Estimarea parametrilor structurali în modelul de credibilitate liniar si neomogen al lui Bühlmann

Asist.drd. Virginia Atanasiu,
Catedra de matematica, A.S.E.

Tema acestui articol vizeaza primele de credibilitate obtinute în cadrul modelului clasic al lui Bühlmann. Estimatorii optimi liniari si neomogeni determinati pentru modelul de credibilitate respectiv, contin parametrii de structura necunoscuti, ceea ce face imposibila folosirea lor ca atare. Deci, considerentul de ordin practic impune necesitatea estimarii caracteristicilor de portofoliu implicate în rezultatele la care s-a ajuns pe plan teoretic. Prezentarea realizata în cadrul materialului de fata, reprezinta solutia modelului clasic al lui Bühlmann în cazul estimatorilor de credibilitate liniari si neomogeni pentru primele contractelor din portofoliile de asigurari non-viata.

***Cuvinte cheie**: parametrii structurali, estimatori nedeplasati, modelul clasic al lui Bühlmann, prima empirica de credibilitate.*

Fie un portofoliu ca cel descris în diagrama 1:

Contract:			1...j...k
Variabila de structur\:			q_j
Variabilele observabile:	d	1	X_{j1}
	u	2	X_{j2}
	r	.	.
	a	.	.
	t	.	.
	a	t	X_{jt}
[ani]			

Diagrama 1. Modelul clasic al lui Bühlmann

constând din k contracte de asigurare non-viata, identice si independente. Fiecare polita j implica un risc X_j asimilat cu o variabila aleatoare nenegativa. Parametrul riscului X_j este notat cu q_j si presupus variabila aleatoare reala. Daca X_j s-a aflat sub cercetare (observatie), iar studiul statistic efectuat a înregistrat t (≥2) ani, atunci dispunem acum de un trecut statistic asupra lui X_j concretizat în variabilele aleatoare observabile $X_{j1}, X_{j2}, ..., X_{jt}$.

Observatii: - parametrul de risc q_j defineste caracteristicile riscului luat în consideratie pentru polita j;
- admitem ca riscul X_j este stabil (constant) în timp, ceea ce înseamna ca poate fi caracterizat prin aceeasi valoare a parametrului de risc q_j de-a lungul anilor de valabilitate ai contractului j;
- contractul j se identifica cu un vector aleator de componente: q_j (parametrul aleator de structura) si $X_{j1}, X_{j2}, ..., X_{jt}$ (observatiile efectuate asupra politei j); deci, contractul j este reprezentat de perechea $(?_j, \underline{X}`_j)$, unde $\underline{X}`_j = (X_{j1}, X_{j2}, ..., X_{jt})$.

Ipoteza care afirma despre contractele $j = \overline{1,k}$ (cuplurile $(?_j, \underline{X}`_j)$, $j = \overline{1,k}$) ca sunt independente si identic distribuite, revine la faptul urmator: pentru $q_j = \theta_j$, variabilele $X_{j1}, X_{j2}, ..., X_{jt}$ sunt independente si identic distribuite conditional.

Prima neta de risc a contractului j, daca parametrul de risc al acestuia este q_j, se defineste astfel:

$$\mu(?_j) = M(X_{jr} \mid ?_j), \forall r = \overline{1,t} \qquad (1),$$

unde $\mu(\bullet)$ este o functie cu valori reale depinzând de q_j, dar independenta de $r=\overline{1,t}$.

Toate contractele au în comun aceea ca variantele lor pot fi exprimate prin functii $\sigma^2(\bullet)$ de parametrul riscului, independente de $r=\overline{1,t}$, cu valori pozitive si anume:

$$\sigma^2(?_j)=Var(X_{jr} \mid ?_j), \forall r=\overline{1,t} \quad (2).$$

Pentru fiecare contract j rezulta urmatoarea matrice de covariante a observatiilor din anii $r=\overline{1,t}$:

$$Cov(\underline{X}_j \mid ?_j)=\sigma^2(?_j)I^{(t,t)} \quad (3),$$

unde $I^{(t,t)}$ indica matricea unitate de ordinul $(t\times t)$. Facem remarca, de altfel intuiutiva, asupra observatiilor X_{jr}, $r=\overline{1,t}$: variabilele observabile ale contractului j au varianta finita.

Afirmatiile de pâna acum alcatuiesc contextul modelului clasic al lui Bühlmann.

Date fiind conditiile acestui model, estimatorii optimi (în sensul celor mai mici patrate) liniari si neomogeni M_j^a pentru $\mu(?_j)$, $j=\overline{1,k}$ sunt:

$$\mu(\hat{?}_j)=M_j^a=zM_j+(1-z)m \quad (4),$$

unde:

$M_j=\overline{X}_j=\frac{1}{t}\sum_{s=1}^{t}X_{js}$ indica estimatorul individual al lui $\mu(?_j)$, iar:

$$z=\frac{at}{at+s^2}$$

factorul de credibilitate rezultat, cu:

$$a=Var[\mu(?_j)] \quad (5),$$

$$s^2=M[\sigma^2(?_j)] \quad (6),$$

$$m=M(X_{jr})=M[\mu(?_j)] \quad (7).$$

<u>Observatie</u>: Marimile definite la (5), (6) si (7) reprezinta asa-numitii parametrii structurali sau caracteristicile de portofoliu si ele sunt necunoscute.

Asadar, primele de credibilitate (4) pentru acest model implica trei parametrii necunoscuti: a, s^2 si m. Pentru a putea folosi rezultatul (4) trebuie sa estimam aceste caracteristici ale portofoliului schisat în diagrama 1.

Deoarece politele sunt incluse într-un colectiv de contracte identice, avem mai mult de o observatie disponibila privind parametrul de risc, încât putem atribui estimatori nedeplasati parametrilor de structura necunoscuti.

<u>Teorema 1.</u> (Estimarea parametrilor structurali din modelul de credibilitate liniar si neomogen al lui Bühlmann)

Variabilele aleatoare:

$$\hat{m}=M_0=\frac{1}{kt}\sum_{j=1}^{k}\sum_{r=1}^{t}X_{jr}=\frac{1}{k}\sum_{j=1}^{k}\overline{X}_j=\frac{1}{k}\sum_{j=1}^{k}M_j \ (8),$$

$$\hat{s^2}=\frac{1}{k(t-1)}\sum_{j=1}^{k}\sum_{r=1}^{t}(X_{jr}-M_j)^2 \quad (9),$$

$$\hat{a}=\frac{1}{k-1}\sum_{j=1}^{k}(M_j-M_0)^2-\frac{1}{t}\hat{s^2} \quad (10)$$

sunt estimatori nedeplasati ai parametrilor structurali corespunzatori: m, s^2 si a, adica:

$$M\left(\hat{m}\right)=m, M\left(\hat{s^2}\right)=s^2, M\left(\hat{a}\right)=a \quad (11).$$

<u>Demonstratie</u>:

A arata ca $M\left(\hat{m}\right)=m$ este trivial. Pentru fiecare polita j varianta empirica:

$$\frac{1}{t-1}\sum_{r=1}^{t}(X_{jr}-\overline{X}_j)^2 \quad (12)$$

este un estimator nedeplasat al lui $Var(X_{jr} \mid ?_j)$ si astfel:

$$\hat{s^2}=\frac{1}{k(t-1)}\sum_{j=1}^{k}\sum_{r=1}^{t}(X_{jr}-\overline{X}_j)^2 \quad (13)$$

este un estimator nedeplasat al lui s^2.

Sa justificam întâi ca:

$$M\left[\left(\frac{1}{t-1}\sum_{r=1}^{t}\left(X_{jr}-\overline{X}_j\right)^2\right)\mid ?_j\right]=\mathrm{Var}\left(X_{jr}\mid ?_j\right)$$ (14).

în acest sens, tinem cont de identitatea urmatoare:

$$\frac{1}{t}\sum_{r=1}^{t}\left(X_{jr}-\overline{X}_j\right)^2=\frac{1}{t}\sum_{r=1}^{t}\left[X_{jr}-M\left(X_{jr}\mid ?_j\right)\right]^2--\left[\overline{X}_j-M\left(X_{jr}\mid ?_j\right)\right]^2 \qquad (15).$$

Deci, are loc sirul de egalitati:

$$M\left[\left(\frac{1}{t}\sum_{r=1}^{t}\left(X_{jr}-\overline{X}_j\right)^2\right)\mid ?_j\right]\overset{(15)}{=}\frac{1}{t}\sum_{r=1}^{t}M\left(\left\{\left[X_{jr}-M\left(X_{jr}\mid ?_j\right)\right]^2\right\}\mid ?_j\right)-$$
$$-M\left(\left\{\left[\overline{X}_j-M\left(X_{jr}\mid ?_j\right)\right]^2\right\}\mid ?_j\right)=\frac{1}{t}\sum_{r=1}^{t}\mathrm{Var}\left(X_{jr}\mid ?_j\right)-\mathrm{Var}\left(\overline{X}_j\mid ?_j\right)\overset{(2)}{=}$$
$$\overset{(2)}{=}\frac{1}{t}t\sigma^2\left(?_j\right)-\frac{1}{t}\mathrm{Var}\left(X_{jr}\mid ?_j\right)\overset{(2)}{=}\mathrm{Var}\left(X_{jr}\mid ?_j\right)-\frac{1}{t}\mathrm{Var}\left(X_{jr}\mid ?_j\right)=$$
$$=\frac{t-1}{t}\mathrm{Var}\left(X_{jr}\mid ?_j\right)$$

Aceasta implica relatia (14); într-adevar:

$$M\left[\left(\frac{1}{t-1}\sum_{r=1}^{t}\left(X_{jr}-\overline{X}_j\right)^2\right)\mid ?_j\right]=$$
$$=\frac{t}{t-1}M\left[\left(\frac{1}{t}\sum_{r=1}^{t}\left(X_{jr}-\overline{X}_j\right)^2\right)\mid ?_j\right]=$$
$$=\frac{t}{t-1}\frac{t-1}{t}\mathrm{Var}\left(X_{jr}\mid ?_j\right)=\mathrm{Var}\left(X_{jr}\mid ?_j\right).$$

Acum stabilim ca: $M\left(\hat{s^2}\right)=s^2$ (16).

Întrucât:

$$M\left[\left(X_{jr}-\overline{X}_j\right)^2\right]=M\left\{M\left[\left(X_{jr}-\overline{X}_j\right)^2\mid ?_j\right]\right\}=$$
$$=M\left\{M\left(X^2{}_{jr}\mid ?_j\right)+M\left(\overline{X}_j^2\mid ?_j\right)-2M\left[\left(X_{jr}\overline{X}_j\right)\mid ?_j\right]\right\}=$$
$$=M\left[\mathrm{Var}\left(X_{jr}\mid ?_j\right)\right]+M\left[M^2\left(X_{jr}\mid ?_j\right)\right]+$$
$$+M\left[M\left(\overline{X}_j^2\mid ?_j\right)\right]-2M\left\{M\left[\left(X_{jr}\frac{1}{t}\sum_{s=1}^{t}X_{js}\right)\mid ?_j\right]\right\}\overset{(1)}{=}$$
$$\overset{(1)}{=}M\left[\sigma^2\left(?_j\right)\right]+M\left[\mu^2\left(?_j\right)\right]+M\left(\overline{X}_j^2\right)-\frac{2}{t}\sum_{s=1}^{t}M\left[M\left(X_{jr}X_{js}\mid ?_j\right)\right]\overset{(6)}{=}$$
$$\overset{(6)}{=}s^2+\mathrm{Vat}\left[\mu\left(?_j\right)\right]+M^2\left[\mu\left(?_j\right)\right]+\mathrm{Var}\left(\overline{X}_j\right)+M^2\left(\overline{X}_j\right)-$$

$$-\frac{2}{t}\left\{M\left[M\left(X^2{}_{jr} \mid ?_j\right)\right]+\sum_{\substack{s=1\\ s\neq r}}^{t} M\left[M\left(X_{jr} \mid ?_j\right)M\left(X_{js} \mid ?_j\right)\right]\right\}\overset{(5)}{\underset{(7)(1)}{=}}$$

$$\overset{(5)}{\underset{(7)(1)}{=}} s^2+a+m^2+\frac{s^2}{t}+a+m^2-\frac{2}{t}\left\{M\left[Var\left(X_{jr} \mid ?_j\right)\right]+\right.$$

$$\left.+M\left[M^2\left(X_{jr} \mid ?_j\right)\right]+\sum_{\substack{s=1\\ s\neq r}}^{t} M\left[\mu^2\left(?_j\right)\right]\right\}\overset{(2)}{=}s^2+2a+2m^2+\frac{s^2}{t}-$$

$$-\frac{2}{t}\left\{M\left[\sigma^2\left(?_j\right)\right]+M\left[\mu^2\left(?_j\right)\right]+(t-1)\left(Var\left[\mu\left(?_j\right)\right]+M^2\left[\mu\left(?_j\right)\right]\right)\right\}\overset{(7)}{\underset{(6)(5)}{=}}$$

$$\overset{(7)}{\underset{(6)(5)}{=}} s^2+2a+2m^2+\frac{s^2}{t}-\frac{2}{t}\left\{s^2+Var\left[\mu\left(?_j\right)\right]+M^2\left[\mu\left(?_j\right)\right]+(t-1)\left(a+m^2\right)\right\}\overset{(5)}{=}$$

$$\overset{(5)}{=}s^2+2a+2m^2+\frac{s^2}{t}-\frac{2}{t}\left\{s^2+a+m^2+(t-1)\left(a+m^2\right)\right\}=$$

$$=s^2+2a+2m^2+\frac{s^2}{t}-\frac{2}{t}s^2-\frac{2}{t}\left(a+m^2\right)-\frac{2}{t}(t-1)\left(a+m^2\right)=$$

$$=s^2+2a+2m^2-\frac{s^2}{t}-\frac{2}{t}\left(a+m^2\right)(1+t-1)=$$

$$=s^2+2a+2m^2-\frac{s^2}{t}-2a-2m^2=s^2-\frac{s^2}{t}=\frac{t-1}{t}s^2$$

sir de egalitati în care s-a avut în vedere formula:

$$Var\left(\overline{X}_j\right)=\frac{s^2}{t}+a \qquad (17).$$

Demonstratia relatiei (17) decurge astfel:

$$Var\left(\overline{X}_j\right)=M\left[Var\left(\overline{X}_j \mid ?_j\right)\right]+Var\left[M\left(\overline{X}_j \mid ?_j\right)\right]=$$

$$=M\left\{Var\left[\left(\frac{1}{t}\sum_{r=1}^{t}X_{jr}\right) \mid ?_j\right]\right\}+Var\left\{M\left[\left(\frac{1}{t}\sum_{r=1}^{t}X_{jr}\right) \mid ?_j\right]\right\}=$$

$$=\frac{1}{t^2}\sum_{r=1}^{t}M\left[Var\left(X_{jr} \mid ?_j\right)\right]+Var\left[\frac{1}{t}\sum_{r=1}^{t}M\left(X_{jr} \mid ?_j\right)\right]\overset{(2)}{\underset{(6)(1)}{=}}$$

$$\overset{(2)}{\underset{(6)(1)}{=}}\frac{1}{t^2}ts^2+Var\left[\frac{1}{t}t\mu\left(?_j\right)\right]\overset{(5)}{=}\frac{s^2}{t}+a$$

Prin urmare:

$$M\left[\hat{s}^2\right]\overset{(9)}{=}\frac{1}{k(t-1)}\sum_{j=1}^{k}\sum_{r=1}^{t}M\left[\left(X_{jr}-\overline{X}_j\right)^2\right]=$$

$$=\frac{1}{k(t-1)}\sum_{j=1}^{k}\sum_{r=1}^{t}\frac{t-1}{t}s^2=\frac{1}{k(t-1)}kt\,\frac{t-1}{t}s^2=s^2,$$

de unde rezulta valabilitatea relatiei (16).

Varianta empirica:

$$\frac{1}{k-1}\sum_{j=1}^{k}\left(\overline{X}_j - M_0\right)^2 \qquad (18)$$

este un estimator nedeplasat al lui $Var\left(\overline{X}_j\right)$, ceea ce înseamna ca:

$$M\left[\frac{1}{k-1}\sum_{j=1}^{k}\left(\overline{X}_j - M_0\right)^2\right] = Var\left(\overline{X}_j\right) \qquad (19),$$

afirmatie ce are la baza urmatorul rezultat din statistica matematica:

“Fie X_1, X_2, ..., X_n un sir de variabile aleatoare independente si identic distribuite, iar $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Atunci varianta empirica:

$$\frac{1}{n-1}\sum_{j=1}^{n}\left(X_j - \overline{X}_n\right)^2 \qquad (20)$$

este un estimator nedeplasat pentru $Var(X_i)$ $(i=\overline{1,n})$”.

Deoarece are loc (17), deducem ca:

$$a = Var\left(\overline{X}_j\right) - \frac{s^2}{t} \qquad (21)$$

si deci putem introduce estimatorul nedeplasat:

$$\hat{a} = \frac{1}{k-1}\sum_{j=1}^{k}\left(\overline{X}_j - M_0\right)^2 - \frac{\hat{s}^2}{t} \qquad (21)$$

pentru a. într-adevar:

$$M\left(\hat{a}\right) = a \qquad (22),$$

asa dupa cum se observa din cele de mai jos:

$$M\left(\hat{a}\right) \overset{(21)}{=} M\left[\frac{1}{k-1}\sum_{j=1}^{k}\left(\overline{X}_j - M_0\right)^2\right] - \frac{1}{t}M\left(\hat{s}^2\right) \overset{(19)}{=}$$

$$\overset{(19)}{=} Var\left(\overline{X}_j\right) - \frac{1}{t}s^2 \overset{(21)}{=} a$$

Acum, demonstratia teoremei 1 este completa.

Remarca: Estimatorul (10) prezinta inconvenientul ca poate lua si valori negative, în - timp ce a este nenegativ. Din acest motiv, se înlocuieste a cu estimatorul:

$$a^* = \max\left\{0, \hat{a}\right\} \qquad (23).$$

Definitia (23) a lui a^* arata ca estimatorul nu mai ramâne nedeplasat pentru a (îsi pierde calitatea de a fi nedeplasat pentru a) însa obtine admisibilitatea (câstiga acceptabilitatea).

Observatii: 1) Daca $k\to\infty$, atunci $\hat{m}$, $\hat{s}^2$ si a^* devin estimatori consistenti pentru parametrii m, s^2 si a, ceea ce din punct de vedere practic este foarte satisfacator. Afirmatiile formulate anterior se justifica astfel:

- din (8), rezulta ca $\hat{m}$ fiind media aritmetica a k variabile aleatoare $\overline{X}_j, j=\overline{1,k}$, care urmeaza aceeasi lege de repartitie, cu aceeasi valoare medie finita: $M\left(\overline{X}_j\right) = m, j=\overline{1,k}$, îi putem aplica legea numerelor mari (conform teoremei lui A.I. Hincin), ceea ce înseamna ca $\hat{m}$ converge în probabilitate catre m, când $k\to\infty$;
- din (9), rezulta ca $\hat{s}^2$ fiind media aritmetica a k variabile aleatoare $Z_j = \frac{1}{t-1}\sum_{r=1}^{t}\left(X_{jr} - \overline{X}_j\right)^2, j=\overline{1,k}$, care urmeaza aceeasi lege de repartitie, cu aceeasi valoare medie finita: $M\left(Z_j\right) = M\left[M\left(Z_j \mid ?_j\right)\right] = M\left[Var\left(X_{jr} \mid ?_j\right)\right] = s^2$, îi putem aplica legea numerelor mari (conform teoremei lui A.I. Hincin), ceea ce înseamna ca $\hat{s}^2$ converge în probabilitate catre s^2, când $k\to\infty$;
- are loc sirul de egalitati urmator:

$$\hat{a}_+ \overset{(10)}{=} \frac{1}{k-1}\sum_{j=1}^{k}\left(\overline{X}_j - M_0\right)^2 - \frac{1}{t}\hat{s}^2 =$$
$$= \frac{k}{k-1}\frac{1}{k}\sum_{j=1}^{k}\left(\overline{X}_j - M_0\right)^2 - \frac{1}{t}\hat{s}^2 =$$
$$= \frac{k}{k-1}\left(\frac{1}{k}\sum_{j=1}^{k}\overline{X}_j^{\,2} - M_0^{\,2}\right) - \frac{1}{t}\hat{s}^2 \qquad (24)$$

Cum: $\frac{k}{k-1}\xrightarrow{P}1;\ \frac{1}{k}\sum_{j=1}^{k}\overline{X}_j^2 \xrightarrow{P} \frac{s^2}{t} + a + m^2$ (deoarece $\left(\frac{1}{k}\sum_{j=1}^{k}\overline{X}_j^2\right)$ este media aritmetica a k variabile aleatoare $\overline{X}_j^2, j=\overline{1,k}$, care urmeaza aceeasi lege de repartitie, cu aceeasi valoare medie finita: $M\left(\overline{X}_j^2\right) = Var\left(\overline{X}_j\right) + M^2\left(\overline{X}_j\right) \overset{(17)}{=}$ $\overset{(17)}{=} \frac{s^2}{t} + a + m^2, j=\overline{1,k}$, rezulta ca îi putem aplica legea numerelor mari, conform teoremei lui A.I. Hincin); $M_0 \overset{(8)}{=} \hat{m} \xrightarrow{P} m; \hat{s}^2 \xrightarrow{P} s^2$ deducem (vezi (24)) ca:

$$\hat{a}_+ \xrightarrow{P} 1\left(\frac{s^2}{t} + a + m^2 - m^2\right) - \frac{1}{t}s^2 = a.$$

2) în conditiile teoremei 1, $z = \frac{at}{at + s^2}$ se poate estima cu $\hat{z} = \frac{a^* t}{a^* t + \hat{s}^2}$. Inserând estimatorii $\hat{z}$ si $\hat{m}$ în prima de credibilitate (4), se obtine prima empirica de credibilitate:

$$M_j^a = \hat{z} M_j + \left(1 - \hat{z}\right) M_0;\ j = \overline{1,k} \qquad (25),$$

care este, asa dupa cum se poate constata, o combinatie liniara omogena a tuturor variabilelor aleatoare observabile. în cazul folosirii formulei (25), se observa ca $M\left(M_j^a\right) \neq m$, deoarece $\hat{z}$ este dependent de M_j si M_0. Prin urmare, proprietatea atractiva din punct de vedere practic a lui $M^a{}_j$ (vezi (25)) si anume caracterul nedeplasat este pierdut, dar înca ne putem astepta ca estimatorii rezultati sa fie buni.

3) Rezultatul (4) si teorema 1. conduc la solutia modelului clasic al lui Bühlmann, în cazul unui estimator optim liniar si neomogen pentru $\mu(?_j)$ (sau, ceea ce este echivalent, pentru $X_{j,t+1}$).

Bibliografie:

[1] Goovaerts, M.J., Kass, R., Van Heerwaarden, A.E., Bauwelinckx, T., - *Insurance Series, volume 3, Efective actuarial methods*, University of Amsterdam, The Netherlands (1990).

[2] Pentikainen, T., Daykin, C.D., Pesonen, M., - *Practical Risk Theory for Actuaries*, Universite Pieere et Marie Curie (1990).

[3] Sundt, B., *An Introduction to Non-Life Insurance Mathematics*, Veroffentlichungen de Instituts für Versicherungswissenschaft de Universitat Mannheim Band 28 (1984 VVW Karlsruhe).