 = \alpha_h k \frac{1}{m_{ih}} + \sum_{t=1}^T \alpha_t, \quad h = 1, \dots, T \end{cases} \Leftrightarrow \begin{cases} \sum_{t=1}^T \alpha_t = 1 - \alpha_0 / \mu \\ \alpha_h = \frac{m_{ih}}{k} (1 - \sum_{t=1}^T \alpha_t), \quad h = 1, \dots, T \end{cases}$$

$$\begin{cases} \sum_{t=1}^T \alpha_t = 1 - \alpha_0 / \mu \\ \alpha_h = \frac{m_{ih}}{k} \frac{\alpha_0}{\mu}, \quad h = 1, \dots, T \end{cases} \Leftrightarrow \begin{cases} m_{i\bullet} \frac{\alpha_0}{\mu k} = 1 - \frac{\alpha_0}{\mu} \\ \alpha_h = \frac{m_{ih}}{k} \frac{\alpha_0}{\mu}, \quad h = 1, \dots, T \end{cases}$$

L'ultimo passaggio è ottenuto sommando membro a membro le equazioni del secondo gruppo e sostituendo nella prima equazione. La soluzione del sistema, che dipende da i , è

$$\alpha_{i0}^* = \mu \frac{k}{m_{i\bullet} + k}, \quad \alpha_{it}^* = \frac{m_{it}}{m_{i\bullet} + k}, \quad t = 1, \dots, T.$$

Lo stimatore di credibilità lineare del premio individuale in $T + 1$ è

$$\begin{aligned}\widetilde{\mu_{i,T+1}(U_i)}^{cred} &= \alpha_{i0}^* + \sum_{t=1}^T \alpha_{it}^* Y_{it} = \mu \frac{k}{m_{i\bullet} + k} + \sum_{t=1}^T \frac{m_{it}}{m_{i\bullet} + k} Y_{it} \\ &= (1 - z_{iT})\mu + z_{iT} \bar{\bar{Y}}_{iT}\end{aligned}$$

dove

- $z_{iT} = \frac{m_{i\bullet}}{m_{i\bullet} + k}$, con $k = \frac{v}{a}$,
- $\bar{\bar{Y}}_{iT} = \sum_{t=1}^T \frac{m_{it}}{m_{i\bullet}} Y_{it}$, media ponderata dei valori della grandezza di interesse con pesi le esposizioni relative.

Vista la storia di sinistrosità $H_{iT} = (Y_{i1} = y_{i1}, \dots, Y_{iT} = y_{iT})$, il **premio di credibilità in $T + 1$** è

$$P_{i,T+1}^{cred}(H_{iT}) = (1 - z_{iT})\mu + z_{iT} \sum_{t=1}^T \frac{m_{it}}{m_{i\bullet}} y_{it},$$

media ponderata del premio a priori (collettivo) μ e della media osservata della grandezza di interesse $\sum_{t=1}^T \frac{m_{it}}{m_{i\bullet}} y_{it}$ con fattore, peso di credibilità z_{iT} .

Il peso di credibilità $z_{iT} = \frac{m_{i\bullet}}{m_{i\bullet} + k}$, dipende da $m_{i\bullet}, k$

- fissato k , z_{iT} è crescente al crescere di $m_{i\bullet}$, che dipende da T, m_{i1}, \dots, m_{iT} ,
- fissato $m_{i\bullet}$, z_{iT} è decrescente al crescere di k , se c'è “poca” eterogeneità residua, il premio individuale $E(Y_{it}|U_i) = \mu(U_i)$ ha poca variabilità, se a è “piccolo” relativamente a ν , $k = \nu/a$ è “grande” e dunque z_{iT} è “piccolo”: non si attribuisce molto peso alla storia individuale; viceversa, se c'è “molta” eterogeneità residua z_{iT} è “grande”: si attribuisce peso maggiore alla sintesi della storia individuale rispetto al premio collettivo.

Modello di Bühlmann-Straub, stima dei parametri

Il premio di credibilità dipende dai parametri di struttura. Tali parametri sono stimati dai dati, tramite stimatori non distorti (stime non parametriche).

Supponiamo di disporre dei dati per r assicurati della collettività, osservati ciascuno per T anni. Per l'assicurato i -esimo, siano Y_{it} la grandezza aleatoria di interesse nell'anno t , U_i il parametro aleatorio di rischio, m_{it} la misura di esposizione e y_{it} valore osservato di Y_{it} .

Per i processi dei diversi assicurati si assumono le ipotesi del modello, con parametri di struttura μ, ν, a .

Spesso sono utilizzati i seguenti stimatori non distorti.

Stimatore per μ . Consideriamo le medie ponderate per rischio

$$\bar{\bar{Y}}_{iT} = \sum_{t=1}^T \frac{m_{it}}{m_{i\bullet}} Y_{it} = \tilde{\mu}_i, \quad i = 1, \dots, r,$$

e la loro media ponderata

$$\tilde{\mu} = \sum_{i=1}^r \frac{m_{i\bullet}}{m_{\bullet\bullet}} \tilde{\mu}_i, \text{ dove } m_{\bullet\bullet} = \sum_{i=1}^r m_{i\bullet}.$$

Stimatore per ν . Consideriamo le seguenti varianze campionarie per rischio

$$\tilde{\sigma}_i^2 = \frac{1}{T-1} \sum_{t=1}^T m_{it} (Y_{it} - \tilde{\mu}_i)^2, \quad i = 1, \dots, r,$$

e la loro media aritmetica

$$\tilde{\nu} = \frac{1}{r} \sum_{i=1}^r \tilde{\sigma}_i^2.$$

Stimatore per a .

$$\tilde{a} = \frac{m_{\bullet\bullet}}{m_{\bullet\bullet}^2 + \sum_{i=1}^r m_{i\bullet}^2} [\sum_{i=1}^r m_{i\bullet} (\tilde{\mu}_i - \tilde{\mu})^2 - (r-1)\tilde{\nu}].$$

Notiamo che, in una collettività non omogenea, il parametro a è positivo. Tale dovrà risultare la stima.

Esempio Bühlmann-Straub. Due gruppi di assicurati. Obiettivo: valutare i risarcimenti attesi

(premi equi) per i due gruppi, nell'anno 4.

		Anno			
		1	2	3	4
Gruppo 1	Risarcimento totale	15 000	22 400	25 200	
	Esposizione	50	70	80	85
	Risarcimento totale per unità di esposizione (quota danni)	300	320	315	
Gruppo 2	Risarcimento totale	46 500	48 000	44 950	
	Esposizione	150	160	155	110
	Risarcimento totale per unità di esposizione (quota danni)	310	300	290	

200
665

Per il gruppo i -esimo nell'anno t , siano

- X_{it} il risarcimento totale
- m_{it} l'esposizione,
- $Y_{it} = X_{it}/m_{it}$ il risarcimento medio per unità di esposizione

Assumendo il modello di Bühlmann-Straub per i risarcimenti per unità di esposizione, si ha

$$m_{1\bullet} = 200, \quad m_{2\bullet} = 465, \quad m_{\bullet\bullet} = 665$$

$$\hat{\mu}_1 = 313.00, \quad \hat{\mu}_2 = 299.89, \quad \hat{\sigma}_1^2 = 6\ 100.00, \quad \hat{\sigma}_2^2 = 15\ 247.31$$

Parametri di struttura

$$\hat{\mu} = 303.83, \quad \hat{v} = 10\ 673.66, \quad \hat{a} = 47.74, \quad \hat{k} = \hat{v}/\hat{a} = 223.57$$

Coefficienti di credibilità

$$z_{1T} = 0.472, \quad 1 - z_{1T} = 0.528 \\ z_{2T} = 0.675, \quad 1 - z_{2T} = 0.325$$

Premi di credibilità nell'anno 4, per unità di esposizione

$$P_{1,T+1}^{cred}(H_{1T}) = 0.528 \times 303.83 + 0.472 \times 313.00 = 308.16 \\ P_{2,T+1}^{cred}(H_{2T}) = 0.325 \times 303.83 + 0.675 \times 299.89 = 301.17$$

Premi di credibilità nell'anno 4

$$308.16 \times 85 = 26\ 193.80 \quad (303.83 \times 85 = 25\ 825.94)$$

$$301.17 \times 110 = 33\ 128.97 \quad (303.83 \times 110 = 33\ 421.80)$$

Esempio. Modello Poisson-mistura semiparametrico per i numeri di sinistri, con personalizzazione a priori

- Aggiunte confronto al precedente

Per l'assicurato i -esimo, siano N_{it} il numero di sinistri nell'anno t e U_i il parametro aleatorio di rischio. Assumiamo

- $\{U_i, N_{i1}, N_{i2}, \dots\}$ processi stocasticamente indipendenti al variare di i ,
- U_1, U_2, \dots identicamente distribuiti,
- fissato i ,

$N_{i1}|U_i = u, N_{i2}|U_i = u, \dots$ stocasticamente indipendenti per ogni u determinazione di U_i ,

$N_{it}|U_i = u \sim P(\lambda_{it}u)$, per ogni t , $\lambda_{it} > 0$,
→ classe tariffaria dell'ass. i nell'anno t
 U_i tale che $E(U_i) = 1$, $var(U_i) = \sigma^2$.

Si ha

$$E(N_{it}) = E[E(N_{it}|U_i)] = E(\lambda_{it}U_i) = \lambda_{it}.$$

Il premio (per la frequenza) *a priori* nell'anno t per l'assicurato i -esimo, dipende da i, t . Se il portafoglio di rischi è ripartito in classi tariffarie, λ_{it} può rappresentare il premio *a priori* valutato per la classe tariffaria alla quale l'assicurato appartiene nell'anno t .

Posto

$$Y_{it} = \frac{N_{it}}{\lambda_{it}},$$

per tali numeri aleatori si ha

- $\{U_i, Y_{i1}, Y_{i2}, \dots\}$ processi stocasticamente indipendenti al variare di i ,
- U_1, U_2, \dots identicamente distribuiti, con $E(U_i) = 1$, $var(U_i) = \sigma^2$,
- fissato i ,

$Y_{i1}|U_i = u, Y_{i2}|U_i = u, \dots$ stocasticamente indipendenti per ogni u determinazione di U_i , con

$$E(Y_{it}|U_i = u) = \frac{1}{\lambda_{it}} E(N_{it}|U_i = u) = u = \mu(u), \quad \mu(\cdot) \text{ non dipende da } i, t,$$

$$var(Y_{it}|U_i = u) = \frac{1}{\lambda_{it}^2} var(N_{it}|U_i = u) = \frac{1}{\lambda_{it}^2} u = \frac{1}{\lambda_{it}^2} v(u), \quad v(\cdot) \text{ non dipende da } i, t,$$

Sono soddisfatte le ipotesi del modello di Bühlmann-Straub con: $\mu(\cdot)$, $v(\cdot)$ funzione identica,
 $m_{it} = \lambda_{it}$.

Lotto

I parametri di struttura sono

$$\begin{aligned}\mu &= E[\mu(U_i)] = E(U_i) = 1, \\ \nu &= E[\nu(U_i)] = E(U_i) = 1, \\ a &= \text{var}[\mu(U_i)] = \text{var}(U_i) = \sigma^2.\end{aligned}$$

Si può ottenere lo stimatore di credibilità per $Y_{i,T+1}$

$$\widetilde{\mu_{i,T+1}}(U_i)^{\text{cred}} = E(\widetilde{Y_{i,T+1}}|U_i)^{\text{cred}} = (1 - z_{iT})\mu + z_{iT}\bar{\bar{Y}}_{iT} = \frac{1 + \sigma^2 \sum_{t=1}^T N_{it}}{1 + \sigma^2 \sum_{t=1}^T \lambda_{it}},$$

dove

$$z_{iT} = \frac{\lambda_{i\bullet}}{\lambda_{i\bullet} + 1/\sigma^2}, \quad \text{con } \lambda_{i\bullet} = \sum_{t=1}^T \lambda_{it}, \quad \text{e} \quad \bar{\bar{Y}}_{iT} = \sum_{t=1}^T \frac{\lambda_{it}}{\lambda_{i\bullet}} Y_{it}.$$

Lo stimatore di credibilità del premio per la frequenza in $T + 1$

$$E(\widetilde{N_{i,T+1}}|U_i)^{\text{cred}} = \lambda_{i,T+1} \frac{1 + \sigma^2 \sum_{t=1}^T N_{it}}{1 + \sigma^2 \sum_{t=1}^T \lambda_{it}}.$$

CREDIBILITÀ ESATTA (CENNO)

Solo da leggere, non da studiare

Nel modello Poisson-gamma, per un fissato assicurato, lo **stimatore bayesiano** del premio individuale in $T + 1$ è

$$E(N_{T+1}|N_1, \dots, N_T) = \lambda \frac{\alpha}{\alpha + \lambda T} + \frac{\sum_{t=1}^T N_t}{T} \frac{\lambda T}{\alpha + \lambda T} = \lambda(1 - z_T) + \frac{\sum_{t=1}^T N_t}{T} z_T,$$

che è una formula di credibilità. Questo succede anche in altri modelli.

Si parla di **credibilità esatta** quando in un modello bayesiano lo stimatore bayesiano del premio individuale coincide con lo stimatore di credibilità.

Riprendiamo lo schema bayesiano, per un fissato assicurato (omettiamo l'indice i),

- $Y_1|U = u, Y_2|U = u, \dots$ stocasticamente indipendenti, di legge assegnata, per ogni u determinazione possibile di U ,
- U di legge assegnata.

Solo leggere

Può accadere che lo stimatore bayesiano del premio individuale sia una funzione lineare (affine) del vettore delle osservazioni ovvero che

$$E(\mu_{T+1}(U)|Y_1, \dots, Y_T) = E(Y_{T+1}|Y_1, \dots, Y_T) = \alpha_0 + \sum_{t=1}^T \alpha_t Y_t.$$

Allora, se si calcola lo stimatore di credibilità si ottiene lo stesso stimatore. Infatti lo stimatore bayesiano è soluzione di

$$\operatorname{argmin}_{g \in \mathcal{G}} E[(\mu_{T+1}(U) - g(Y_1, \dots, Y_T))^2],$$

lo stimatore di credibilità è soluzione di

$$\operatorname{argmin}_{g \in \mathcal{L}} E[(\mu_{T+1}(U) - g(Y_1, \dots, Y_T))^2],$$

dove $\mathcal{L} \subset \mathcal{G}$, $\mathcal{L} = \{g : g(y_1, \dots, y_T) = \alpha_0 + \sum_{t=1}^T \alpha_t y_t\}$.

Tale situazione si verifica in riferimento alle distribuzioni delle famiglie esponenziali lineari e delle loro coniugate naturali.

Si assuma che

- $Y_t|U = u \sim f_{Y_t|U=u} \in \mathcal{F}$ **famiglia esponenziale lineare**, con

$$f_{Y_t|U=u}(y) = \exp \left\{ \omega_t \frac{yu - b(u)}{\phi} \right\} c(y, \phi, \omega_t), \quad y \in \mathcal{Y},$$

dove

$b(\cdot)$ è la funzione cumulante,

u è il parametro canonico,

ϕ è il parametro di dispersione,

$\omega_t > 0$ peso.

Si noti che le distribuzioni, al variare di t , possono differire solo per i pesi.

- $U \sim f_U \in \mathcal{U}$, con

$$f_U(u) = \exp \left\{ \frac{\mu u - b(u)}{\delta} \right\} d(\mu, \delta), \quad u \in]u_0, u_1[,$$

dove

μ, δ sono parametri,

$d(\mu, \delta)$ è una costante di normalizzazione.

Consideriamo la densità *a posteriori* del parametro aleatorio U , condizionata a $H_T = (Y_1 = y_1, \dots, Y_T = y_T)$, dal Teorema di Bayes per densità

$$\begin{aligned}
f_{U|H_T}(u) &\propto f_U(u) \prod_{t=1}^T f_{Y_t|U=u}(y_t) \\
&\quad \uparrow \qquad \uparrow \\
&\text{densità *a priori*} \qquad \text{verosimiglianza dell'osservazione} \\
&\propto e^{\frac{\mu u - b(u)}{\delta}} \prod_{t=1}^T e^{\omega_t \frac{y_t u - b(u)}{\phi}} = e^{\frac{\mu u - b(u)}{\delta}} e^{\sum_{t=1}^T \omega_t \frac{y_t u - b(u)}{\phi}} \\
&= \exp \left\{ \frac{\mu u}{\delta} - \frac{b(u)}{\delta} + \sum_{t=1}^T \frac{\omega_t y_t u}{\phi} - \frac{b(u)}{\phi} \sum_{t=1}^T \omega_t \right\} \\
&= e^{\left(\frac{\mu}{\delta} + \omega_\bullet \frac{\bar{y}_T}{\phi} \right) u} e^{-b(u) \left(\frac{1}{\delta} + \frac{\omega_\bullet}{\phi} \right)},
\end{aligned}$$

dove $\omega_\bullet = \sum_{t=1}^T \omega_t$, $\bar{y}_T = \sum_{t=1}^T \frac{\omega_t}{\omega_\bullet} y_t$.

Con semplici passaggi si ottiene

$$f_{U|H_T}(u) \propto \exp \left\{ \frac{\mu^* u - b(u)}{\delta^*} \right\},$$

dove $\mu^* = \left(\frac{\phi}{\delta} \mu + \omega_\bullet \bar{y}_T \right) \left(\frac{\phi}{\delta} + \omega_\bullet \right)^{-1}$, $\delta^* = \phi \left(\frac{\phi}{\delta} + \omega_\bullet \right)^{-1}$.

Pertanto,

$$f_U(u) = \exp\left\{\frac{\mu u - b(u)}{\delta}\right\} d(\mu, \delta) \quad \rightarrow \quad f_{U|H_T}(u) = \exp\left\{\frac{\mu^* u - b(u)}{\delta^*}\right\} d(\mu^*, \delta^*)$$

La distribuzione *a posteriori* di $U|H_T$ appartiene ancora alla famiglia \mathcal{U} , ma con parametri aggiornati

$$\mu \rightarrow \mu^*, \quad \delta \rightarrow \delta^*$$

La famiglia \mathcal{U} è **coniugata** della famiglia \mathcal{F} : se $f_U \in \mathcal{U}$ e $f_{Y_t|U=u} \in \mathcal{F}$ comportano che $f_{U|H_T} \in \mathcal{U}$, per ogni H_T , \mathcal{U} è detta famiglia coniugata della famiglia \mathcal{F} .

Si prova che se

- la famiglia \mathcal{U} nella quale si sceglie la distribuzione di U è coniugata della famiglia esponenziale lineare \mathcal{F} nella quale si scelgono le distribuzioni di $Y_1|U = u, Y_2|U = u, \dots,$
- $\exp\{\mu u - b(u)\} = 0$, per $u = u_0, u = u_1$, per ogni μ ,

allora il premio bayesiano coincide con il premio di credibilità per ogni vettore di osservazioni: si ha **credibilità esatta**. Riesce

$$E(Y_{T+1}|Y_1, \dots, Y_T) = \mu(1 - z_T) + \bar{\bar{Y}}_T z_T,$$

dove $\mu = E(\mu_{T+1}(U)) = E(Y_{T+1})$, $z_T = \frac{\omega_\bullet}{\omega_\bullet + \phi/\delta}$, $\bar{\bar{Y}}_T = \sum_{t=1}^T \frac{\omega_t}{\omega_\bullet} Y_t$.

Sono ad esempio famiglie coniugate:

Poisson e Gamma

Gamma e Gamma

Normale e Normale

Binomiale e Beta

Alcuni riferimenti

H. Bühlmann, A. Gisler (2005), *A Course in Credibility Theory and its Applications*, Springer

T. Mikosch (2004), *Non-life insurance mathematics. An introduction with stochastic processes*, Springer

S.A. Klugman, H.H. Panjer, G.E. Willmot (1998), *Loss models. From data to decisions*, Wiley
(prima edizione e seguenti)

SISTEMI BONUS-MALUS (BM) IN ASSICURAZIONI RCA

Gli assicurati di ogni classe tariffaria sono ripartiti in **classi di merito** o **classi BM**. Il premio annuo per un rischio dipende dalla classe tariffaria e dalla classe di merito che gli sono assegnate.

Tipicamente, la classe di merito in un dato anno, t , è determinata dalla **classe di merito** c_{t-1} e dal **numero di sinistri causati dall'assicurato** nell'anno precedente n_{t-1} : $(c_{t-1}, n_{t-1}) \rightarrow c_t$.

Si tiene conto solo del numero di sinistri e non anche dei loro importi. Sistemi Bonus-Malus basati anche sugli importi di danno non sono generalmente usati

- per motivi di semplicità,
- perché, mentre il numero di sinistri avvenuti in un anno è usualmente noto, la valutazione degli importi di risarcimento può richiedere anche molti anni,
- perché si assume che le caratteristiche non osservabili influiscano più sul numero dei sinistri che sulla loro entità, in quanto il comportamento dell'assicurato può influire sul numero di sinistri, in misura inferiore sui loro costi.

Osservazione. Ci sono sistemi nei quali non si ha $(c_{t-1}, n_{t-1}) \rightarrow c_t$: per determinare la classe nell'anno t non si tiene conto solo dell'informazione relativa alla classe occupata nell'anno $t - 1$, ma anche in alcuni anni precedenti. Spesso, in tali sistemi, si possono introdurre classi ausiliarie e l'evoluzione può essere ancora descritta da $(c_{t-1}, n_{t-1}) \rightarrow c_t$.

Inoltre, in pratica, non tutti i sinistri sono “penalizzanti”.

Nel seguito supponiamo che sussistano le $(c_{t-1}, n_{t-1}) \rightarrow c_t$.

Per assegnare un sistema BM, si devono assegnare

- le **classi di merito**, siano $1, \dots, J$,
- le **regole di ingresso**: la classe iniziale per i nuovi assicurati e la classe da assegnare ad un assicurato proveniente da un'altra tariffa,
- le **regole di transizione o regole evolutive**, che determinano la classe di arrivo, noti la classe di provenienza e il numero di sinistri nell'anno precedente, $(c_{t-1}, n_{t-1}) \rightarrow c_t$,
- la **scala dei coefficienti di premio**, sia $\pi_1 \leq \dots \leq \pi_J$.

Nei sistemi BM il premio nell'anno t , per un assicurato in classe tariffaria k e in classe BM j è dato da una formula moltiplicativa

$$P_{t,k,j}^{BM} = P_{t,k} \gamma_k \pi_j,$$

dove

- $P_{t,k}$ è un **premio base** per la classe tariffaria k , nell'anno t , usualmente $P_{t,k} = P_t$,
- γ_k è una “relatività” legata alla classe tariffaria k (ottenuta con la personalizzazione a priori),
- π_j è il coefficiente della scala per la classe BM j .

Il fattore $P_{t,k} \gamma_k$ ($= P_t \gamma_k$) è detto **premio di riferimento** per la classe tariffaria k , nell'anno t .

Dalla

$$P_{t,k,j}^{BM} = P_{t,k}\gamma_k\pi_j,$$

i coefficienti della scala BM determinano la riduzione o l'incremento del premio, rispetto al premio di riferimento della classe tariffaria k ,

- se $\pi_j < 1$, allora $P_{t,k,j}^{BM} < P_{t,k}\gamma_k$: gli assicurati ricevono uno sconto rispetto al premio di riferimento, j è una **classe di bonus**,
- se $\pi_j > 1$, allora $P_{t,k,j}^{BM} > P_{t,k}\gamma_k$: gli assicurati pagano un premio maggiore rispetto al premio di riferimento, j è una **classe di malus**.

In alcuni sistemi c'è una classe BM, sia h , tale che $\pi_h = 1$, tale classe è detta classe BM di riferimento, si ha $P_{t,k,h}^{BM} = P_{t,k}\gamma_k$: il premio è uguale al premio di riferimento della classe tariffaria.

Esempio. Sistema “italiano”: in vigore in Italia prima della liberalizzazione tariffaria 1/7/1994

Classe BM j	Numero di sinistri					Scala BM π_j
	0	1	2	3	4 o >4	
1	1	3	6	9	12	0.50
2	1	4	7	10	13	0.53
3	2	5	8	11	14	0.56
4	3	6	9	12	15	0.59
5	4	7	10	13	16	0.62
6	5	8	11	14	17	0.66
7	6	9	12	15	18	0.70
8	7	10	13	16	18	0.74
9	8	11	14	17	18	0.78
10	9	12	15	18	18	0.82
11	10	13	16	18	18	0.88
12	11	14	17	18	18	0.94
13	12	15	18	18	18	1.00
14	13	16	18	18	18	1.15
15	14	17	18	18	18	1.30
16	15	18	18	18	18	1.50
17	16	18	18	18	18	1.75
18	17	18	18	18	18	2.00

Classe di ingresso per i nuovi assicurati 14.

Un assicurato che riporti, in media, un sinistro ogni 3 anni, rimane nella stessa “zona” del sistema, assicurati con frequenze sinistri $<1/3$ tendono a spostarsi nelle classi di bonus, assicurati con frequenze sinistri $>1/3$ tendono a spostarsi nelle classi di malus.

In presenza di un sistema BM ci sono diversi aspetti importanti da considerare:

- la **capacità del sistema di differenziare adeguatamente i premi** in modo che gli assicurati virtuosi paghino premi inferiori rispetto ai non virtuosi,
- l'effetto di **personalizzazione/solidarietà** che è introdotto dal sistema,
- l'effetto di **autoliquidazione** dei sinistri (***hunger for bonus***): un assicurato può decidere di pagare per proprio conto un sinistro per non incorrere negli aggravi di premio previsti dal sistema: tale effetto dipende, in particolare, dalla “severità” del sistema,
- la **stabilità finanziaria del sistema**: se il sistema non è adeguatamente severo, rispetto alla sinistrosità degli assicurati, in tempi brevi la maggior parte degli assicurati si sposterà nelle classi più basse e allora il totale dei premi potrebbe essere insufficiente per coprire il risarcimento totale atteso. Per riportare in equilibrio tecnico il portafoglio è necessario aumentare il premio di riferimento.

Altri aspetti interessanti sono,

- confrontare sistemi,
- come assegnare le regole evolutive e la scala dei coefficienti di premio.

Alcuni elementi per la valutazione di un sistema BM

Consideriamo dapprima una collettività di assicurati omogenei rispetto a caratteristiche osservabili *a priori* ed “indipendenti”: una classe tariffaria. Omettiamo l’indice della classe.

Supponiamo che il sistema BM sia assegnato. Indichiamo con

P_t il premio di riferimento nell’anno t .

Consideriamo un fissato assicurato nella classe tariffaria e indichiamo con

N_t il numero di sinistri nell’anno t ,

$X_t = \sum_{h=1}^{N_t} Y_{th}$ il risarcimento totale nell’anno t ,

C_t la classe BM nell’anno t .

Ipotizziamo

- indipendenza stocastica del processo $\{N_t, t = 1, 2, \dots\}$ dal processo degli importi di risarcimento, nel senso che $(N_{T+1}|N_1 = n_1, \dots, N_T = n_T) =^d (N_{T+1}|N_1 = n_1, \dots, N_T = n_T, H_T)$, per ogni $H_T = \bigwedge_{\substack{t=1, \dots, T \\ h=1, \dots, n_t}} (Y_{th} \in A_{th})$, con $A_{th} \subset \mathbb{R}$, per ogni T ,
- X_t abbia distribuzione composta con $E(X_t) = E(N_t)E(Y_t)$,
- i processi dei diversi assicurati siano iid.

Poiché C_t è funzione di C_1, N_1, \dots, N_{t-1} , la legge del processo $\{C_t, t = 1, 2, \dots\}$ dipende dalla legge del processo $\{C_1, N_1, N_2, \dots\}$ e dalle regole evolutive del sistema.

Per la valutazione di un sistema, consideriamo i seguenti elementi.

- **Distribuzione di probabilità di C_t :**

$$Pr(C_t = j), \quad j = 1, \dots, J.$$

Tale distribuzione è comune ai rischi della collettività. La sua evoluzione al variare di t permette di valutare l'evoluzione nel tempo degli assicurati tra le classi BM, per effetto delle regole evolutive in relazione alla sinistrosità.

- **Coefficiente medio di premio:**

$$b_t = \sum_{j=1}^J \pi_j Pr(C_t = j).$$

E' il premio atteso nell'anno t se il premio di riferimento è unitario. Dipende dalla distribuzione di C_t e dalla scala dei coefficienti di premio $\pi_j, j = 1, \dots, J$.

Se gli assicurati tendono a portarsi nelle classi più basse, b_t è decrescente al crescere di t .

Descriviamo gli effetti di tale situazione.

$\{C_t\}$ è una catena di Markov se $[C_t]$ rappresenta la classe al tempo t .
 1) $\{C_t\}$ è un processo markoviano
 $Pr\{C_{t+1} = j | C_t = i, K_{t-1}\} =$
 $Pr\{C_{t+1} = j | C_t = i\} = P_{ij}(t)$

2) Il processo markoviano è omogeneo, per cui: $P_{ij}(t) = P_{ij}$
 Probabilità indiscutibile di transizione dello stato: allo stato j in passo

Processo
Markovian

Nell'anno $t = 1$ si fissi il premio di riferimento P_1 pari al **premio di equilibrio** $P_1^{(e)}$, cioè al premio che realizza l'**equilibrio tecnico**, in termini attesi, tra premi e risarcimenti,

$$\sum_{j=1}^J P_1^{(e)} \pi_j Pr(C_1 = j) = E(X_1)$$

↑ ↑
 premio atteso risarcimento atteso
 $P_1^{(e)} b_1$

Assumiamo che le distribuzioni di N_t, X_t siano invarianti al variare di t (stazionarietà).

Se nell'anno t si fissa il premio di riferimento $P_t = P_1 = P_1^{(e)}$, si ha equilibrio tecnico solo se $b_t = b_1$.

Se invece il coefficiente medio di premio ha un andamento decrescente nel tempo, mantenendo fisso il premio di riferimento si ha insufficienza dei premi

$$P_1^{(e)} b_t < P_1^{(e)} b_1 = E(X_1) = E(X_t),$$

nonostante non ci sia un peggioramento della sinistrosità. Per riportare equilibrio, è necessario incrementare il premio di riferimento: $P_t > P_1$.

Ciò ha effetto sulla capacità del sistema di premiare gli assicurati concedendo gli sconti previsti.

Esempio. All'epoca 1,

- assicurato (a) in classe BM h , premio $P_1\pi_h$,
- assicurato (b) in classe BM j , premio $P_1\pi_j$.

Sia $j < h$. L'assicurato (b) riceve uno sconto di premio, rispetto ad (a), pari a

$$\frac{P_1\pi_h - P_1\pi_j}{P_1\pi_h} = 1 - \frac{\pi_j}{\pi_h}.$$

Se (a) si trova in classe BM j , nell'anno t , si attende di ricevere tale sconto rispetto al premio in 1. Ma se il premio di riferimento P_t è aumentato rispetto al premio di riferimento in 1, lo sconto effettivo è

$$1 - \frac{P_t\pi_j}{P_1\pi_h} < 1 - \frac{\pi_j}{\pi_h}.$$

Nel sistema “italiano”, se $h = 14$, $\pi_h = 1.15$. Dopo 5 anni senza sinistri l'assicurato passa in classe $j = 9$ con $\pi_j = 0.78$. Lo sconto atteso sul premio è $1 - 0.78/1.15 \approx 0.32$. Se però, ad esempio, $P_6 = 1.15P_1$, lo sconto effettivo è $1 - 0.78 \approx 0.22$.

- **Premio di equilibrio:**

$$\sum_{j=1}^J P_t^{(e)} \pi_j Pr(C_t = j) = E(X_t)$$

E' il valore del premio di riferimento in t che consente di avere equilibrio tecnico, in termini attesi, tra premi e risarcimenti nell'anno t .

E' un importante indicatore per l'analisi dei sistemi BM che tiene conto: della distribuzione degli assicurati tra le classi di merito, dei coefficienti di premio, del risarcimento atteso.

L'analisi dell'evoluzione nel tempo dei premi di riferimento permette di valutare gli incrementi del premio di riferimento richiesti per mantenere la stabilità finanziaria del sistema.

Per potere calcolare i tre precedenti indicatori si deve introdurre una valutazione probabilistica per il processo $\{C_1, N_1, N_2, \dots\}$ e assegnare $E(Y_t)$.

Richiami. Catene markoviane finite (cenno)

Def. Un processo stocastico $\{C_1, C_2, \dots\}$, dove le determinazioni possibili di C_t sono in numero finito, siano $1, \dots, J$, le medesime per ogni $t = 1, 2, \dots$, è detto **processo markoviano finito** se per ogni i, j e per ogni evento $K_{t-1} = (C_1 = i_1, \dots, C_{t-1} = i_{t-1})$,

$$Pr(C_{t+1} = j | C_t = i, K_{t-1}) = Pr(C_{t+1} = j | C_t = i).$$

Il processo markoviano è detto **omogeneo** o **catena markoviana** se, in aggiunta,

$$Pr(C_{t+1} = j | C_t = i) = p_{ij},$$

ovvero le precedenti probabilità non dipendono da t .

Le determinazioni $1, \dots, J$ sono dette **stati del sistema**. La

$$\mathbf{P} = \begin{bmatrix} p_{11} & \cdots & p_{1J} \\ \cdots & & \cdots \\ p_{J1} & \cdots & p_{JJ} \end{bmatrix}$$

è una matrice stocastica: $p_{ij} \geq 0$, $\sum_{j=1}^J p_{ij} = 1$, per ogni i , detta **matrice delle probabilità condizionate di transizione**.

Posto $\mathbf{a}_t = (a_t(1), \dots, a_t(J)) = (Pr(C_t = 1), \dots, Pr(C_t = J))$, si ha

$$a_2(j) = Pr(C_2 = j) = \sum_{i=1}^J Pr(C_1 = i) \Pr(C_2 = j | C_1 = i) = \sum_{i=1}^J a_1(i) p_{ij}.$$

Disintegabilità
elemento di posto j
del prodotto $\mathbf{a}_1 P$

Segue

$$\mathbf{a}_2 = \mathbf{a}_1 \mathbf{P}.$$

Analogamente,

$$a_3(j) = Pr(C_3 = j) = \sum_{i=1}^J Pr(C_2 = i) Pr(C_3 = j | C_2 = i) = \sum_{i=1}^J a_2(i) p_{ij}.$$

elemento di posto j
del prodotto $\mathbf{a}_2 P$

Segue

$$\mathbf{a}_3 = \mathbf{a}_2 \mathbf{P} = (\mathbf{a}_1 \mathbf{P}) \mathbf{P} = \mathbf{a}_1 \mathbf{P}^{(2)},$$

dove $\mathbf{P}^{(2)} = \mathbf{P} \times \mathbf{P}$.

Per induzione su t , si prova che

$$\mathbf{a}_t = \mathbf{a}_{t-1} \mathbf{P} = \mathbf{a}_1 \mathbf{P}^{(t-1)}.$$

Quindi, assegnate la distribuzione iniziale \mathbf{a}_1 e la matrice \mathbf{P} sono assegnate le distribuzioni marginali unidimensionali.

Si considera solo l'ultimo istante perché si è deciso di fare così

Più in generale, assegnate la distribuzione iniziale a_1 e la matrice P è assegnata la legge del processo. Si ha, applicando prima il teorema delle probabilità composte e poi l'ipotesi markoviana,

$$\begin{aligned} Pr(C_1 = i_1, \dots, C_t = i_t) &= Pr(C_1 = i_1) Pr(C_2 = i_2 | C_1 = i_1) Pr(C_3 = i_3 | C_1 = i_1, C_2 = i_2) \times \dots \\ &\quad \times Pr(C_t = i_t | C_1 = i_1, \dots, C_{t-1} = i_{t-1}) \\ &= Pr(C_1 = i_1) Pr(C_2 = i_2 | C_1 = i_1) Pr(C_3 = i_3 | C_2 = i_2) \dots Pr(C_t = i_t | C_{t-1} = i_{t-1}) \\ &\quad a_1(i_1) \qquad \qquad p_{i_1 i_2} \qquad \qquad p_{i_2 i_3} \qquad \qquad p_{i_{t-1} i_t} \end{aligned}$$

Dalle probabilità sui “segmenti iniziali”, si ottengono le $Pr(C_{t_1} = i_1, \dots, C_{t_n} = i_n)$, per ogni $t_1 < \dots < t_n$. Ad esempio,

$$\begin{aligned} (C_2 = i, C_4 = j) &= (C_2 = i, C_4 = j) \wedge \Omega = (C_2 = i, C_4 = j) \wedge V_{i_1, i_3}(C_1 = i_1, C_3 = i_3) \\ &= V_{i_1, i_3}(C_1 = i_1, C_2 = i, C_3 = i_3, C_4 = j) \end{aligned}$$

Segue

$$Pr(C_2 = i, C_4 = j) = \sum_{i_1, i_3} Pr(C_1 = i_1, C_2 = i, C_3 = i_3, C_4 = j).$$

Gli elementi della potenza $\mathbf{P}^{(\tau)}$ della matrice delle probabilità condizionate di transizione, che indichiamo con $p_{ij}^{(\tau)}$, sono interpretabili come le probabilità di passare in τ passi (anni) dallo stato i allo stato j . Ad esempio,

$$\begin{aligned} Pr(C_{t+2} = j | C_t = i) &= \sum_{h=1}^J Pr(C_{t+2} = j, C_{t+1} = h | C_t = i) \\ &= \underbrace{\sum_{h=1}^J Pr(C_{t+1} = h | C_t = i)}_{p_{ih}} \underbrace{Pr(C_{t+2} = j | C_{t+1} = h, C_t = i)}_{p_{hj} (*)} = p_{ij}^{(2)} \\ &\quad \text{elemento di posto } i,j \text{ della matrice } \mathbf{P} \times \mathbf{P} = \mathbf{P}^{(2)} \end{aligned}$$

Per (*), notiamo che

$$(C_{t+1} = h, C_t = i) = \vee_{i_1, \dots, i_{t-1}} (C_{t+1} = h, C_t = i, C_1 = i_1, \dots, C_{t-1} = i_{t-1}).$$

Allora per la proprietà di disintegrità della probabilità si ha

$$\begin{aligned} Pr(C_{t+2} = j | C_{t+1} = h, C_t = i) &= \sum_{i_1, \dots, i_{t-1}} Pr(C_1 = i_1, \dots, C_{t-1} = i_{t-1} | C_{t+1} = h, C_t = i) \times \\ &\quad \times Pr(C_{t+2} = j | C_{t+1} = h, C_t = i, C_1 = i_1, \dots, C_{t-1} = i_{t-1}) \\ &\quad p_{hj} \\ &= p_{hj} \sum_{i_1, \dots, i_{t-1}} Pr(C_1 = i_1, \dots, C_{t-1} = i_{t-1} | C_{t+1} = h, C_t = i) = p_{hj}. \end{aligned}$$

Per induzione su τ , si prova che

$$Pr(C_{t+\tau} = j | C_t = i) = p_{ij}^{(\tau)}.$$

Def. Una catena markoviana è detta **regolare** se, esiste t_0 tale che, per ogni $t \geq t_0$, $p_{ij}^{(t)} > 0$, per ogni i, j .

Interpretazione: se il numero di passi t è sufficientemente elevato, è positiva la probabilità di raggiungere un qualunque stato del sistema, a partire da qualunque stato.

Teorema di Markov. In una catena markoviana regolare, per ogni i, j esiste

$$\lim_{t \rightarrow +\infty} p_{ij}^{(t)} = v_j > 0, \quad \text{non dipende da } i. \quad (*)$$

è positiva la probabilità di passare da i, j a un altro

Da $\sum_{j=1}^J p_{ij}^{(t)} = 1$, segue che $\sum_{j=1}^J v_j = 1$.

Catena perché dipende dalla sua storia passata
Markoviana perché dipende dall'ultimo passo
Regolare perché c'è la probabilità positiva

Dunque in una **catena markoviana regolare** la potenza $P^{(t)}$ converge, al divergere di t , (limite elemento per elemento) ad una matrice stocastica, con elementi positivi e con righe tutte uguali

$$P^{(t)} \xrightarrow{t \rightarrow +\infty} V = \begin{bmatrix} v_1 & \dots & v_J \\ \dots & \dots & \dots \\ v_1 & \dots & v_J \end{bmatrix}$$

Si noti che la (*) è anche sufficiente perché la catena sia regolare.

Conseguenze.

- Il processo ammette una distribuzione asintotica, la $\nu = (\nu_1, \dots, \nu_J)$, indipendente dalla distribuzione iniziale. Infatti, da $\mathbf{a}_t = \mathbf{a}_1 \mathbf{P}^{(t-1)}$, segue

$$\lim_{t \rightarrow +\infty} \mathbf{a}_t = \mathbf{a}_1 \lim_{t \rightarrow +\infty} \mathbf{P}^{(t-1)} = \mathbf{a}_1 \mathbf{V} = \nu.$$

- La distribuzione asintotica soddisfa la

$$\nu = \nu \mathbf{P}.$$

Segue da

$$\begin{array}{ccc} \mathbf{a}_t & = & \mathbf{a}_{t-1} \mathbf{P}, \\ \downarrow & \downarrow & \downarrow \\ \nu & = & \nu \mathbf{P} \end{array} \quad t \rightarrow +\infty$$

- Si può allora determinare la distribuzione asintotica risolvendo il seguente sistema lineare

$$\begin{cases} \nu = \nu \mathbf{P} \\ \sum_{j=1}^J \nu_j = 1 \end{cases} \Leftrightarrow \begin{cases} \nu(\mathbf{P} - \mathbf{I}) = \mathbf{0} \\ \sum_{j=1}^J \nu_j = 1 \end{cases}$$

Si noti che la matrice $\mathbf{P} - \mathbf{I}$ ha rango minore di J e dunque determinante nullo.

- Se $\mathbf{a}_1 = \nu$, dalla $\nu = \nu \mathbf{P}$, segue che $\mathbf{a}_t = \nu$, per ogni t : stazionarietà



Introduciamo due modelli probabilistici per il processo $\{C_1, N_1, N_2, \dots\}$, per lo studio di sistemi BM.

Modello (a) Si assume che

- $N_1|C_1 = j, N_2|C_1 = j, \dots$ siano iid, per ogni $j = 1, \dots, J$, indipendenza subordinata alla classe
- $Pr(N_t = n|C_1 = j) = p_n$, $n = 0, 1, \dots$, indipendenti da j , assegnate, Dipendono solo da n
- $Pr(C_1 = j)$, $j = 1, \dots, J$, assegnate.

Proposizione 1. Nelle ipotesi del Modello (a), si ha

$$Pr(N_t = n|K_t) = p_n,$$

per ogni n , per ogni t , per ogni $K_t = (C_1 = j_1, \dots, C_t = j_t)$, purché $K_t \neq \emptyset$.

Dim. Fissiamo n , t e $K_t = (C_1 = j_1, C_2 = j_2, \dots, C_t = j_t)$. Osserviamo che l'evento K_t è logicamente dipendente dal processo $\{C_1, N_1, N_2, \dots\}$:

$$K_t = (C_1 = j_1, C_2 = j_2, \dots, C_t = j_t) = \bigvee_{H_{t-1} \rightarrow K_t} H_{t-1},$$

↗ L'unione di tutte le traiettorie che implicano K_t

dove $H_{t-1} = (C_1 = j_1, N_1 = n_1, \dots, N_{t-1} = n_{t-1})$ è una realizzazione iniziale del processo $\{C_1, N_1, N_2, \dots\}$ che, a partire dalla classe BM j_1 al tempo 1, conduce ad occupare le classi j_2 al tempo 2, ..., j_t al tempo t .

Ad esempio. Nel sistema “italiano”:

$$(\mathcal{C}_1 = 14, \mathcal{C}_2 = 13) = (\mathcal{C}_1 = 14, N_1 = 0)$$

$$(\mathcal{C}_1 = 17, \mathcal{C}_2 = 18) = (\mathcal{C}_1 = 17, N_1 = 1) \vee (\mathcal{C}_1 = 17, N_1 = 2) \vee \dots$$

$$(\mathcal{C}_1 = 14, \mathcal{C}_2 = 16, \mathcal{C}_3 = 15) = (\mathcal{C}_1 = 14, N_1 = 1, N_2 = 0)$$

Poiché gli eventi del tipo H_{t-1} sono a due a due incompatibili, per la proprietà di disintegrità

$$\begin{aligned} Pr(N_t = n | K_t) &= \underbrace{\sum_{H_{t-1} \rightarrow K_t} Pr(H_{t-1} | K_t) Pr(N_t = n | H_{t-1})}_{\substack{= Pr(N_t = n | \mathcal{C}_1 = j_1, N_1 = n_1 \dots, N_{t-1} = n_{t-1}) \\ = Pr(N_t = n | \mathcal{C}_1 = j_1) \quad \text{per l'indip. condizionata}}} \\ &= p_n \quad \text{per la seconda ipotesi} \\ \\ &= p_n \underbrace{\sum_{H_{t-1} \rightarrow K_t} Pr(H_{t-1} | K_t)}_{Pr(K_t | K_t) = 1} = p_n \end{aligned}$$



Proposizione 2. Nelle ipotesi del Modello (a), il processo $\{C_1, C_2, \dots\}$ è una catena markoviana.

Dim. Fissiamo t, i, j e $K_{t-1} = (C_1 = j_1, C_2 = j_2, \dots, C_{t-1} = j_{t-1})$ e proviamo che

$$Pr(C_{t+1} = j | C_t = i, K_{t-1}) = Pr(C_{t+1} = j | C_t = i) = p_{ij}.$$

Il primo membro

$$\begin{aligned} Pr(C_{t+1} = j | C_t = i, K_{t-1}) &= \sum_n Pr(C_{t+1} = j, N_t = n | C_t = i, K_{t-1}) \\ &= \underbrace{\sum_n Pr(N_t = n | C_t = i, K_{t-1})}_{p_n} \underbrace{Pr(C_{t+1} = j | N_t = n, C_t = i, K_{t-1})}_{\text{se } \text{\'e } \text{possibile}} \\ &\quad (\text{Proposizione 1}) \\ &\quad \text{non dipende da } t, K_{t-1} \quad \text{non dipende da } t, K_{t-1} \\ &\triangleq \alpha(i, j). \end{aligned}$$

Da quanto sopra, $Pr(C_{t+1} = j | C_t = i, K'_{t-1}) = \alpha(i, j)$, per qualunque K'_{t-1} storia di occupazione delle classi negli anni da 1 a $t - 1$, purché $C_t = i, K'_{t-1} \neq \emptyset$.

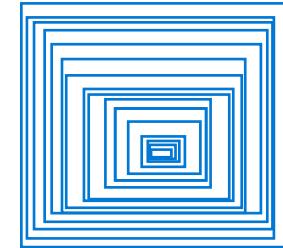
Il secondo membro

$$\begin{aligned} Pr(C_{t+1} = j | C_t = i) &= \sum_{K'_{t-1}} Pr(C_{t+1} = j, K'_{t-1} | C_t = i) \\ &= \sum_{K'_{t-1}} Pr(K'_{t-1} | C_t = i) Pr(C_{t+1} = j | C_t = i, K'_{t-1}) \\ &= \alpha(i, j) \sum_{K'_{t-1}} Pr(K'_{t-1} | C_t = i) = \alpha(i, j). \end{aligned}$$

Pertanto,

$$Pr(C_{t+1} = j | C_t = i, K_{t-1}) = Pr(C_{t+1} = j | C_t = i)$$

e tale probabilità non dipende da t .



■

Osservazioni.

- Per assegnare la legge del processo $\{C_1, C_2, \dots\}$, basta conoscere
 - la distribuzione iniziale $a_1 = (Pr(C_1 = 1), \dots, Pr(C_1 = J))$, che è supposta assegnata,
 - la matrice P delle probabilità condizionate di transizione; notiamo che se sono date le regole evolutive del sistema, la P si ottiene dalle assegnate p_n , $n = 0, 1, \dots$

Esempio. Sistema “italiano”

		n_{t-1}									
		0	1	2	3	4+					
c_t	$c_{t-1} - 1$	$c_{t-1} + 2$	$c_{t-1} + 5$	$c_{t-1} + 8$	$c_{t-1} + 11$						
		1	2	3	4	5	6	...	16	17	18
<i>Claesse anno t-1</i>		1	p_0	0	p_1	0	0	p_2	...	0	0
<i>Claesse anno t</i>		2	p_0	0	0	p_1	0	0	...	0	0
<i>Claesse anno t+1</i>		3	0	p_0	0	0	p_1	0	...	0	0
...		Prob. di fare 0 incidente				Prob. di fare 1 incidente					
...						
17		0	0	0	0	0	0	...	p_0	0	$1 - p_0$
18		0	0	0	0	0	0	...	0	p_0	$1 - p_0$

Le $(Pr(C_1 = 1), \dots, Pr(C_1 = J), p_n, n = 0, 1, \dots)$, possono essere stimate tramite le frequenze relative osservate.

- Se la catena $\{C_1, C_2, \dots\}$ è regolare, esiste la distribuzione asintotica, determinabile risolvendo il sistema $\begin{cases} v(P - I) = 0 \\ \sum_{j=1}^J v_j = 1 \end{cases}$
- Il modello è spesso utilizzato per la valutazione dei sistemi BM, tipicamente pensando di collocare inizialmente tutti gli assicurati in una medesima classe, la classe d'ingresso, e studiando, in particolare, il sistema “a regime” (la distribuzione asintotica).

- Nel modello si calcolano facilmente gli indicatori per la valutazione di un sistema BM, al variare di t ,
 - la distribuzione degli assicurati tra le classi $Pr(C_t = j)$, $j = 1, \dots, J$, $\mathbf{a}_2 = \mathbf{a}_1 \mathbf{P}$, $\mathbf{a}_3 = \mathbf{a}_2 \mathbf{P}, \dots$
 - i coefficienti medi di premio $b_t = \sum_{j=1}^J \pi_j Pr(C_t = j)$,
 - i premi di equilibrio, se sono assegnati i risarcimenti attesi per sinistro $E(Y_t)$; da

$$\sum_{j=1}^J P_t^{(e)} \pi_j Pr(C_t = j) = E(X_t)$$

segue

$$P_t^{(e)} = \frac{E(X_t)}{b_t} = \frac{E(N_t)E(Y_t)}{b_t},$$

con

$$E(N_t) = \sum_{j=1}^J Pr(C_1 = j) \underbrace{E(N_t | C_1 = j)}_{= \sum_n np_n} = \sum_n np_n.$$

- Commento sulle ipotesi del Modello (a)
 - $N_1|C_1 = j, N_2|C_1 = j, \dots$ identicamente distribuiti: si presta alle usuali critiche sulla condizione di stazionarietà,
 - $N_1|C_1 = j, N_2|C_1 = j, \dots$ stocasticamente indipendenti: benché si tratti di indipendenza condizionata all'informazione sulla classe BM iniziale, non è adeguata se, come accade nei modelli con personalizzazione basata sull'esperienza, si ritenga influente, per la valutazione probabilistica, l'informazione sulla sinistrosità passata,
 - $N_t|C_1 = j$ ha legge indipendente dalla classe BM occupata al tempo iniziale: non adeguata se gli assicurati sono già ripartiti in classi di merito all'istante iniziale, in tale caso è naturale pensare di fare dipendere dalla classe BM occupata la valutazione probabilistica.

Modello (b)

Si fa dipendere la legge del processo $\{C_1, N_1, N_2, \dots\}$ da un parametro aleatorio di rischio U che tiene conto del profilo di rischio dell'assicurato.

Si assume che

- $N_1|U = u, C_1 = j, N_2|U = u, C_1 = j, \dots$ siano iid, per ogni $j = 1, \dots, J$, per ogni u determinazione di U ,
- $N_t|U = u, C_1 = j \sim P(\lambda_j u)$, $\lambda_j > 0$, assegnati,
- $U|C_1 = j \sim \text{gamma}(\alpha, \alpha)$, $\alpha > 0$, assegnato,
- C_1 di distribuzione assegnata.

Dalle prime tre condizioni, poiché $(N_t|U = u, C_1 = j) = (N_t|C_1 = j)|(U = u|C_1 = j)$, si ha che

$$\{N_1|C_1 = j, N_2|C_1 = j, \dots\}$$

è un processo Poisson-gamma (λ_j, α) . Pertanto,

$$\{N_1, N_2, \dots\}$$

mistura di processi Poisson-gamma con misturante la distribuzione di C_1 .

In particolare, se $Pr(C_1 = j_0) = 1$, processo Poisson-gamma (λ_{j_0}, α) .

Dal fatto che $\{N_1|C_1 = j, N_2|C_1 = j, \dots\}$ processo Poisson-gamma (λ_j, α) , segue

- $N_t|C_1 = j \sim BN\left(\alpha, \frac{\alpha}{\alpha + \lambda_j}\right)$, con

↗ Sarà λ_j per disintegrobilità

$E[N_{t+1}|N_t = n]$ $\forall t \in \{1, \dots, T\}$

$$E(N_t|C_1 = j) = \lambda_j, \quad var(N_t|C_1 = j) = \lambda_j + \alpha^{-1} \lambda_j^2,$$
- $Pr(N_1 = n_1, \dots, N_T = n_T | C_1 = j) = \frac{\Gamma(\alpha + \sum_{t=1}^T n_t)}{\Gamma(\alpha) \prod_{t=1}^T n_t!} \left(\frac{\alpha}{\alpha + \lambda_j T}\right)^\alpha \left(\frac{\lambda_j}{\alpha + \lambda_j T}\right)^{\sum_{t=1}^T n_t},$
- $N_{T+1}|N_1 = n_1, \dots, N_T = n_T, C_1 = j \sim BN\left(\alpha + \sum_{t=1}^T n_t, \frac{\alpha + \lambda_j T}{\alpha + \lambda_j(T+1)}\right).$

Anche con questo modello si possono calcolare

- $Pr(C_t = h), \quad h = 1, \dots, J$
- $b_t = \sum_{h=1}^J \pi_h Pr(C_t = h),$
- $P_t^{(e)}$ tale che $\sum_{h=1}^J P_t^{(e)} \pi_h Pr(C_t = h) = E(X_t),$

ma per t non troppo elevato.

Quello che importa è la fissazione del processo

Si ha infatti

$$(C_t = h) = \bigvee_{H_{t-1} \rightarrow (C_t=h)} H_{t-1},$$

con $H_{t-1} = (C_1 = j, N_1 = n_1 \dots, N_{t-1} = n_{t-1})$: l'evento $(C_t = h)$ può essere descritto come somma logica di tutte le traiettorie iniziali del processo $\{C_1, N_1, N_2, \dots\}$ che portano l'assicurato in classe BM h nell'anno t . Ad esempio, nel sistema "italiano"

$$(C_2 = 7) = (C_1 = 8, N_1 = 0) \vee (C_1 = 5, N_1 = 1) \vee (C_1 = 2, N_1 = 2).$$

Si devono dunque considerare tutte tali traiettorie. Se t non è troppo elevato, si può ricostruire la storia fino a t e si ha allora

$$Pr(C_t = h) = \sum_{H_{t-1} \rightarrow (C_t=h)} Pr(H_{t-1}),$$

con

$$\begin{aligned} Pr(H_{t-1}) &= Pr(C_1 = j, N_1 = n_1 \dots, N_{t-1} = n_{t-1}) \\ &= Pr(C_1 = j) Pr(N_1 = n_1 \dots, N_{t-1} = n_{t-1} | C_1 = j). \end{aligned}$$

Date le $Pr(C_t = h)$, si trovano i coefficienti medi di premio e, se sono assegnate le $E(Y_t)$, anche i premi di equilibrio, dove $E(X_t) = E(N_t)E(Y_t)$, con

$$E(N_t) = \sum_{j=1}^J Pr(C_1 = j) E(N_t | C_1 = j) = \sum_{j=1}^J Pr(C_1 = j) \lambda_j.$$

Osservazioni.

- Nel processo $\{N_1|C_1 = j, N_2|C_1 = j, \dots\}$ i numeri aleatori non sono stocasticamente indipendenti, il processo è scambiabile.
- Essendo $N_t|C_1 = j \sim BN\left(\alpha, \frac{\alpha}{\alpha + \lambda_j}\right)$, la distribuzione dipende dalla classe BM iniziale tramite il parametro λ_j .
- Il modello richiede di assegnare i parametri λ_j , $j = 1, \dots, J$. Poiché $\lambda_j = E(N_t|C_1 = j)$, si potrebbero stimare tali parametri dai dati, tramite le frequenze sinistri nelle diverse classi BM. Per avere stime affidabili si dovrebbe però disporre di ampie basi di dati in ciascuna classe BM, in pratica la condizione non è usualmente soddisfatta. Notiamo tuttavia, che si può ipotizzare che per alcuni gruppi di classi BM il parametro sia il medesimo, ad es. $\lambda^{(1)}$ per le classi “basse”, $\lambda^{(2)}$ per le classi “medie”, $\lambda^{(3)}$ per le classi “alte”, riducendo in tal modo il numero di parametri da stimare.

Indicatori della capacità di personalizzazione di un sistema BM (cenno)

Consideriamo ancora una collettività di assicurati omogenei, rispetto a caratteristiche osservabili *a priori*, ed “indipendenti”: una classe tariffaria. E supponiamo che il sistema BM sia assegnato. Supponiamo che in $t = 1$, tutti gli assicurati siano posti nella classe iniziale del sistema.

Si possono confrontare i seguenti premi in t

Premio BM	Altri premi
$P_t \pi_{C_t}$	$E(X_t)$ premio <i>a priori</i> , non personalizzazione
	$E(X_t U)$ premio individuale, massima personalizzazione
	$E(X_t C_t)$ premio “ <i>a posteriori</i> ”, stimatore del premio individuale

Si considerano i seguenti elementi

$$A = E[(E(X_t) - P_t \pi_{C_t})^2] \quad \text{se “piccolo”, il sistema realizza poca personalizzazione}$$

$$B = E[(E(X_t|U) - P_t \pi_{C_t})^2] \quad \text{se “piccolo”, il sistema realizza una buona personalizzazione}$$

$$C = E[(E(X_t|C_t) - P_t \pi_{C_t})^2] \quad \text{se “piccolo”, il sistema realizza una buona personalizzazione,}\\ \text{detto } \textit{predictive accuracy}$$

$$\text{E gli indici } i_1 = \left(\frac{A}{A+B} \right)^{1/2}, i_2 = \left(\frac{A}{A+C} \right)^{1/2}.$$

Costruzione di scale di coefficienti di premio (cenno)

Consideriamo ancora una collettività di assicurati omogenei, rispetto a caratteristiche osservabili *a priori*, ed “indipendenti”: una classe tariffaria.

Supponiamo che il sistema BM sia assegnato, a parte la scala, e che sia prevista una classe BM di riferimento: h tale che $\pi_h = 1$.

Supponiamo che in $t = 1$, tutti gli assicurati siano posti nella classe iniziale del sistema.

Con riferimento ad un fissato assicurato, l’evoluzione tra le classi del sistema BM è allora determinata dalle traiettorie del processo $\{N_1, N_2, \dots\}$.

Consideriamo come riferimento metodologico l’approccio bayesiano e, per semplicità, trattiamo solo la componente “premio per la frequenza”.

Assumiamo che la legge del processo $\{N_1, N_2, \dots\}$ dipenda da un parametro aleatorio di rischio U e che si abbia

- $N_1|U = u, N_2|U = u, \dots$ stocasticamente indipendenti, di legge assegnata, per ogni u determinazione possibile di U ,
- U di legge assegnata.

Nell'anno t , il premio individuale è $E(N_t|U) = \mu_t(U)$. Obiettivo è ottenere il premio BM in modo da “stimare” il premio individuale.

1) Il premio BM è una funzione della classe BM, sia $g(C_t)$. Una possibilità per determinare g è richiedere che il premio BM accosti il premio individuale "al meglio" nel senso della perdita quadratica attesa (Norberg 1976)

→ Rappresenta la massima personalizzazione

$$\underset{g \in \mathcal{G}}{\operatorname{argmin}} E[(\mu_t(U) - g(C_t))^2] \equiv \underset{g \in \mathcal{G}}{\operatorname{argmin}} E[(N_t - g(C_t))^2]$$

La soluzione è $g^*(j) = E [\mu_t(U)|C_t = j] = E [N_t|C_t = j]$.

Essendo h la classe BM di riferimento, si pone

$$\pi_j = \frac{g^*(j)}{g^*(h)} = \frac{E [N_t|C_t = j]}{E [N_t|C_t = h]}.$$

→ faccio varire j
↳ h è bloccata

Infatti, i coefficienti della scala BM hanno il significato di rapporti tra il premio di ciascuna classe BM e il premio della classe BM di riferimento.

ipotesi più generali rispetto al modello b

Esempio. Sia $\{N_1, N_2, \dots\}$ è un processo Poisson-gamma (λ, α) . Poiché si ha

$$(C_t = j) = V_{H_{t-1} \rightarrow (C_t=j)} H_{t-1},$$

con $H_{t-1} = (N_1 = n_1, \dots, N_{t-1} = n_{t-1})$, per la proprietà di disintegrità della speranza matematica

$$E [N_t | C_t = j] = \sum_{H_{t-1} \rightarrow (C_t=j)} Pr(H_{t-1} | C_t = j) E(N_t | H_{t-1})$$

↳ La sommatoria su tutte le lotterie

Non chiede

con

$$- E(N_t | H_{t-1}) = E(N_t | N_1 = n_1, \dots, N_{t-1} = n_{t-1}) = \lambda \frac{\alpha + \sum_{s=1}^{t-1} n_s}{\alpha + \lambda(t-1)}$$

$$- Pr(H_{t-1} | C_t = j) = \frac{Pr(H_{t-1})}{Pr(C_t=j)} = \frac{Pr(N_1=n_1, \dots, N_{t-1}=n_{t-1})}{Pr(C_t=j)}$$

$$= \frac{\frac{\Gamma(\alpha + \sum_{s=1}^{t-1} n_s)}{\Gamma(\alpha) \prod_{s=1}^{t-1} n_s!} \left(\frac{\alpha}{\alpha + \lambda(t-1)} \right)^\alpha \left(\frac{\lambda}{\alpha + \lambda(t-1)} \right)^{\sum_{s=1}^{t-1} n_s}}{\sum_{H_{t-1} \rightarrow (C_t=j)} Pr(H_{t-1})}$$

■

Come è evidente dall'ultimo membro,

$$\pi_j = \frac{g^*(j)}{g^*(h)} = \frac{E [N_t | C_t = j]}{E [N_t | C_t = h]}.$$

tale scala dipende da t . Per ovviare a tale inconveniente si può effettuare le valutazioni per t “elevato”, quando il sistema è a regime.

2) Per svincolarsi da t , per determinare g si può richiedere che (Borgan, Hoem, Norberg 1981)

$$\underset{g \in \mathcal{G}}{\operatorname{argmin}} \sum_{t=1}^T w_t E[(N_t - g(C_t))^2],$$

dove $w_t \geq 0$, sono pesi opportuni.

Si prova che la soluzione è

$$g^*(j) = \frac{\sum_{t=1}^T w_t Pr(C_t=j) E[N_t | C_t=j]}{\sum_{t=1}^T w_t Pr(C_t=j)}.$$

Si pone

$$\pi_j = \frac{g^*(j)}{g^*(h)}.$$

3) Per determinare la scala si può ragionare come segue (Coene, Doray 1996, Grasso 1997).

Nell'approccio bayesiano il premio individuale è stimato dal premio bayesiano, vista la storia

$$H_{t-1} = (N_1 = n_1 \dots, N_{t-1} = n_{t-1}),$$

$$E(N_t | H_{t-1}) = E(N_t) \underbrace{\frac{E(N_t | H_{t-1})}{E(N_t)}}_{\text{coefficiente di aggiornamento bayesiano}} \\ \triangleq c_t(n_1 \dots, n_{t-1})$$

Ad esempio, nel modello Poisson gamma

$$c_t(n_1 \dots, n_{t-1}) = \frac{\alpha + \sum_{s=1}^{t-1} n_s}{\alpha + \lambda(t-1)}.$$

Nel sistema BM, vista la storia $H_{t-1} = (N_1 = n_1 \dots, N_{t-1} = n_{t-1})$, a partire dalla classe d'ingresso, l'assicurato è collocato in una classe BM che indichiamo con $j(n_1 \dots, n_{t-1})$ alla quale corrisponde il coefficiente di premio $\pi_{j(n_1 \dots, n_{t-1})}$.

Notiamo che i due coefficienti

$$c_t(n_1 \dots, n_{t-1}) \quad \pi_{j(n_1 \dots, n_{t-1})}$$

hanno significato analogo

- $c_t(n_1 \dots, n_{t-1})$ è il coefficiente che moltiplicato per il premio *a priori* permette di ottenere il premio basato sull'esperienza individuale,
- $\pi_{j(n_1 \dots, n_{t-1})}$ è il coefficiente che moltiplicato per il premio di riferimento permette di ottenere il premio basato sull'esperienza individuale.

Allora, per determinare i coefficienti della scala BM, si può porre come obiettivo l'accostamento tra i due coefficienti

$$\min_{\pi_1, \dots, \pi_J} \sum_{t=1}^T w_{t,n_1 \dots, n_{t-1}} \left(\pi_{j(n_1 \dots, n_{t-1})} - c_t(n_1 \dots, n_{t-1}) \right)^2,$$

dove $w_{t,n_1 \dots, n_{t-1}} \geq 0$, sono pesi opportuni.

Si considerano dei vincoli, quali $\pi_{j+1} \geq \pi_j$ per ogni j ; $\pi_h = 1$; una condizione di sufficienza dei premi,... Il problema può essere risolto con procedimento numerico.

Concludiamo con le seguenti osservazioni.

Abbiamo considerato una collettività di rischi omogenei rispetto a caratteristiche osservabili a priori, una classe tariffaria. Un portafoglio è costituito da diverse classi tariffarie.

Le analisi di valutazione dei sistemi BM possono essere sviluppate in un portafoglio ripartito in classi tariffarie considerando gli elementi introdotti per ciascuna classe tariffaria.

La costruzione di una scala di coefficienti BM di portafoglio richiede invece di tenere conto congiuntamente dei due tipi di personalizzazione, introducendo anche un processo per descrivere l'evoluzione tra le classi tariffarie.

Alcuni riferimenti

M. Denuit, X. Marechal, S. Pitrebois, J. Walhin (2007), *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-malus Systems*, Wiley

R. Kaas, M. Goovaerts, J. Dhaene, M. Denuit (2001), *Modern actuarial risk theory Using R*, (prima e seconda edizione), Kluwer

J. Lemaire (1995), *Bonus-malus systems in automobile insurance*, (seconda edizione), Kluwer