PRACE NAUKOWE UNIWERSYTETU EKONOMICZNEGO WE WROCŁAWIU
Nr 1 (1201) 2008
Statystyka aktuarialna – teoria i praktyka

Anna Nikodem
Akademia Ekonomiczna we Wrocławiu

APROKSYMACJA ZAGREGOWANEJ WIELKOŚCI SZKÓD W ZALEŻNYM INDYWIDUALNYM MODELU RYZYKA

1. Zależny indywidualny model ryzyka

W klasycznym indywidualnym modelu ryzyka zakłada się, że wypłaty generowane z portfela polis ubezpieczeniowych są niezależne. Jest to jednak silne założenie w przypadku, gdy w skład tego portfela wchodzą polisy wykupione przez małżeństwa. Zależność spowodowana jest chociażby tym, że mąż i żona są narażeni na to samo ryzyko. Rozważmy zatem portfel jednorocznych polis na życie z wypłatą na koniec roku w wysokości 1. W skład tego portfela wchodzą $2m$ polisy wykupione przez m małżeństw. Dla tak określonego zależnego indywidualnego modelu ryzyka, zagregowana wielkość wypłaty definiuje się następująco:

$$S = \sum_{i=1}^{m} \widetilde{X}_i + \sum_{i=m+1}^{n} X_i, \tag{1}$$

gdzie $\widetilde{X}_i = \left(X_i + X_{i'}\right)$ określa wysokość świadczenia z dwóch polis wykupionych przez i-te małżeństwo. Jeśli przez q_i oznaczymy prawdopodobieństwo wystąpienia wypłaty z i-tej polisy, to rozkład zmiennej losowej X_i i rozkład zmiennej $X_{i'}$ pod warunkiem, że wypłata z polisy nastąpi dla $i = 1, 2, \ldots, m+n$, mają odpowiednio postać

$$P\left(X_i = x\right) = \begin{cases} p_i = 1 - q_i, & x = 0, \\ q_i, & x = 1, \end{cases} \quad f(x) = \begin{cases} 1, & x = 1, \\ 0, & \text{poza tym.} \end{cases} \tag{2}$$

W przypadku dwuwymiarowej zmiennej losowej $\widetilde{X}_i$ dla $i = 1, 2, \ldots, m$, gdy prawdopodobieństwa łączne o rozkładach brzegowych określonych w (2) oznaczymy przez:

$$\begin{gathered} p_i^{[0]} = P(X_i = 0, X_{i'} = 0),\ p_i^{[2]} = P(X_i = 1, X_{i'} = 1), \\ p_i^{[1]} = P(X_i = 1, X_{i'} = 0) + P(X_i = 0, X_{i'} = 1), \end{gathered} \tag{3}$$

to rozkład $\widetilde{X}_i$ oraz rozkład zmiennej $\widetilde{X}_{i'}$ pod warunkiem że nastąpi co najmniej jedna wypłata przybiorą odpowiednio postać:

$$P\left(\widetilde{X}_i = x\right) = \begin{cases} p_i^{[0]}, & x = 0, \\ p_i^{[1]}, & x = 1, \\ p_i^{[2]}, & x = 2, \end{cases} \quad \widetilde{f}(x) = \begin{cases} \dfrac{p_i^{[1]}}{1 - p_i^{[0]}}, & x = 1, \\ \dfrac{p_i^{[2]}}{1 - p_i^{[0]}}, & x = 2, \\ 0, & \text{poza tym.} \end{cases} \tag{4}$$

W praktyce zależność między wypłatami z polis małżeństw jest dodatnia. Stąd będzie rozważany wyłącznie dodatni współczynnik korelacji. Dla takiej zależności zachodzi $\bar{p}_i^{[k]} \le p_i^{[k]} \le \breve{p}_i^{[k]}$ (zob. [4]), gdzie $\bar{p}_i^{[k]}$ oznacza prawdopodobieństwo, gdy zmienne X_i oraz $X_{i'}$ są niezależne, a $\breve{p}_i^{[2]} = \min\{q_i, q_{i'}\}$, gdy zmienne X_i, $X_{i'}$ są współmonotoniczne (ang. *comonotonic*). Jeśli zmienne są współmonotoniczne, oznacza to, że jeśli młodsza osoba z małżeństwa umrze, to oboje ubezpieczeni umrą w okresie ubezpieczenia. Rozważmy parę $(X_i, X_{i'})$ i niech T_x oznacza przyszły czas życia x-latka. Dla $i, i' = 1, 2, \ldots, m$ współczynnik korelacji spełnia warunek

$$r(X_i, X_{i'}) = r\left(I\left(T_{x_i} \le 1\right), I\left(T_{x_{i'}} \le 1\right)\right),$$

gdzie I jest indykatorem (zob. [4]). Jeśli $r(X_i, X_{i'}) \ge 0$, wtedy dla każdego $i = 1, 2, \ldots, m$ i wartości $s_{ii'} \in [0, 1]$, takiego że

$$r\left(I\left(T_{x_i} \le 1\right), I\left(T_{x_{i'}} \le 1\right)\right) = s_{ii'} r\left(I\left(\breve{T}_{x_i} \le 1\right), I\left(\breve{T}_{x_{i'}} \le 1\right)\right)$$

definiuje się

$$p_i^{[0]} = s_{ii'} \breve{p}_i^{[0]} + (1 - s_{ii'}) \bar{p}_i^{[0]},\ p_i^{[1]} = s_{ii'} \breve{p}_i^{[1]} + (1 - s_{ii'}) \bar{p}_i^{[1]}, \tag{5}$$

$$p_i^{[2]} = s_{ii'} \breve{p}_i^{[2]} + (1 - s_{ii'}) \overline{p}_i^{[2]}.$$

W dalszej części artykułu będzie zakładane, że współczynnik $s_{ii'}$ jest dany. Zatem łączne prawdopodobieństwa dwuwymiarowej zmiennej losowej $\widetilde{X}_i$ są również znane, stąd będzie można wyznaczyć rozkład zagregowanej wielkości szkód S. W podpunkcie drugim zostaną przedstawione metody wyznaczania tego rozkładu.

2. Wyznaczanie rozkładu zagregowanej wielkości szkód

2.1. Metoda rekurencyjna wyznaczania rozkładu zagregowanych szkód

Metoda rekurencyjna wyznacza dokładny rozkład zagregowanej wielkości szkód wypłacanej z portfela. Funkcja tworząca prawdopodobieństwa zmiennej S określonej w (1) ma postać:

$$G_S(t) = \mathrm{E}\left[t^S\right] = \prod_{i=1}^{m} \left(p_i^{[0]} + \left(1 - p_i^{[0]}\right) K_i(t)\right) \cdot \prod_{i=m+1}^{n} \left(p_i + (1 - p_i) H_i(t)\right) \tag{6}$$

gdzie $K_i(t)$ oraz $H_i(t)$ są odpowiednio funkcjami tworzącymi warunkowych zmiennych losowych o rozkładach $f_i(x)$ oraz $\widetilde{f}_i(x)$. Wiedząc, że

$$P(S = s) = \frac{G^{(s)}(0)}{s!},$$

otrzymujemy bezpośrednio z (6), wstawiając $t = 0$, prawdopodobieństwo

$$P(S = 0) = \prod_{i=1}^{m} p_i^{[0]} \prod_{i=m+1}^{n} p_i .$$

W celu wyznaczenia prawdopodobieństwa w pozostałych punktach obliczamy pochodną funkcji tworzącej

$$G_S'(t) = \underbrace{\sum_{i=1}^{m} \left(\frac{\left(1 - p_i^{[0]}\right) K_i'(t) G_S(t)}{p_i^{[0]} + \left(1 - p_i^{[0]}\right) K_i(t)} \right)}_{V_i(t)} + \underbrace{\sum_{i=m+1}^{n} \left(\frac{(1 - p_i) H_i'(t) G_S(t)}{p_i + (1 - p_i) H_i(t)} \right)}_{W_i(t)}.$$

Niech $G_S'(t) = V_i(t) + W_i(t)$, gdzie funkcje $V_i(t)$ oraz $W_i(t)$ zdefiniowane są następująco

$$V_i(t)=\sum_{x=0}^{\infty} v_i(x+1)t^x=\left(\frac{\left(1-p_i^{[0]}\right)K_i'(t)G_S(t)}{p_i^{[0]}+\left(1-p_i^{[0]}\right)K_i(t)}\right),$$

$$W_i(t)=\sum_{x=0}^{\infty} w_i(x+1)t^x=\left(\frac{(1-p_i)H_i'(t)G_S(t)}{p_i+(1-p_i)H_i(t)}\right).$$

Pochodna rzędu s funkcji $G_S(t)$ ma postać:

$$G_S^{(s)}(t)=\sum_{i=1}^{m} V_i^{(s-1)}(t)+\sum_{i=m+1}^{n} W_i^{(s-1)}(t),$$

gdzie

$$V_i^{(s-1)}(t)=\sum_{x=s-1}^{\infty} v_i(x+1)s!t^{x-(s-1)},\ W_i^{(s-1)}(t)=\sum_{x=s-1}^{\infty} w_i(x+1)s!t^{x-(s-1)}.$$

Zatem prawdopodobieństwo w punkcie s wyznacza się ze wzoru

$$sP(S=s)=\sum_{i=1}^{m} v_i(s)+\sum_{i=m+1}^{n} w_i(s).$$

Wzór rekurencyjny na funkcję $v_i(s)$ wyznaczamy, obliczając pochodną rzędu $(s-1)$ funkcji

$$V_i(t)=\frac{1-p_i^{[0]}}{p_i^{[0]}}\left[K_i'(t)G_S(t)-K_i(t)V_i(t)\right]$$

i podstawiając $t=0$ otrzymamy

$$v_i(s)=\frac{1-p_i^{[0]}}{p_i^{[0]}}\sum_{x=1}^{2}\tilde{f}_i(x)\left(x\cdot P(S=s-x)-v_i(s-x)\right).$$

Podobnie uzyskamy rekurencję dla funkcji

$$w_i(s)=\frac{1-p_i}{p_i}\left(P(S=s-1)-w_i(s-1)\right).$$

Zatem wzór rekurencyjny wyznaczający dokładny rozkład zmiennej S ma postać:

$$P(S=0)=\prod_{i=1}^{m} p_i^{[0]}\prod_{i=m+1}^{n} p_i, \tag{7}$$

$$sP(S=s)=\sum_{i=1}^{m} v_i(s)+\sum_{i=m+1}^{n} w_i(s) \text{ dla } s=1,2,\ldots,M,$$

gdzie:

$$v_i(s)=\frac{1-p_i^{[0]}}{p_i^{[0]}}\sum_{x=1}^{2}\tilde{f}_i(x)\big(x\cdot P(S=s-x)-v_i(s-x)\big), \tag{8}$$

$$w_i(s)=\frac{1-p_i}{p_i}\big(P(S=s-1)-w_i(s-1)\big).$$

a M jest maksymalną wielkością zagregowanej wypłaty z portfela.

Przykład 1

Rozważmy portfel zawierający 500 polis. W tabeli 1. przedstawione są prawdopodobieństwa wypłaty (p_i^{M}; $p_i^{Ż}$) dla małżonków oraz liczba małżeństw. Łącznie, 230 małżeństw wykupiło 460 polis. Prawdopodobieństwa wypłaty dla pozostałych ubezpieczonych przedstawione są w tab. 2.

Tabela 1. Struktura portfela w wypadku małżeństw

Prawdopodobieństwa wypłaty dla małżonków	Liczba małżeństw
0,02; 0,02	90
0,02; 0,03	80
0,03; 0,03	60
Łączna liczba polis	460

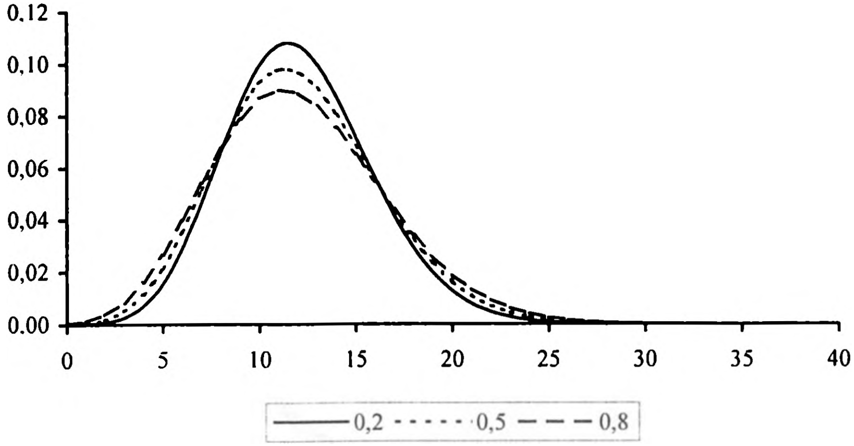
Tabela 2. Struktura portfela dla pozostałych ubezpieczonych

Prawdopodobieństwo wypłaty	Liczba polis
0,02	30
0,03	10
Łączna liczba polis	40

Wartość oczekiwana i wariancja zagregowanej wielkości wypłat z tego portfela wynoszą w zależności od współczynnika $s_{ii'}$:

$$s_{ii'}=0,\quad E[S]=12{,}1,\ V[S]=11{,}795,$$
$$s_{ii'}=0{,}2,\quad E[S]=12{,}1,\ V[S]=13{,}819,$$
$$s_{ii'}=0{,}5,\quad E[S]=12{,}1,\ V[S]=16{,}857,$$
$$s_{ii'}=0{,}8,\quad E[S]=12{,}1,\ V[S]=19{,}894,$$
$$s_{ii'}=1,\quad E[S]=12{,}1,\ V[S]=21{,}191.$$

Można zauważyć, że gdy zakłada się, iż wszystkie zmienne są niezależne (przypadek, gdy $s_{ii'}=0$) to wariancja jest najmniejsza. Największa wariancja jest wtedy, gdy zmienne dotyczące małżeństw są współmonotoniczne. Dokładny rozkład zagregowanej wielkości wypłat przy założeniu współczynnika $s_{ii'}=0{,}2$; 0,5; 0,8 przedstawia rys. 1.



Rys. 1. Dokładny rozkład zagregowanej wielkości szkód

2.2. Aproksymacja rozkładu zagregowanych szkód

Aproksymacje stosuje się w celu skrócenia czasu potrzebnego na wyznaczenie dokładnego rozkładu wypłat S. De Pril zaproponował algorytm dla indywidualnego modelu ryzyka (zob. [1; 2]). Korzystając z tego wzoru, otrzymujemy bezpośrednio wzór na aproksymację rozkładu zagregowanych szkód określonych w (1):

$$\begin{gathered} P_{AP}(S=0)=\prod_{i=1}^{m} p_i^{[0]} \prod_{i=m+1}^{n} p_i, \\ s \cdot P_{AP}(S=s)=\sum_{j=1}^{2} A(j,1) \sum_{x=1}^{M_j} x f_{ij} P_{AP}(S=s-x) \text{ dla } x=1,2,\ldots,M \end{gathered} \tag{9}$$

gdzie $A(1,1)=\sum_{i=1}^{m}\left(\frac{1-p_i^{[0]}}{p_i^{[0]}}\right)$, $A(2,1)=\sum_{i=m+1}^{n}\left(\frac{1-p_i}{p_i}\right)$, $f_{i1}(x)=\tilde{f}_i(x)$, $f_{i2}(x)=$ $=f_i(x)$, $M_1=2$, $M_2=1$. Wstawiając do wzoru (9) za rozkład $f_i(x)$ i $\tilde{f}_i(x)$ odpowiednio (2) i (4), otrzymujemy rekurencyjną postać aproksymacji zmiennej S

$$P_{AP}(S=0)=\prod_{i=1}^{m} p_i^{[0]} \prod_{i=m+1}^{n} p_i,$$

$$s\cdot P_{AP}(S=s)=\sum_{i=1}^{m}\left[\frac{p_i^{[1]}}{p_i^{[0]}}P_{AP}(S=s-1)+2\frac{p_i^{[2]}}{p_i^{[0]}}P_{AP}(S=s-2)\right]+ \tag{10}$$
$$+\sum_{i=m+1}^{n}\left[\frac{1-p_i}{p_i}\right]P_{AP}(S=s-1),\text{ dla } x=1,2,\ldots,M.$$

Błąd tej aproksymacji zdefiniowanej jako $\sum_{s=0}^{M}\left|P(S=s)-P_{AP}(S=s)\right|$ nie przekracza $e^{\sigma}-1$, tzn.

$$\sum_{s=0}^{M}\left|P(S=s)-P_{AP}(S=s)\right|\le e^{\sigma}-1,$$

gdzie

$$\sigma=\frac{1}{2}\left[\sum_{i=1}^{m}\frac{p_i^{[0]}}{2p_i^{[0]}-1}\left(\frac{1-p_i^{[0]}}{p_i^{[0]}}\right)^2+\sum_{i=m+1}^{n}\frac{p_i}{2p_i-1}\left(\frac{1-p_i}{p_i}\right)^2\right].$$

Rozkład zagregowanej wielkości szkód może być aproksymowany złożonym rozkładem Poissona (zob. [3])

$$F_{cP}(s)=\sum_{n=0}^{s}P(N=n)F^{*n}(s)\text{ dla s}=0,1,\ldots, \tag{11}$$

gdzie N ma rozkład Poissona z parametrem $\lambda=\sum_{i=1}^{n}q_i$, a $F^{*n}(s)$ oznacza n-splot dystrybuant określonych następująco $F(s)=\frac{1}{\lambda}\sum_{i=1}^{n}q_iP\left(X_i\le s\middle|X>0\right)$.

Błąd aproksymacji definiowanej jako $\sup_J\left|P(S\in J)-P\left(S_{cP}\in J\right)\right|$ wynosi

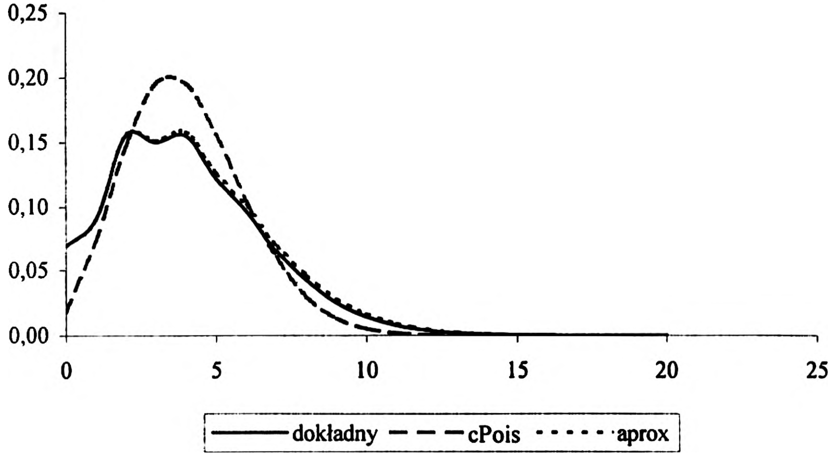
$$d\left(F_{ind},F_{cP}\right)=\sup_J\left|P(S\in J)-P\left(S_{cP}\in J\right)\right|=\frac{1}{\lambda}\left\{\sum_{i=1}^{m}\left[\left(q_i+q_{i'}\right)^2\right]+\sum_{i=m+1}^{n}q_i^2\right\},$$

gdzie

$$\lambda=\sum_{i=1}^{m}\left(q_i+q_{i'}\right)+\sum_{i=m+1}^{n}q_i.$$

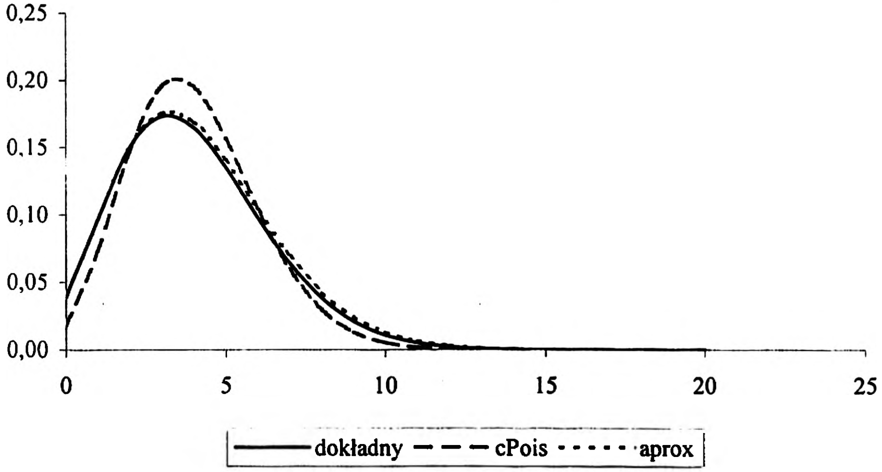
Przykład 2

a) Rozważmy portfel składający się z 200 identycznych polis, w tym 170 jest wykupionych przez 85 małżeństw. Prawdopodobieństwo śmierci wynosi 0,02. Współczynnik $s_{ii'} = 0{,}8$.



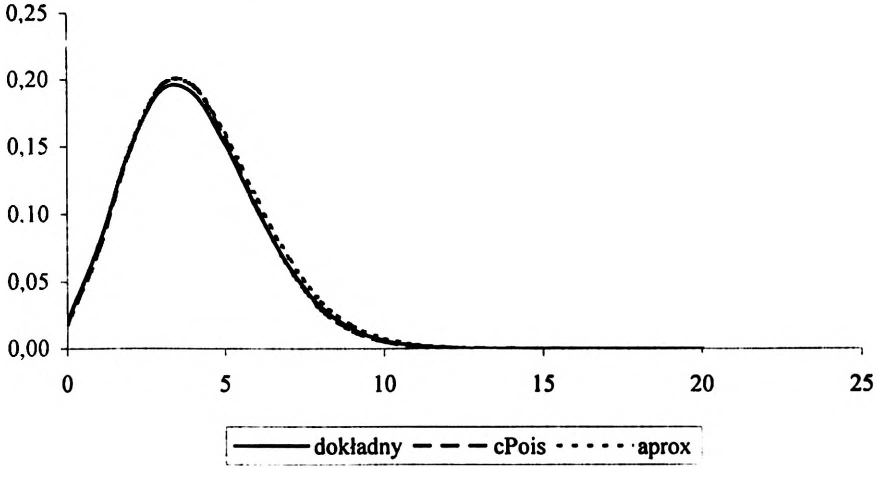
Rys. 2. Dokładny rozkład zagregowanej wielkości szkód oraz aproksymacje

b) Portfel składa się z 200 polis, w tym liczba małżeństw wynosi 50. Prawdopodobieństwo śmierci wynosi 0,02. Współczynnik $s_{ii'} = 0{,}8$.



Rys. 3. Dokładny rozkład zagregowanej wielkości szkód oraz aproksymacje

c) Portfel składa się z 200 polis, w którym prawdopodobieństwo śmierci wynosi 0,02. Współczynnik $s_{ii'} = 0{,}8$, liczba małżeństw $m = 10$.



Rys. 4. Dokładny rozkład zagregowanej wielkości szkód oraz aproksymacje

Na rysunkach 2, 3 i 4 przedstawiono dokładny rozkład zagregowanej wielkości szkód oraz rozkłady aproksymujące wyznaczone zgodnie z (10) i (11). Można zauważyć, że aproksymacja złożonym rozkładem Poissona, dla tak określonego portfela, może być stosowana, gdy w portfelu jest niewielka liczba małżeństw (rys. 4), a aproksymacja wyznaczana zgodnie z (10) była dokładniejsza we wszystkich przypadkach.

Literatura

[1] De Pril N., *The aggregate claims distribution in the individual model with arbitrary positive claims*, ASTIN BULLETIN 1989 vol. 19, no. 1, s. 9-24.

[2] Dhaene J., Vandebroek M., *Recursions for the indivudual model*, „Insurance: Mathematics and Economics" 1995, no. 16, s. 31-38.

[3] Goovaerts M.J., Dhaene J., *The compound Poisson approximation for a portfolio of dependent risks*, „Insurance: Mathematics and Economics" 1996, no. 18, s. 81-85.

[4] Ribas C., Marin-Solano J., Alegre A., *On the computation of the aggregate claims distribution in the individual life model with bivariate dependencies*, „Insurance: Mathematics and Economics" 2003, no. 32, s. 201-215.

APPROXIMATION OF THE AGGREGATE CLAIM DISTRIBUTION IN THE INDIVIDUAL RISK MODEL WITH DEPENDENT RISKS

Summary

In the classical individual risk model it is assumed that the risks are mutually independent. In this paper, a portfolio consisting of some independent pairs of dependent risks is considered. The method of computing the aggregate claim distribution for this portfolio is also presented.