Exemplul lui Norberg ca mod de ilustrare a aplic\rii teoriei credibilit\]ii ^n practic\

Asist. Virginia ATANASIU,
Catedra de matematic\, A.S.E. Bucuresti

~n acest articol, realiz\m o descriere a unui model de portofoliu eterogen, din perspectiva practicii asigur\rilor non-via]\, mai precis, a exemplului elaborat de actuarul Norberg, analizându-l din punct de vedere al teoriei credibilit\]ii, cu ajutorul instrumentului matematic pus la dispozi]ie de teoria probabilit\]ilor. Presupunem cunoscut\ no]iunea de eterogenitate, care intervine ^n mod esen]ial cu aceast\ ocazie (vezi articolul "Verificarea eterogenit\]ii unui portofoliu de contracte de asigurare prin metode statistice clasice" din Revista "Informatica Economic\", num\rul 6). Exemplul supus studiului constituie, ^n acela[i timp, o ilustrare a modului ^n care statistica clasic\ poate verifica eterogenitatea unui portofolui (colectiv) de contracte (poli]e) de asigurare.

Cuvint cheie: *portofoliu eterogen, Norberg, frecven]\ relativ\, testul hi-p\trat.*

Exemplu datorat lui Norberg prezint\ ca-zul unui portofoliu eterogen de contracte de asigurare. ~n acest sens, se consider\ 20 de poli]e, ale c\ror riscuri urmeaz\ o reparti]ie de tip Bernoulli standard, dat\ prin func]ia de probabilitate:

$$f(x,q) = q^x (1-q)^{1-x},\ x \in \{0,1\}, 0 < q < 1,$$

unde parametrii "θ" sunt egali cu: θ_1 relativ la contractul 1, θ_2 relativ la contractul 2,..., θ_{20} relativ la contractul 20 (de fapt θ_i, $i = \overline{1,20}$ reprezint\ parametrii abstrac]i ai celor 20 de riscuri); θ_1, θ_2,...θ_{20} diferi]i ^ntre ei.

Riscurile implicate de cele 20 contracte s-au aflat sub observa]ie statistic\ timp de zece ani, astfel ^ncât, acum dispunem de un trecut statistic pentru ele de forma celui din diagrama care urmeaza.

Poli]a nr.→																				
An	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1			1								1						1	1		
2							1		1		1						1			
3			1				1		1	1										
4																				
5											1						1			
6						1			1			1								
7									1					1						
8									1			1					1			
9						1					1		1							
10									1			1					1		1	
$\hat{q}_j$	.0	.0	.2	.0	.0	.2	.2	.0	.6	.1	.4	.3	.1	.1	.0	.0	.5	.1	.1	.0

Diagrama se interpreteaz\ astfel: ^n pri-mul an s-au produs riscurile (sinistrele) 3,11,17 [i 18, deci ^n total au ap\rut patru pagube, ... , ^n al patrulea an nu s-a produs nici un risc (sinistru, pagub\), ^n al cincilea an s-au produs riscurile (sinistrele) 11 [i 17, deci ^n total au ap\rut dou\ pagube, [.a.m.d.

Riscurile j, $j = \overline{1,20}$ au fost presupuse de tip Bernoulli standard, ceea ce ^nseamn\ c\ riscul implicat de contractul j s-a considerat o variabil\ aleatoare X_j, c\reia i-am atribuit doar valori booleene; deci X_j ia valoarea 1, dac\ apare riscul de sinistru j pentru bunul supus asigur\rii, ^ntr-unul din anii de valabilitate a con-tractului j, respectiv valoarea 0, dac\ nu apare riscul de sinistru j pentru bunul supus asigur\rii ^ntr-unul din anii de va-labilitate a contractului j; s-a notat cu θ_j probabilitatea cu care X_j ia valoarea 1 (unde $j=\overline{1,20}$). Prin urmare, variabila aleatoare X_j are reparti]ia:

$$X_j : \begin{pmatrix} 0 & 1 \\ 1-q_j & q_j \end{pmatrix}, \quad \text{unde} \quad j=\overline{1,20} \quad \text{fixat.}$$

Trecut statistic al riscului X_j, consta ^n variabilele aleatoare observabile X_{j1}, X_{j2}, ..., X_{j10}; observa]iile X_{j1}, X_{j2},..., X_{j10} sunt variabilele aleatoare independente [i identic distribuite cu X_j, care au condus la valorile numerice concrete x_{j1}, x_{j2},..., x_{j10} din diagrama anterioara ($j=\overline{1,20}$).

Contractele din exemplul lui Norberg ar putea fi, de pild\, poli]e de asigurare contra avariilor auto cu tip de ma[in\ necunoscut, caz ^n care se poate presupune despre X_j c\ are reparti]ia de tip Bernoulli standard, cu $j=\overline{1,20}$. Pro-babilitatea ca (X_j=1) este diferit\ de la un tip de ma[in\ la altul; dac\ ma[ina este de tipul j (necunoscut) atunci aceas-t\ probabilitate este egal\ cu θ_j (la rândul ei necunoscut\), unde $j=\overline{1,20}$.

Parametrii de risc θ_j, $j=\overline{1,20}$ reprezint\ caracteristicile de risc necunoscute ale poli]elor; pentru cazul particular consi-derat, aceste caracteristici sunt legate de tipul ma[inii. ~ntrebarea care se pune ^n leg\tur\ cu exemplul lui Norberg este urm\toarea: *sunt cele 20 de riscuri omo-gene sau nu?* R\spunsul poate fi dat de statistica clasic\.

Ipoteza omogenit\]ii riscurilor se formu-leaz\ astfel: $\theta_1=\theta_2=...=\theta_{20}$ (adic\ parame-trii abstrac]i θ_j pentru fiecare risc j, $j=\overline{1,20}$ sunt aceea[i, nu difer\ semni-ficativ din punct de vedere statistic, ci doar ^ntâmpl\tor). Not\m ipoteza pe care o avem de verificat cu H_0 [i o numim ipoteza nul\ sau fundamental\.

Ipoteza contrar\ ipotezei H_0 se formu-leaz\ astfel: $\theta_1\neq\theta_2\neq...\neq\theta_{20}$ (adic\ parame-trii abstrac]i θ_j pentru fiecare risc j, $j=\overline{1,20}$ nu sunt aceea[i; ei difer\ sem-nificativ din punct de vedere statistic) [i se noteaz\ cu H_1, numindu-se ipoteza alternativ\ ipotezei H_0.

Prin urmare, avem de testat cuplul de ipoteze (H_0, H_1), unde: H_0: $\theta_1=\theta_2=...=\theta_{20}$, iar H_1: $\theta_1\neq\theta_2\neq...\neq\theta_{20}$.

Acceptarea ipotezei nule H_0 ^nseamn\ c\ riscurile implicate de cele 20 contracte sunt omogene sau c\ parametrii abstrac]i θ_j ai fiec\rui risc j, $j=\overline{1,20}$, sunt ace-ea[i, cu alte cuvinte, cele 20 de con-tracte pot fi grupate ^mpreun\ ^ntr-un portofoliu omogen.

Respingerea ipotezei nule ^nseamn\ c\ riscurile implicate de cele 20 de con-tracte nu sunt omogene sau c\ parametrii abstrac]i θ_j pentru fiecare risc j, $j=\overline{1,20}$, nu sunt aceea[i; cu alte cuvin-te, portofoliul format din cele 20 con-tracte nu poate fi considerat omogen (este eterogen), datorit\ realiz\rilor diferi-te ale parametrului de risc q_j pentru fie-care contract (unde $j=\overline{1,20}$).

Verificarea perechii de ipoteze (H_0, H_1) se face la un prag de semnifica]ie α>0 foarte mic, dat [i pe baza testului statis-tic clasic *hi-p\trat*. ~n cele ce urmeaz\, ne ocup\m de determinarea statisticii testului.

Se [tie c\ estima]ia probabilit\]ii θ_j din reparti]ia lui X_j este frecven]a relativ\ a evenimentului A_j=("apari]ia riscului de sinistru j pentru bunul supus asigur\rii ^ntr-unul din anii de valabilitate ai con-tractului j"), ^n cei zece ani, cât a durat studiul statistic asupra variabilei aleatoare X_j (unde $j=\overline{1,20}$). ~ntr-adev\r, aplicarea metodei verosimilit\]ii maxime conduce la estima]ia:

$$\hat{\theta}_j = \frac{1}{n}\sum_{r=1}^{n} x_{jr} \overset{not}{=} \mathrm{f}_n(A_j), \quad (n=10; \quad j=\overline{1,20} \text{ fi-xat}),$$

a lui θ_j.

~n continuare acord\m estimatorului $\hat{\theta}_j$ interpretarea de variabil\ aleatoare, adic\ se consider\:

$$\hat{\theta}_j = \frac{1}{n}\sum_{r=1}^{n} X_{jr} \overset{not}{=} \mathrm{f}_n(A_j) \quad (n=10; \quad j=\overline{1,20}$$

fixat). Variabila aleatoare $f_n(A_j)$ – vezi expresia anterioara – desemneaz\ frec-ven]a relativ\ a evenimentului A_j ^n cei n(=10) ani. Frecven]a absolut\ a eveni-mentului A_j ^n cei n(=10) ani este notat\ $K_n(Aj)$ [i este egal\ cu $\sum_{r=1}^{n} X_{jr}$; ea are semnifica]ia num\rului de realiz\ri ale evenimentului A_j ^n cei zece ani. Evident, $K_n(A_j)$ are o reparti]ie binomial\ de parametrii n [i θ_j; scriem acest lucru astfel: $K_n(A_j) \in B(n, \theta_j)$ (n=10; $j = \overline{1,20}$ fixat).

Atragem aten]ia asupra faptului c\ *n* nu reprezint\ ^n mod necesar num\rul de ani ^n care contractul j a fost observat, ci poate reprezenta, de asemenea, num\-rul total al anilor de observa]ie pentru un grup de contracte, toate având θ_j drept parametru.

A[adar, dac\ ipoteza H_0 este adev\rat\, atunci $\theta_1=\theta_2=...=\theta_{20}$ [i putem considera, ^n baza observa]iei anterioare, c\ num\rul observa]iilor efectuate asupra contractului j este egal cu 10x20=200, ceea ce ^nseamn\ c\ dispunem de un num\r foarte mare de observa]ii pentru contractul j. Deci, conform teoremei Moivre-Laplace: $K_n(A_j) \underset{\approx}{\in} N\left(n\theta, \sqrt{n\theta(1-\theta)}\right)$, ($j=\overline{1,20}$ fixat), unde θ este valoarea comun\ necunoscut\ a celor 20 de pro-babilit\]i $\theta_1, \theta_2, ..., \theta_{20}$ din ipoteza H_0, pe care o estim\m cu m\rimea:

$$\hat{\theta} = \frac{\hat{\theta}_1 + \hat{\theta}_2 ... + \hat{\theta}_{20}}{20}.$$

Deoarece:

$$K_n(A_j) \underset{\approx}{\in} N\left(n\hat{\theta}, \sqrt{n\hat{\theta}\left(1-\hat{\theta}\right)}\right) (j=\overline{1,20} \quad \text{fixat}),$$

rezult\ c\:

$$\hat{\theta}_j = f_n(A_j) = \frac{K_n(A_j)}{n} \underset{\approx}{\in} N\left(\hat{\theta}, \sqrt{\frac{\hat{\theta}\left(1-\hat{\theta}\right)}{n}}\right) (j=\overline{1,20}$$

fixat), astfel ^ncât variabila aleatoare

$$\chi^2 = \frac{\sum_{j=1}^{20}\left(\hat{\theta}_j - \hat{\theta}\right)^2}{\frac{\hat{\theta}(1-\hat{\theta})}{10}}$$

are o reparti]ie hi-p\trat cu (20-1)=19 grade de libertate, fapt ce justific\ notarea acestei variabile aleatoa-re cu litera "χ^2". ~n concluzie, statistica pe baza c\reia test\m ipoteza nul\ H_0 este χ^2.

Testul "*hi-p\trat*" respinge ipoteza nul\ H_0 a omogenit\]ii riscurilor, adic\ accep-t\ H_1, ceea ce ^nseamn\ c\ portofoliul de 20 contracte nu poate fi considerat omogen, deci este eterogen (datorit\ realiz\rilor diferite ale parametrului de risc θ_j ($j=\overline{1,20}$) pentru fiecare contract $\leftarrow$ vezi ipoteza H_1), dac\:

$$\chi^2_{calc} = \frac{\sum_{j=1}^{20}\left(\hat{\theta}_j - \hat{\theta}\right)^2}{\frac{\hat{\theta}\left(1-\hat{\theta}\right)}{10}} \quad \text{unde} \quad \hat{\theta}_j = \frac{\sum_{r=1}^{n} x_{jr}}{n}$$

($j=\overline{1,20}$), dep\[e[te un anumit nivel (prag, limit\) c: $\chi^2_{calc} > c$.

Num\rul c se determin\ din defini]ia pragului de semnifica]ie α, astfel: $\alpha = P\left(\chi^2 > c\right)$ sau $1-\alpha = 1 - P\left(\chi^2 > c\right)$ sau ^nc\ $1-\alpha = P\left(\chi^2 \le c\right)$, de unde deducem c\ $c = \chi^2_{1-\alpha;19}$.

Pentru valorile lui $\chi^2{}_{calc}$ din intervalul $[c, +\infty)$, ipoteza nul\ H_0 se accept\, deci se poate presupune c\ cele 20 de riscuri au acela[i parametru. Prin urmare, are sens gruparea ^mpreun\ a contrac-telor, deoarece sunt caracterizate prin parametrul de risc egal (comun).

~n cazul exemplului lui Norberg, $\chi^2_{calc} = 49{,}16 \left(\hat{\theta} = 0{,}145\right)$ [i dac\ lu\m α=0,05 atunci $\chi^2_{1-\alpha;19} = \chi^2_{1-0{,}05;19} = \chi^2_{0{,}95;19} = 30{,}1$; ^ntrucât $\chi^2_{calc} > \chi^2_{1-\alpha;19}$, conchidem c\ ipo-teza omogenit\]ii riscurilor este fals\, adic\ portofoliul de 20 contracte este eterogen.

Bibliografie

1. Goovaerts, M.J., Kaas, R., Van Heerwaarden, A.E., Bauwelinckx, T. (1990); Insurance Series, volume 3, Effective Actuarial Methods, University of Amsterdam, The Netherlands.
2. Pentikäinen, T., Daykin, C.D., Pesonen, M. (1990); Practical Risk Theory for Actuaries, Université Pieré et Marie Curie.
3. Sundt, B. (1984); An Introduction to Non-Life Insurance Mathematisc, Veröffentlichungen des Instituts für Versicherungswissenschaft der Universität Mannheim Band 28 (V.V.W Karlsruhe).