A C T A   U N I V E R S I T A T I S   L O D Z I E N S I S
FOLIA OECONOMICA 296, 2013

Anna Szymańska[*]

ROZKŁADY LICZBY SZKÓD W UBEZPIECZENIACH KOMUNIKACYJNYCH OC

1. WPROWADZENIE

Zgodnie z *Ustawą z dnia 28 lipca 1990 r. o działalności ubezpieczeniowej*[1] ubezpieczenia komunikacyjne należą do działu II ubezpieczeń majątkowych i pozostałych osobowych. Przy czym ubezpieczenia komunikacyjne odpowiedzialności cywilnej OC stanowią grupę 10. tego działu. W Polsce w 2012 r. składka przypisana brutto z tytułu ubezpieczeń komunikacyjnych stanowiła około 56% składki ubezpieczeń majątkowych i osobowych, z czego 34% to ubezpieczenia grupy 10.[2] Ponadto składki z tytułu ubezpieczeń komunikacyjnych OC stanowiły średnio 39% składek portfela ubezpieczeń majątkowych każdego towarzystwa ubezpieczeniowego. Od roku 2008 wynik techniczny[3] w ubezpieczeniach grupy 10. na rynku jest ujemny, co wskazuje na potrzebę zmian taryf w tych ubezpieczeniach. Zadaniem ubezpieczyciela jest utrzymanie równowagi finansowej składek i świadczeń poprzez ustalenie wielkości funduszu ubezpieczeniowego, potrzebnego do wywiązania się z przyjętych zobowiązań oraz właściwe rozłożenie pomiędzy ubezpieczanych kosztów tworzenia tego funduszu. Wielkość funduszu ustala się na podstawie przewidywanych liczby i wielkości szkód. Rozłożenie kosztów tworzenia funduszu polega na różnicowaniu wysokości składek ubezpieczeniowych. Pełna równowaga operacji finansowych to stan, w którym składki pokrywają wszystkie koszty ubezpieczyciela.

W ubezpieczeniach komunikacyjnych zakłada się, że rozkład liczby szkód w portfelu dla każdego ubezpieczonego jest tego samego typu. Zmienna losowa opisująca liczbę szkód w jednostce czasu jest zmienną losową skokową. W lite-

[*] Dr, Katedra Metod Statystycznych, Uniwersytet Łódzki.

[1] *Ustawa z dnia 28 lipca 1990 r. o działalności ubezpieczeniowej*, DzU 1990, nr 59, poz. 344.

[2] *Raport o stanie sektora ubezpieczeń po I półroczu 2012 roku*, KNF, Warszawa 2012, www.knf.gov.pl.

[3] Wynik techniczny rozumiany jak w *Rozporządzeniu Ministra Finansów z dnia 28 grudnia 2009 r. w sprawie szczególnych zasad rachunkowości zakładów ubezpieczeń i zakładów reasekuracji*, DzU 2009, nr 226, poz. 1825.

raturze aktuarialnej dotyczącej metod wyznaczania składek w ubezpieczeniach komunikacyjnych najczęściej do opisu liczby szkód w danym okresie czasu stosuje się rozkład dwumianowy, Poissona lub ujemny dwumianowy. Rozkład liczby szkód opisywano za pomocą rozkładu ujemnego dwumianowego m. in. w pracach Lemaire[4], Ibiwoye, Adeleke i Aduloju[5]. Zastosowanie rozkładu Poissona – logarytmiczno normalnego przedstawia np. praca Aitchison i Ho[6]. W pracach Tremblay[7], Willmot[8] do modelowania liczby szkód w ubezpieczeniach komunikacyjnych wykorzystano rozkład Poissona – odwrotny normalny, a w pracy Sarabia i Gomez[9] rozkład Poissona – beta. Rozkład Poissona – gamma–gamma do opisu liczby szkód użyto np. w pracy Gomez, Sarabia, Perez i Vazquez[10]. Uogólniony rozkład Poissona – gamma zastosowano np. w opracowaniu Sarabia, Gomez i Vazquez[11], rozkład dwumianowy – beta np. w pracy Griffiths[12], ujemny dwumianowy – Pareto np. w pracy Shengwang, Yuan i Whitmore[13], geometryczny np. w opracowaniu Mert i Saykan[14]. Rozkład Neymana typu A oraz uogólniony rozkład Poissona–Pascala prezentuje praca Panjer i Willmot[15].

[4] J. Lemaire, *Automobil Insurance. Actuarial Models*, Kluwer, Boston 1985; J. Lemaire, *Bonus-Malus Systems in Automobile Insurance,* Kluwer, Boston 1995.

[5] A. Ibiwoye, A. Adeleke, S. A. Aduloju, *Quest for Optimal Bonus-Malus in Automobile Insurance in Developing Economies: An Actuarial Perspective,* „International Business Research" 2011, vol. 4, no. 4, s. 74–83.

[6] J. Aitchison, C. H. Ho, *The Multivariate Poisson – Lognormal Distribution*, „Biometrika" 1989, vol. 76, s. 643–653.

[7] L. Tremblay, *Using the Poisson Inverse Gaussian in bonus-malus systems*, „ASTIN Bulletin" 1992, vol. 22, no.1, s. 97–106.

[8] G. E. Willmot, *The Poisson-inverse Gaussian distributionas an alternative to the negative binominal*, „Scandinavian Actuarial Journal" 1987, s. 113–127.

[9] J. M. Sarabia, E. Gomez-Deniz, *Distribuciones multivariantes Poisson – beta con aplicaciones a datos de seguros*, Investigaciones en Seguros y Gestion de Riesgos, Riesgo 2007, s. 15–26.

[10] E. Gomez-Deniz, J. M. Sarabia, J. M. Perez, F. Vazquez, *Using a Bayesian hierarchical model for fitting automobile claim frequency data*, „Communications in Statistics: Theory and Methods" 2008, vol. 37, s. 1425–1435.

[11] J. M. Sarabia, E. Gomez-Deniz, F. Vazquez-Polo, *On the Use of Conditional Models in Claim Count Distributions: An Application to Bonus-Malus Systems*, „ASTIN Bulletin" 2004, vol. 34, no. 1, s. 85–98.

[12] D. A. Griffiths, *Maximum Likelihood Estimation for the Beta-Binominal Distribution and Application to the Household Distribution of the Total Number of Cases of Disease*, „Biometrics" 1973, vol. 29, s. 637–648.

[13] M. Shengwang, W. Yuan, G. A. Whitmore, *Accounting for Individual Over-Dispersion in a Bonus-Malus Automobile Insurance System*, „ASTIN Bulletin" 1999, vol. 29, no. 2, s. 327–337.

[14] M. Mert, Y. Saykan, *On a Bonus-Malus System where the Claim Frequency Distribution is Geometric and the Claim Severity Distribution is Pareto*, „Hacettepe Journal of Mathematics and Statistics" 2005, vol. 34, s. 75–81.

[15] H. H. Panjer, G. E. Willmot, *Insurance risk models,* Society of Actuaries, Schaumburg 1992.

2. WYBRANE ROZKŁADY LICZBY SZKÓD

W niniejszej części pracy dokonano przeglądu rozkładów najczęściej stosowanych w literaturze aktuarialnej do opisu liczby szkód w ubezpieczeniach komunikacyjnych OC. Niech zmienna losowa *X* oznacza liczbę szkód z indywidualnej polisy lub portfela polis.

Rozkład dwumianowy Bernouliego opisany jest funkcją rozkładu prawdopodobieństwa:

$$P(X=k)=\binom{n}{k}p^k q^{n-k},\ gdzie\ \ k=0,1,...,n\ oraz\ \ \binom{n}{k}=\frac{n!}{k!(n-k)!} \tag{1}$$

Wartość oczekiwana i wariancja rozkładu wynoszą:

$$EX=np\ \text{ i }\ D^2X=npq,\ \text{gdzie}\ q=1-p.$$

Rozkład Poissona to rozkład o funkcji prawdopodobieństwa określonej wzorem:

$$P(X=k)=\exp(-\lambda)\frac{\lambda^k}{k!},\ k=0,1,... \tag{2}$$

z wartością oczekiwaną i wariancją postaci: $EX=\lambda$ i $D^2X=\lambda$.

Zmienna losowa *X* ma **rozkład ujemny dwumianowy (Polya)**, jeżeli jej funkcja rozkładu prawdopodobieństwa ma postać:

$$P(X=k)=\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)k!}\left(\frac{\beta}{1+\beta}\right)^{\alpha}\left(\frac{1}{1+\beta}\right)^{k} \tag{3}$$

Przyjmując $p=\dfrac{\beta}{1+\beta}$, $\beta>0$, funkcja rozkładu prawdopodobieństwa rozkładu ujemnego dwumianowego z parametrami α i p dana jest równaniem:

$$P(X=k)=\binom{\alpha+k-1}{k}p^{\alpha}(1-p)^k,\ k=0,1,2,...;\ \alpha>0 \tag{4}$$

gdzie: $\binom{\alpha+k-1}{k} = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\Gamma(k+1)} = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)k!}$.

Wartość oczekiwana i wariancja rozkładu ujemnego dwumianowego opisanego równaniami (3) i (4) odpowiednio wynoszą:

$$EX = \frac{\alpha(1-p)}{p} = \frac{\alpha}{\beta} \text{ oraz } D^2X = \frac{\alpha(1-p)}{p^2} = \frac{\alpha}{\beta}\left(1+\frac{1}{\beta}\right).$$

Rozkład ujemny dwumianowy z całkowitą wartością parametru α nazywany jest **rozkładem Pascala**. Dla $\alpha = 1$ rozkład ujemny dwumianowy nazywany jest **rozkładem geometrycznym**.

Niech zmienna losowa X ma rozkład Poissona z parametrem λ oraz niech parametr λ ma rozkład odwrotny normalny. Wówczas zmienna losowa X ma **rozkład Poissona – odwrotny normalny**. Funkcja rozkładu prawdopodobieństwa rozkładu Poissona – odwrotnego normalnego (Poisson – Inverse Gaussian) dana jest wzorem:

$$P(X=k) = \sqrt{\frac{2\alpha}{\pi}}\exp\left(\alpha\sqrt{1-\theta}\right)\frac{(\alpha\theta/2)^k}{k!}K_{k-1/2}(\alpha),\ k = 0,1,... \tag{5}$$

gdzie: $K_{k-1/2}(\alpha)$ jest zmodyfikowaną funkcją Bessela trzeciego rodzaju (dla dodatnich i rzeczywistych argumentów) postaci:

$$K_{k-1/2}(\alpha) = \sqrt{\frac{\pi}{2\alpha}}\exp(-\alpha)\left(\sum_{i=0}^{k-1}\frac{(k-1+i)!}{(k-1-i)!\,i!}(2\alpha)^{-i}\right),\ k = 1,2.... \tag{6}$$

Wartość oczekiwana i wariancja rozkładu są postaci:

$$EX = \frac{\alpha\theta}{2\sqrt{1-\theta}} \text{ i } D^2X = \frac{\alpha\theta(2-\theta)}{4(1-\theta)^{3/2}},\ 0 < \theta \le 1, \alpha > 0.$$

Rozkład można przedstawić, używając innej parametryzacji, za pomocą funkcji tworzącej prawdopodobieństwa postaci:

$$P(t) = \exp\left[\frac{\mu}{\beta}\left(1-\sqrt{1+2\beta(1-t)}\right)\right] \tag{7}$$

Wartość oczekiwana i wariancja rozkładu Poissona – odwrotnego normalnego mają wówczas postać: $EX = \mu$ i $D^2 X = \mu(1+\beta)$, a wartości prawdopodobieństw można obliczyć rekurencyjnie ze wzorów:

$$p_0 = P(X = 0) = \exp\left(\frac{\mu}{\beta}\left(1 - \sqrt{1+2\beta}\right)\right),$$

$$p_1 = P(X = 1) = \frac{\mu}{\sqrt{1+2\beta}} p_0, \tag{8}$$

$$p_k = P(X = k) = \frac{\beta}{1+2\beta}\frac{2k-3}{k} p_{k-1} + \frac{\mu^3}{1+2\beta}\frac{1}{k(k-1)} p_{k-2}, \quad k = 2,3,...$$

Ponadto zachodzi związek: $\lambda = \frac{\mu}{\beta}\left[(1+2\beta)^{\frac{1}{2}} - 1\right]$.

Niech zmienna losowa X ma rozkład Poissona z parametrem λ oraz niech parametr λ będzie zmienną losową o rozkładzie logarytmiczno-normalnym. Wówczas zmienna losowa X ma **rozkład Poissona – logarytmiczno normalny** z funkcją rozkładu prawdopodobieństwa daną równaniem:

$$P(X = k) = \frac{1}{\sigma\sqrt{2\pi}k!}\int_0^\infty \exp(-\lambda)\lambda^{k-1} \exp\left(-\frac{(\ln\lambda - \mu)^2}{2\sigma^2}\right)d\lambda, \tag{9}$$

$$k = 0,1,...,\mu \in R, \sigma > 0$$

Wartość oczekiwana i wariancja rozkładu mają postać:

$$EX = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

oraz $D^2 X = \exp(2\mu + 2\sigma^2) + \exp\left(\mu + \frac{1}{2}\sigma^2\right) - \exp(2\mu + \sigma^2)$,

gdzie μ i σ^2 są odpowiednio wartością oczekiwaną i wariancją zmiennej losowej Λ o rozkładzie logarytmiczno normalnym.

Funkcja rozkładu prawdopodobieństwa **rozkładu Poissona–Poissona (Neymana typu A)** wyraża się wzorem:

$$P(X=k)=\exp(-\lambda_1)\frac{\lambda_2^{\,k}}{k!}\sum_{n=0}^{\infty}\frac{n^k}{n!}(\lambda_1\exp(-\lambda_2))^n,\quad k=0,1,2,... \tag{10}$$

Wartość oczekiwana i wariancja rozkładu Neymana typu A wynoszą:

$$EX=\lambda_1\lambda_2 \text{ oraz } D^2X=\lambda_1\lambda_2(1+\lambda_2).$$

Prawdopodobieństwa rozkładu Neymana typu A można wyznaczyć rekurencyjnie za pomocą wzorów:

$$p_k=\frac{\lambda_1}{k}\sum_{j=1}^{k} jq_j p_{k-j},\quad k=1,2,...,$$

$$p_0=\exp\{-\lambda_1[1-\exp(-\lambda_2)]\}, \tag{11}$$

$$q_k=\frac{\lambda_2}{k}q_{k-1}.$$

Uogólniony rozkład Poisona–Pascala to rozkład opisywany za pomocą funkcji tworzącej momenty postaci:

$$M_X(t)=\exp\left\{\lambda\left[\frac{[1-\beta(t-1)]^{-\alpha}-(1+\beta)^{-\alpha}}{1-(1+\beta)^{-\alpha}}-1\right]\right\},\quad \alpha>-1,\lambda>0,\beta>0 \tag{12}$$

Wartość oczekiwana, wariancja oraz skośność rozkładu wynoszą odpowiednio:

$$EX=\mu_1=\lambda[1-(1+\beta)^{-\alpha}]^{-1}\alpha\beta,\ D^2X=\mu_2=\mu_1[1+(\alpha+1)\beta]$$

$$\text{oraz } \gamma=\mu_2^{-\frac{3}{2}}\left[3\mu_2-2\mu_1+\frac{\alpha+2}{\alpha+1}\frac{(\mu_2-\mu_1)^2}{\mu_1}\right].$$

Dla $\alpha > 0$ rozkład jest nazywany **rozkładem Poissona–Pascala**. Dla $\alpha = 1$ rozkład jest nazywany **rozkładem Polya–Aeppli**. Dla $\alpha = -0,5$ rozkład nazywa się rozkładem Poissona – odwrotnym normalnym. Wartości prawdopodobieństw uogólnionego rozkładu Poissona–Pascala można wyznaczyć ze wzorów rekurencyjnych postaci:

$$p_k = \frac{\lambda}{k}\sum_{j=1}^{k} j q_j p_{k-j}, \quad k = 1,2,\ldots,$$

$$p_0 = \exp(-\lambda),$$

$$q_k = \frac{k+\alpha-1}{k}\frac{\beta}{1+\beta}q_{k-1}, \quad k = 2,3,\ldots, \tag{13}$$

$$q_1 = \frac{\alpha}{(1+\beta)^{\alpha}-1}\frac{\beta}{1+\beta}.$$

3. WYBÓR ROZKŁADU LICZBY SZKÓD

Heilmann sugeruje wybór rozkładu liczby szkód w ubezpieczeniach komunikacyjnych odpowiedzialności cywilnej w zależności od relacji między wartością oczekiwaną i wariancją z próby[16] rozważając trzy rozkłady: dwumianowy, Poissona oraz ujemny dwumianowy. Według pracy Panjer i Willmot[17] wstępny wybór teoretycznego rozkładu liczby szkód może być oparty na obliczonych momentach z próby oraz współczynnikach częstości.

Niech $X_1, X_2, \ldots, X_n$ będzie próbą prostą oraz $X_1, X_2, \ldots, X_n$ niezależnymi zmiennymi losowymi o tym samym rozkładzie skokowym. Momenty zwykłe rzędu r z próby mają postać:

$$M_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r, \quad r = 1,2,3,\ldots \tag{14}$$

W przypadku danych zagregowanych, gdzie znamy tylko liczbę polis dla danej liczby szkód, momenty zwykłe z próby wynoszą:

[16] W. R. Heilmann, *Fundamentals of Risk Theory*, Verlag Versiecherungswirtschaft e.V., Karlsruhe 1988, s. 46.

[17] H. H. Panjer, G. E. Willmot, *op. cit.*, s. 292.

$$M_r = \frac{1}{n}\sum_{l=1}^{\infty} k^r N_k, \quad r = 1,2,... \tag{15}$$

gdzie N_k jest liczbą X_i, dla których:

$$X_i = k,\ (k = 0,1,2...),\ n = \sum_{k=0}^{\infty} N_k \ \text{ oraz } \ \overline{X} = \frac{1}{n}\sum_{k=0}^{\infty} kN_k$$

Pierwsze trzy momenty centralne z próby wynoszą:

$$\overline{X} = M_1;\ S^2 = M_2 - M_1^2;\ K = M_3 - 3M_2M_1 + 2M_1^3.$$

Współczynniki częstości opisuje równanie:

$$T_k = (k+1)\frac{N_{k+1}}{N_k}, \quad k = 0,1,2,... \tag{16}$$

Wybierając rozkład liczby roszczeń najpierw rozpatruje się rozkłady z klasy $(a,b,0)$[18], czyli rozkład Poissona, dwumianowy lub ujemny dwumianowy. Niech:

$$T_k = (a+b) + ak, \quad k = 0,1,2,... \tag{17}$$

będzie równaniem pewnej funkcji. W przypadku, gdy funkcja dana równaniem (17) jest prostą, której współczynnik kierunkowy:

– wynosi zero oraz $\overline{X} = S^2$, wówczas do modelowania rozkładu liczby szkód nadaje się rozkład Poissona;

– jest ujemny oraz $\overline{X} > S^2$, wówczas rozkład liczby szkód można modelować rozkładem dwumianowym;

– jest dodatni oraz $\overline{X} < S^2$, powinno się wybrać rozkład ujemny dwumianowy.

W przypadku, gdy funkcja opisana równaniem (17) rośnie szybciej niż liniowo należy rozważyć skośność rozkładu. Jeżeli spełnione jest równanie:

$$K = 3S^2 - 2\overline{X} + 2\frac{(S^2 - \overline{X})^2}{\overline{X}},$$

[18] Por. W. Otto, *Matematyka w ubezpieczeniach. Ubezpieczenia majątkowe*, WNT, Warszawa 2002.

to rozkład ujemny dwumianowy powinien dobrze modelować liczbę szkód. Jeżeli prawdziwa jest nierówność:

$$K < 3S^2 - 2\overline{X} + 2\frac{(S^2 - \overline{X})^2}{\overline{X}},$$

to do opisu rozkładu liczby szkód można użyć uogólnionego rozkładu Poissona –Pascala lub jego specjalnego przypadku rozkładu Poissona – odwrotnego normalnego. W przypadku, gdy prawdziwa jest nierówność:

$$K > 3S^2 - 2\overline{X} + 2\frac{(S^2 - \overline{X})^2}{\overline{X}},$$

to do modelowania rozkładu liczby szkód nadają się rozkłady: Neymana typu A, Polya–Aeppli, Poissona–Pascala lub ujemnego dwumianowego.

4. STATYSTYCZNE METODY OCENY DOPASOWANIA ROZKŁADÓW EMPIRYCZNYCH I TEORETYCZNYCH

Test zgodności χ^2 i test λ-Kołmogorowa to najczęściej stosowane w literaturze aktuarialnej testy, służące do oceny stopnia dopasowania rozkładu teoretycznego do danych empirycznych. Warto jednak zauważyć, że w przypadku rozkładu liczby szkód w ubezpieczeniach komunikacyjnych OC liczba klas często nie przekracza czterech, co sprawia, że liczba stopni swobody testu zgodności chi-kwadrat jest zbyt mała. Ponadto w ubezpieczeniach komunikacyjnych w klasie z liczbą szkód zero koncentruje się większość polis powodując zniekształcenie rozkładu. W przypadku całych portfeli powstaje także problem z liczebnością próby. Test zgodności chi-kwadrat z reguły przy próbach powyżej 900 odrzuca hipotezę zerową nawet mimo dużej zgodności danych empirycznych z badanym rozkładem teoretycznym. W takich przypadkach w literaturze statystycznej można znaleźć miary oceny stopnia dopasowania rozkładu teoretycznego do danych empirycznych, takie jak: odchylenie standardowe różnic częstości względnych, wskaźnik podobieństwa struktur, wskaźnik podobieństwa rozkładów, wskaźnik maksymalnej różnicy częstotliwości względnych, czy wskaźnik maksymalnej różnicy dystrybuant[19].

[19] Por. J. Kordos, *Metody analizy i prognozowania rozkładów płac i dochodów ludności*, PWE, Warszawa 1973, s. 115–118.

Odchylenie standardowe różnic częstości względnych to miara dana wzorem:

$$S_r = \sqrt{\frac{1}{k}\sum_{i=1}^{k}\left(\gamma_i - \hat{\gamma}_i\right)^2} \tag{18}$$

gdzie:

k – liczba klas,
γ_i – częstości empiryczne,
$\hat{\gamma}_i$ – częstości teoretyczne (oszacowane).

Miara jest równa zeru w przypadku pełnej zgodności rozkładu empirycznego i teoretycznego. Z praktyki wynika, że: wartość $S_r \leq 0{,}005$ świadczy o wysokiej zgodności rozkładów; jeżeli $0{,}005 \leq S_r < 0{,}01$, to zgodność badanych rozkładów jest zadawalająca; $S_r \geq 0{,}01$ świadczy o znacznych odchyleniach między badanymi rozkładami.

Wskaźnik podobieństwa struktur przedstawia wzór:

$$w_p = \sum_{i=1}^{k}\min\left(\gamma_i, \hat{\gamma}_i\right) \tag{19}$$

gdzie:
oznaczenia jak wyżej.

Wskaźnik przyjmuje wartości z przedziału [0,1]. Im jego wartość bliższa jedności tym bardziej podobne struktury badanych rozkładów.

Wskaźnik podobieństwa rozkładów określa równanie:

$$W_p = 1 - \frac{1}{2}\sum_{i=1}^{k}\left|\gamma_i - \hat{\gamma}_i\right| \tag{20}$$

gdzie:
oznaczenia jak wyżej.

Wskaźnik podobieństwa rozkładów jest równy 100% dla rozkładów całkowicie zgodnych. Rozkłady wykazują wysoką zgodność gdy $W_p \geq 0{,}97$. Jeżeli $W_p < 0{,}95$ to rozkłady wykazują znaczne rozbieżności.

Wskaźnik maksymalnej różnicy częstotliwości względnych jest dany wzorem:

$$r_{\max} = \max_i \left| \gamma_i - \hat{\gamma}_i \right| \tag{21}$$

gdzie:
oznaczenia jak wyżej.

Wskaźnik ten jest równy zeru dla rozkładów całkowicie zgodnych. Jeżeli $r_{\max} < 0{,}02$, to uważa się, że rozkłady są dość zgodne.

Wskaźnik maksymalnej różnicy dystrybuant definiuje równanie:

$$D_{\max} = \max_i \left| F_i - \hat{F}_i \right| \tag{22}$$

gdzie:

$F_i = \sum_{j=1}^{i} \gamma_j$ – wartość dystrybuanty empirycznej,

$\hat{F}_i = \sum_{j=1}^{i} \hat{\gamma}_i$ – wartość dystrybuanty teoretycznej.

Wskaźnik ten jest równy zeru dla rozkładów całkowicie zgodnych.

We wstępnych analizach zgodności rozkładów liczby szkód z rozkładami teoretycznymi można oprócz wskaźników danych wzorami (18)–(22) stosować porównania parametrów badanych rozkładów, takich jak: średnia, mediana, pierwszy i trzeci kwartyl, pierwszy i dziewiąty decyl, miary zróżnicowania rozkładów. Przyjmuje się, że jeżeli różnice względne ocen wszystkich parametrów rozkładu empirycznego i teoretycznego nie przekraczają 5%, to rozkłady są dość zgodne. Można również stosować testy statystyczne weryfikujące hipotezy o wartościach parametrów rozkładu empirycznego. Zastosowanie wszystkich wymienionych metod pozwala dość dokładnie wybrać rozkład teoretyczny dopasowany do danych empirycznych.

5. PRZYKŁADY EMPIRYCZNE

W części dotyczącej zastosowań dokonano oceny rozkładów liczby szkód w ubezpieczeniach komunikacyjnych na dwóch przykładowych zbiorach danych rzeczywistych. Dokonano estymacji parametrów rozkładów metodą momentów oraz dopasowano rozkład teoretyczny do empirycznego za pomocą miar przedstawionych w części 4 pracy.

Przykład 1. Rozkład liczby szkód dla około 35 000 polis ubezpieczeń komunikacyjnych OC dla rynku niemieckiego w roku 2000 przedstawia tab. 1. Zgodnie z wcześniej przedstawioną zasadą wyboru rozkładu liczby szkód oceniono relację wartości oczekiwanej i wariancji rozkładu empirycznego. Dla danych z tab. 1: $\overline{X} = 0{,}04065, S^2 = 0{,}04051$, czyli $\overline{X} > S^2$. Zatem rozkład liczby szkód można opisać rozkładem dwumianowym. Ponieważ różnica pomiędzy wartością oczekiwaną, a wariancją w rozkładzie empirycznym nie jest duża, zbadano również zgodność z rozkładem Poissona. Wartości miar oceny stopnia dopasowania rozkładów empirycznego i teoretycznych przedstawia tab. 2.

Tabela 1

Rozkłady liczby szkód

Liczba szkód	Liczba polis	Prawdopodobieństwo		
		empiryczne	rozkład dwumianowy	rozkład Poissona
0	338330	0,960084677	0,960167572	0,960167574
1	13816	0,039205893	0,039028371	0,039028367
2	243	0,000689565	0,0007932	0,000793202
3	6	1,70263E-05	1,07471E-05	1,07472E-05
4	1	2,83772E-06	1,0921E-07	1,09212E-07
Suma	352396	1	1	1

Źródło: obliczenia własne na podstawie danych z pracy: S. Gschlößl, C. Czado, *Spatial modelling of claim frequency and claim size in non-life insurance*, „Scandinavian Actuarial Journal" 2007, vol. 3, s. 202–225.

Tabela 2

Miary stopnia dopasowania rozkładów

Miara	Rozkład teoretyczny	
	dwumianowy	Poissona
S_r	0,00016419	0,0000991716
w_p	0,99971900	0,999813466
W_p	0,99999993	0,999813466
r_{max}	0,00000007	0,000177527
D_{max}	0,00024700	9,463E-05

Źródło: badania własne.

Dokonując analizy miar z tab. 2 nieco lepiej do danych empirycznych pasuje rozkład Poissona.

Przykład 2. Tabela 3 przedstawia rozkład liczby szkód w ubezpieczeniach komunikacyjnych dla około 100 tys. polis z rynku belgijskiego z lat 1975/76. W przypadku przykładu 2 parametry rozkładu liczby szkód wynoszą: $\overline{X} = 0{,}1011$, $S^2 = 0{,}1074$, czyli $\overline{X} < S^2$. Dodatkowo oceniono skośność rozkładu $K = 0{,}121647$ oraz wartość wyrażenia:

$$3S^2 - 2\overline{X} + 2\frac{(S^2 - \overline{X})^2}{\overline{X}} = 0{,}120981.$$

Ponieważ $K \geq 3S^2 - 2\overline{X} + 2\frac{(S^2 - \overline{X})^2}{\overline{X}}$, to rozważono następujące rozkłady teoretyczne: Poissona, ujemny dwumianowy, Poissona – odwrotny normalny oraz Neymana typu A. Uogólniony rozkład Poissona–Pascala w tym przypadku nie mógł być rozważany ze względu na niespełnienie przez rozkład empiryczny założeń dotyczących parametrów rozkładu.

T a b e l a 3

Rozkłady liczby szkód

Liczba szkód	Liczba polis	Prawdopodobieństwo				
		empiryczne	Poissona	ujemny dwumianowy	Poissona – odwrotny normalny	Neymana typu A
0	96978	0,906556733	0,903860146	0,906626067	0,90657319	0,906682146
1	9240	0,086376129	0,091362759	0,086212574	0,086359261	0,008552395
2	704	0,006581038	0,004617503	0,006653076	0,002831039	0,004278738
3	43	0,000401967	0,00015558	0,000473678	0,000171557	0,002852519
4	9	8,41326E-05	3,93153E-06	3,23096E-05	1,22115E-05	0,002139389
	106974	1	1	1	1	1

Ź r ó d ł o: obliczenia własne na podstawie danych z pracy: J. L e m a i r e, *Automobil Insurance. Actuarial Models*, Kluwer, Boston 1985.

T a b e l a 4

Miary stopnia dopasowania rozkładów

Miara	Rozkład teoretyczny			
	Poissona	ujemny dwumianowy	Poissona – odwrotny normalny	Neymana typu A
S_r	0,002685542	0,00009442	0,00192820	0,03491716
w_p	0,995013289	0,99978462	0,99623313	0,91542730
W_p	0,99501333	0,99999998	0,99999071	0,99695198
r_{max}	0,00498663	0,00000003	0,00001406	0,00605653
D_{max}	0,002696588	0,00009422	0,00375041	0,08444729

Ź r ó d ł o: badania własne.

Na podstawie tab. 4 należy stwierdzić, że do danych empirycznych najlepiej dopasowany jest rozkład ujemny dwumianowy.

6. PODSUMOWANIE

W przypadku rozkładu liczby szkód w ubezpieczeniach komunikacyjnych OC nie można w sposób jednoznaczny określić typu rozkładu teoretycznego. Dla każdego rynku ubezpieczeniowego rozkład liczby szkód może być zgodny z innym rozkładem teoretycznym. Należy również podkreślić, że na ocenę stopnia dopasowania rozkładów na pewno ma wpływ metoda estymacji parametrów rozkładu, czego nie badano w niniejszej pracy. Na uwagę zasługuje również fakt, że nie wszystkie rozkłady teoretyczne mogą być w ogóle rozważane dla konkretnych danych empirycznych, co wynika z formalnych definicji tych rozkładów.

BIBLIOGRAFIA

Aitchison J., Ho C. H., *The Multivariate Poisson – Lognormal Distribution*, „Biometrika" 1989, vol. 76.

Gomez-Deniz E., Sarabia J. M., Perez J. M., Vazquez F., *Using a Bayesian hierarchical model for fitting automobile claim frequency data*, „Communications in Statistics: Theory and Methods" 2008, vol. 37.

Griffiths D. A., *Maximum Likelihood Estimation for the Beta-Binominal Distribution and Application to the Household Distribution of the Total Number of Cases of Disease*, „Biometrics" 1973, vol. 29.

Gschlőβl S., Czado C., *Spatial modelling of claim frequency and claim size in non-life insurance*, „Scandinavian Actuarial Journal" 2007, vol. 3.

Heilmann W. R., *Fundamentals of Risk Theory*, Verlag Versiecherungswirtschaft e.V., Karlsruhe 1988.

Ibiwoye A., Adeleke A., Aduloju S. A., *Quest for Optimal Bonus-Malus in Automobile Insurance in Developing Economies: An Actuarial Perspective*, „International Business Research" 2011, vol. 4, no. 4.

Kordos J., *Metody analizy i prognozowania rozkładów płac i dochodów ludności*, PWE, Warszawa 1973.

Lemaire J., *Automobil Insurance. Actuarial Models*, Kluwer, Boston 1985.

Lemaire J., *Bonus-Malus Systems in Automobile Insurance*, Kluwer, Boston 1995.

Mert M., Saykan Y., *On a Bonus-Malus System where the Claim Frequency Distribution is Geometric and the Claim Severity Distribution is Pareto*, „Hacettepe Journal of Mathematics and Statistics" 2005, vol. 34.

Otto W., *Matematyka w ubezpieczeniach. Ubezpieczenia majątkowe*, WNT, Warszawa 2002.

Panjer H. H., Willmot G. E., *Insurance risk models*, Society of Actuaries, Schaumburg 1992.

Raport o stanie sektora ubezpieczeń po I półroczu 2012 roku, KNF, Warszawa 2012, www.knf.gov.pl.

Sarabia J. M., Gomez-Deniz E., *Distribuciones multivariantes Poisson – beta con aplicaciones a datos de seguros*, Investigaciones en Seguros y Gestion de Riesgos, Riesgo 2007.

Sarabia J. M., Gomez-Deniz E., Vazquez-Polo F., *On the Use of Conditional Models in Claim Count Distributions: An Application to Bonus-Malus Systems*, „ASTIN Bulletin" 2004, vol. 34, no. 1.

Shengwang M., Yuan W., Whitmore G. A., *Accounting for Individual Over-Dispersion in a Bonus-Malus Automobile Insurance System*, „ASTIN Bulletin" 1999, vol. 29, no. 2.

Tremblay L., *Using the Poisson Inverse Gaussian in bonus-malus systems*, „ASTIN Bulletin" 1992, vol. 22, no.1.

Willmot G. E., *The Poisson-inverse Gaussian distributionas an alternative to the negative binominal*, „Scandinavian Actuarial Journal" 1987.

Dane źródłowe:

Rozporządzenie Ministra Finansów z dnia 28 grudnia 2009 r. w sprawie szczególnych zasad rachunkowości zakładów ubezpieczeń i zakładów reasekuracji, DzU 2009, nr 226, poz. 1825.

Ustawa z dnia 28 lipca 1990 r. o działalności ubezpieczeniowej, DzU 1990, nr 59, poz. 344.

Anna Szymańska

ROZKŁADY LICZBY SZKÓD W UBEZPIECZENIACH KOMUNIKACYJNYCH OC

Ustalenie wysokości składki ubezpieczeniowej jest podstawowym zadaniem każdego towarzystwa ubezpieczeniowego. W ubezpieczeniach komunikacyjnych OC wyznaczenie składki wymaga znajomości rozkładów liczby i wartości szkód w portfelu. W pracy przedstawiono najczęściej wykorzystywane w ubezpieczeniach komunikacyjnych rozkłady modelujące liczbę szkód oraz zwrócono uwagę na metodologiczne problemy dotyczące oceny dopasowania danych empirycznych do rozkładów teoretycznych.

Słowa kluczowe: rozkład liczby szkód, ubezpieczenia komunikacyjne OC.

DISTRIBUTIONS OF THE NUMBER OF CLAIMS IN MOTOR LIABILITY CAR INSURANCE

Determining the amount of the insurance premium is the primary task of each insurance company. In motor liability car insurance determining of premiums requires the knowledge of distribution of the number and amount of claims in the portfolio. The paper presents the most commonly used in motor insurance distributions modeling the number of claims and highlights the methodological problems to evaluate the fit of empirical data to theoretical distributions.

Key words: distribution of the number of claims, motor liability car insurance.