*Roczniki Kolegium Analiz Ekonomicznych*
*Zeszyt 31/2013*

Arkadiusz Filip
Marcin Wienke

# Odporność składki kwantylowej ze względu na zaburzenia rozkładu wielkości pojedynczej szkody w modelu ryzyka łącznego

**Streszczenie**

*Artykuł porusza problem kalkulacji składki przez zakład ubezpieczeń za pomocą metody kwantylowej przy zastosowaniu kolektywnego modelu ryzyka. W ustalonym modelu matematycznym wyliczenie składki wiąże się z błędami aproksymacji wynikającymi z zastosowania jednej z wielu dostępnych metod aproksymacji łącznego rozkładu szkód i zagadnieniem odporności składki na różnego rodzaju zaburzenia. W wyliczeniach uwzględniono możliwość popełnienia różnych błędów w procesie kalkulacji składki, wynikających z faktu, że zakład ubezpieczeń nie dysponuje pełną wiedzą o charakterystykach procesu szkodowego i jest zmuszony przyjmować o nim pewne założenia. Wrażliwość otrzymywanych wyników została zbadana ze względu na zmiany wielkości portfela ubezpieczeniowego, rozkładów prawdopodobieństwa używanych do modelowania wielkości pojedynczej szkody, metody aproksymacji, siły zaburzenia i rzędu kwantyla używanego do wyliczenia składki.*

## 1. Wstęp

Podstawa działalności zakładu ubezpieczeń to wypłata odszkodowań i świadczeń ze środków zgromadzonych w trakcie poboru składki. Relacja zebranych składek do wypłat określa wynik operacyjny ubezpieczyciela. Analizując obie te wielkości, możemy zauważyć, że jedna z nich (szkody) jest w dużej mierze wielkością losową, wielkość zebranej składki zależy natomiast w dużym stopniu od zakładu ubezpieczeń. Losowość wypłacanych odszkodowań oznacza, że możliwe jest zastosowanie narzędzi modelowania matematycznego. Przez proces szkodowości rozumieć będziemy taki proces stochastyczny, którego realizacjami są kolejne wypłaty odszkodowań przez zakład na rzecz ubezpieczonych. Proces szkodowości bywa zwykle dekomponowany na dwa wymiary, które są analizowane oddzielnie – liczba/częstość zgłaszanych szkód (ang. *frequency*) i ich wielkość (ang. *severity*). Składki są najczęściej kalkulowane w taki sposób, że aktuariusz lub dział wyceny produktów mają pewne wyobrażenie i oczekiwania co do tego, w jaki sposób szkody będą się pojawiały dla określonych produktów ubezpieczeniowych, i na podstawie tych oczekiwań dokonują taryfikacji produktu. Oczywiste wydaje się to, że jeśli oba wymiary procesu szkodowego podlegają estymacji przez zakład ubezpieczeń, to możliwe jest wystąpienie błędów, prowadzących do sytuacji, w których pobrana składka jest niewystarczająca do pokrycia zobowiązań z tytułu roszczeń ze strony ubezpieczonych. Sytuację taką można określić w dużym uproszczeniu jako niewypłacalność. Badanie wystarczalności składki w kontekście błędów popełnionych przy jej kalkulacji nazywamy badaniem odporności składki. Przeprowadzenie analizy odporności składki przy założeniu określonych błędów popełnionych przy estymacji jest celem badawczym niniejszej pracy. Jako metodę badawczą przyjęto metodę symulacyjną.

## 2. Wyniki podobnych badań

Przy przeprowadzaniu badania odporności składki korzystaliśmy z wielu narzędzi rachunku prawdopodobieństwa, statystyki matematycznej i matematycznej teorii ryzyka. Dziedzina ta była obiektem częstych analiz w literaturze naukowej. Szczególnie istotnym i często poruszanym zagadnieniem była kwestia wyboru aproksymacji łącznego rozkładu szkód, ponieważ wybór konkretnej metody aproksymacji ma znaczny wpływ na wielkość składki, a w związku z tym również na ewentualne wystąpienie niewypłacalności zakładu.

Oprócz najprostszej aproksymacji normalnej jest wiele innych metod aproksymacyjnych, w szczególności aproksymacja gamma oraz aproksymacja za pomocą szeregu potęgowego stopnia drugiego standardowej zmiennej normalnej (oznaczana dalej jako NP) (więcej o tej metodzie aproksymacji zob. Ramsey, 1991). Aproksymacja normalna najczęściej jest przytaczana w literaturze jako przykład relatywnie niskiej jakości dopasowania przybliżonego rozkładu do danych rzeczywistych (zob. Kaas, Goovaerts, Dhaene, Denuit, 2001). Różnice pomiędzy aproksymacjami gamma i NP były niejednokrotnie przedmiotem sporów w kwestii lepszego odzwierciedlania danych i poświęcono im szereg publikacji naukowych.

Za aproksymacją gamma opowiadał się m.in. amerykański badacz i wykładowca na Yale University – H. Seal. Starał się on dowieść, że aproksymacja za pomocą szeregu potęgowego standardowej zmiennej normalnej (stopnia zarówno drugiego, jak i trzeciego) zyskała niezasłużoną reputację metody relatywnie dobrze przybliżającej nieznany rozkład zmiennej i że powinna zostać zarzucona na rzecz mniej skomplikowanej aproksymacji gamma – więcej szczegółów na temat jego badań można odnaleźć w Seal (1977). Przewaga aproksymacji gamma była szczególnie wyraźna przy modelowaniu prawdopodobieństwa osiągnięcia przez zmienną wartości oddalonych od wartości oczekiwanej o ponad 4, 5 i 6 odchyleń standardowych. Za aproksymacją NP z kolei opowiadali się m.in. T. Pentikäinen (1977), fiński matematyk i aktuariusz, oraz C. Ramsay (1991), amerykański uczony i wykładowca na University of Nebrasca. W swoich pracach ukazali oni fakt przeszacowywania wartości prawdopodobieństw w prawym ogonie rozkładu przez tę aproksymację, co wpływa wprawdzie niekorzystnie na jej precyzję, korzystnie jednak na odporność wyznaczonej za jej pomocą składki.

G. Berger pokazał w swoim opracowaniu, że zastosowanie aproksymacji szeregiem potęgowym stopnia trzeciego standardowej zmiennej normalnej, mimo większego stopnia skomplikowania, nie daje wyraźnie lepszych wyników w porównaniu do NP (zob. Berger, 1972).

Różnym metodom aproksymacji łącznego rozkładu szkód poświęcili też uwagę R. Kaas, holenderski specjalista od teorii ryzyka, ze współpracownikami w książce *Modern Actuarial Risk Theory* (2001) oraz Y. Chaubey wraz z J. Garrido oraz S. Trudeau (1998), z grupą matematyków z Concordia University w Montrealu. Podobnie jak pozostali, pokazali oni słabą jakość aproksymacji normalnej i porównywalną jakość aproksymacji gamma i NP. Jest też wiele innych, rzadziej stosowanych w praktyce ubezpieczeniowej, metod aproksymacji. Chaubey przywołuje w swoim opracowaniu kilka z nich, jak np. aproksymację Edgewortha, Esschera, odwrotną gaussowską czy mieszaną (gamma z odwrotną gaussowską) (Chaubey, Garrido, Trudeau, 1998, s. 217–226). Wspomniane metody aproksymacji zastosował w swoich badaniach również matematyk holenderski R. Reijnen i jego współpracownicy (2005). Są to jednak metody rzadko pojawiające się w praktyce ubezpieczeniowej, nie zostały więc one zastosowane w niniejszej pracy.

W literaturze można też odnaleźć wyniki badań zależności dopasowania rozkładu od rzędu kwantyla. Polscy naukowcy, J. Iwanik i J. Nowicka-Zagrajek, w swoim opracowaniu (2005) porównali składki w modelu ryzyka łącznego wyznaczone przy użyciu aproksymacji gamma ze składkami wyznaczonymi na podstawie rzeczywistych rozkładów w zależności od przyjętego kwantyla. Jakość aproksymacji okazuje się bardzo dobra przy kwantylu rzędu 0,9; przy kwantylu rzędu 0,98 różnice wynikające z zastosowania przybliżeń aproksymacyjnych są już dość wyraźne.

## 3. Narzędzia analityczne i struktura modelu

Struktura modelu oraz użyte narzędzia z zakresu rachunku prawdopodobieństwa i matematyki aktuarialnej zostały omówione kolejno poniżej.

1. Model jest jednookresowy, tzn. przedmiotem symulacji jest liczba i wielkość szkód wygenerowanych przez portfel identycznych polis ubezpieczeniowych na przestrzeni jednej jednostki czasu. Efektem końcowym jest jednorazowe porównanie łącznej wysymulowanej wielkości szkód z zebraną składką.
2. Składka jest kalkulowana metodą kwantylową. Istota tej metody polega na takim wyborze poziomu składki globalnej $H$ dla całego portfela, by prawdopodobieństwo przekroczenia składki przez łączną wartość szkód w portfelu było

odpowiednio małe, np. wyniosło 1%. Na potrzeby badania wprowadzono symbol $B$ na oznaczenie zakładanego przez ubezpieczyciela poziomu bezpieczeństwa, czyli rzędu kwantyla używanego do kalkulacji składki. Zatem składka kwantylowa jest to wartość $H$ spełniająca zależność:

$$P(W > H) = 1 - B,$$

gdzie $W$ to zmienna losowa oznaczająca łączną wartość wypłaconych szkód w jednostce czasu.

3. Rozkład Poissona z parametrem 0,01[1] jest używany do modelowania liczby szkód zgłoszonych przez pojedynczą polisę. Charakterystyczną cechą rozkładu Poissona, która czyni go użytecznym w procesach modelowania liczby zgłaszanych szkód, jest jego addytywność (tzn. dokonanie symulacji 1000 polis, z których każda generuje liczbę szkód zgodną z rozkładem Poissona o parametrze $\lambda$, jest równoznaczne z dokonaniem pojedynczej symulacji dla zmiennej losowej o rozkładzie Poissona z parametrem $1000 \cdot \lambda$).
4. Użycie rozkładu Poissona do modelowania liczby zgłaszanych szkód implikuje zastosowanie modelu ryzyka łącznego, tzn. dopuszcza się możliwość wygenerowania przez jedną polisę liczby szkód większej niż jeden. W modelu ryzyka łącznego rozkład zmiennej losowej $W$ można określić poprzez wskazanie rozkładów odpowiadających odpowiednio za liczbę zgłaszanych szkód i ich wielkość.
5. W zakresie wielkości zgłaszanych szkód dopuszczono możliwość popełniania przez zakład trzech rodzajów błędów, nazywanych również zaburzeniami:
   a) Zaburzenie typu I – towarzystwo ubezpieczeniowe posiada prawidłową i pełną wiedzę na temat procesu szkodowego w okresie poprzedzającym okres symulacji i na tej podstawie dokonuje kalkulacji składki. W okresie, którego ryzyko składka pokrywa, następuje jednak zmiana procesu szkodowego – wielkości szkód, które poprzednio były realizacjami zmiennej losowej o rozkładzie $\Psi$, w okresie składkowym są realizacjami zmiennej losowej o rozkładzie mieszanym. Szkoda z prawdopodobieństwem $1 - \varepsilon$ pochodzi z rozkładu $\Psi$, natomiast z prawdopodobieństwem $\varepsilon$ pochodzi z innego rozkładu $\Xi$. W modelu założono występowanie z prawdopodobieństwem $\varepsilon$ szkód pochodzących z rozkładów o znacznie większej wariancji i silnej dodatniej skośności. Parametr $\varepsilon$ będzie określany jako natężenie zaburzenia.
   b) Zaburzenie typu II – towarzystwo ubezpieczeniowe posiada prawidłową informację na temat dwóch pierwszych momentów rozkładu wielkości pojedynczej szkody[2]. Błędnie zakłada jednak postać funkcyjną $\Psi$ rozkładu

[1] Parametr dla rozkładu Poissona oznacza jednocześnie jego wartość oczekiwaną.

[2] Oznacza to, że zakład prawidłowo oszacował wartość oczekiwaną i wariancję tego rozkładu.

pojedynczej szkody. Szkoda bowiem z prawdopodobieństwem 1 pochodzi z rozkładu Ξ, który przy całkowicie innej postaci funkcyjnej niż rozkład Ψ posiada identyczne z nim wartość oczekiwaną i wariancję. Do kalkulacji składki zakład wykorzystuje informację o rozkładzie Ψ.

c) Zaburzenie typu III – podobnie jak w zaburzeniu typu II towarzystwo ubezpieczeniowe posiada prawidłową informację na temat dwóch pierwszych momentów rozkładu wielkości pojedynczej szkody. Popełniany błąd dotyczy postaci funkcyjnej rozkładu, tzn. na podstawie informacji o dwóch pierwszych momentach zakład założył, że szkody pochodzą z „czystego" rozkładu Ψ, podczas gdy rzeczywista natura procesu szkodowego jest taka, że szkody pochodzą z rozkładu mieszanego zdefiniowanego poprzez częstości $1-\varepsilon$ (rozkład Ψ) i $\varepsilon$ (rozkład Ξ), podobnie jak w zaburzeniu I. Składka kalkulowana jest na podstawie znanych zakładowi ubezpieczeń momentów rozkładu mieszanego opisującego wielkość pojedynczej szkody.

6. Rozkłady główne i zaburzające.
W badaniu odporności składki ubezpieczeniowej jako rozkładów głównych zdecydowano się użyć rozkładów gamma, logarytmiczno-normalnego (LN) oraz Weibulla. Do modelowania zaburzeń wybrano rozkłady Pareto, uogólniony Pareto (ang. *Generalized Pareto* – GP) i Burra. W dalszej części pracy rozkłady te są określane jako zaburzające. Tabele 1 i 2 przedstawiają parametry i momenty użytych rozkładów[3].
7. Ze względu na fakt, że łączna wartość szkód jest zmienną losową trudną do statystycznej analizy, przy ustalaniu wysokości składki w symulacjach korzystano z aproksymacji normalnej, aproksymacji rozkładem gamma i aproksymacji za pomocą szeregu potęgowego standardowej zmiennej normalnej (ang. *normal power* – NP). Wykorzystanie tych aproksymacji, przy założeniu wiedzy zakładu ubezpieczeń na temat określonych parametrów procesu szkodowego, pozwala na kalkulację przybliżonej wartości kwantyla zmiennej *W*, czyli kalkulację składki dla portfela polis. Aproksymacje gamma i NP wykorzystują

[3] Na potrzeby pracy, chcąc otrzymać porównywalne wyniki dla trzech podstawowych rozkładów, z których pochodzą szkody, przyjęto arbitralnie wartość oczekiwaną pojedynczej szkody równą 20 000 zł oraz jej wariancję równą 4 000 000 (zł)$^2$. W celu uzyskania parametrów użyto metody momentów. W przypadku dwuparametrowych rozkładów estymacja parametrów wymagała rozwiązania układu dwóch równań z dwiema niewiadomymi. W przypadku rozkładów zaburzających postanowiono przyjąć istotnie wyższą wariancję, taką by wpływ zaburzeń był odczuwalny. Przy niezmienionej wartości oczekiwanej wariancja została ustalona na poziomie 800 000 000 (zł)$^2$. Ze względu na to, że rozkład Burra i rozkład GP są trójparametrowymi uogólnieniami rozkładu Pareto, aby je w pewien sposób zróżnicować, postanowiono zmienić ich skośność. Skośność rozkładu Burra ustalono na poziomie o 25% niższym niż dla rozkładu Pareto, a dla rozkładu GP na poziomie o 25% wyższym.

informacje o momentach trzech pierwszych rzędów, aproksymacja normalna wykorzystuje informacje o momentach dwóch pierwszych rzędów[4].

Po uwzględnieniu różnych możliwych wielkości portfela i poziomów bezpieczeństwa w tabeli 3 zestawiono parametry opisujące model symulacyjny.

**Tabela 1. Parametry rozkładów użytych w modelu**

| **Rozkład** | **Parametry** | | |
|---|---|---|---|
| Gamma | $\alpha = 100$ | $\beta = 0{,}005$ | |
| LN | $\mu = 9{,}8985$ | $\sigma^2 = 0{,}01$ | |
| Weibulla | $c = 3{,}2012 \cdot 10\text{-}53$ | $\tau = 12{,}1534$ | |
| Pareto | $\lambda = 60\ 000$ | $\theta = 4$ | |
| Burra | $\lambda = 36\ 975{,}42$ | $\tau = 0{,}9172$ | $\theta = 5{,}4067$ |
| GP | $a = -0{,}2682$ | $b = 14\ 092{,}9377$ | $c = 741{,}9607$ |

Źródło: opracowanie własne.

**Tabela 2. Wartości momentów rozkładów użytych w modelu**

| **Rozkład** | **Momenty** | | |
|---|---|---|---|
| | Wartość oczekiwana | Wariancja | Skośność |
| Gamma | 20 000 | 4 000 000 | 0,200 |
| LN | 20 000 | 4 000 000 | 0,301 |
| Weibulla | 20 000 | 4 000 000 | 0,156 |
| Pareto | 20 000 | 800 000 000 | 7,071 |
| Burra | 20 000 | 800 000 000 | 5,303 |
| GP | 20 000 | 800 000 000 | 8,839 |

Źródło: opracowanie własne.

Do osiągnięcia celu badawczego pracy niezbędne było wygenerowanie liczb losowych stanowiących realizację rozkładów opisujących liczbę i wielkość zgłaszanych szkód. W przypadku rozkładu Poissona stworzono tabelę, w której kolejnym wartościom przyjmowanym przez zmienną losową przyporządkowano prawdopodobieństwa ich wystąpienia, a następnie dokonano ich kumulacji. Generując wielkości pojedynczych szkód, skorzystano z faktu, iż dla dowolnego rozkładu ciągłego o dystrybuancie $F$ zmienna losowa $V = F(Y)$ ma rozkład jednostajny na odcinku $[0, 1]$. Losowe wielkości szkód otrzymano więc poprzez odwracanie dystrybuant poszczególnych rozkładów.

[4] Aproksymacja normalna wykorzystuje centralne twierdzenie graniczne, aproksymacja gamma zakłada, że rozkład łącznej wielkości szkód jest rozkładem przesuniętym gamma, aproksymacja NP postuluje standaryzację zmiennej $W$, a następnie przybliżenie jej kwantyla poprzez rozwinięcie w szereg potęgowy. Szczegółowe informacje na temat aproksymacji patrz: Serfling, 1991, s. 36–39; Ramsay, 1991.

**Tabela 3. Parametry modelu symulacyjnego**

| **Rodzaj zaburzenia** | Typ I | Typ II | Typ III | | |
|---|---|---|---|---|---|
| **Rozkład główny** | Gamma | LN | Weibulla | Pareto | |
| **Rozkład zaburzający** | Pareto | GP | Burra | | |
| **Metoda aproksymacji** | normalna | Gamma | NP | | |
| **Siła zaburzenia** | 0%* | 1% | 5% | 10% | 100%** |
| **Poziom bezpieczeństwa** | 98% | 99% | 99,5% | | |
| **Wielkość portfela** | 500 | 1000 | 5000 | | |

* Przeprowadzono również symulacje, w których założono prawidłową wiedzę zakładu o procesie szkodowym. Symulacje takie umożliwiły ocenę jakości metod aproksymacyjnych.

** Zaburzenie na poziomie 100% odpowiada przypadkowi zaburzenia typu II.

Źródło: opracowanie własne.

Pojedyncza symulacja w badaniu oznaczała przeprowadzenie 100 000 losowań wartości dla wszystkich modelowanych zmiennych losowych dla jednego ustalonego zestawu parametrów, zgodnie z danymi z tabeli 3. Pojedyncze losowanie jest określane mianem iteracji. Kolejnym krokiem było obliczenie wartości oczekiwanej i wariancji rzeczywistego procesu szkodowego. Następnie dokonywano za pomocą przyrównania założonych i teoretycznych momentów estymacji parametrów rozkładu rozpoznawanego przez ubezpieczyciela i odpowiedniej przy danej aproksymacji składki globalnej. Zaburzenia zostały zaimplementowane w taki sposób, że dla każdej szkody generowano liczbę losową, która decydowała o tym, czy szkoda będzie pochodzić z rozkładu głównego, czy zaburzającego[5], a następnie niezależnie losowano wielkość szkody poprzez metodę odwracania dystrybuanty. W dalszej kolejności sumę wszystkich zaszłych szkód porównywano ze składką $H$ i w przypadku niewystarczalności składki stan ten był rejestrowany przez kod programu wraz z informacją o wielkości nadwyżki zgłoszonych szkód ponad zebraną składkę.

## 4. Mierniki odporności

Wynik pojedynczej symulacji składa się z trzech elementów: informacji o liczbie przypadków, kiedy wystąpiła niewypłacalność, średniej wartości nadwyżki zgłoszonych szkód ponad pobraną składkę i zaobserwowanej wariancji nadwyżki szkód ponad zebraną składkę. Na bazie tych informacji opracowano mierniki, które dalej posłużą nam do oceny odporności składki. Miernikami tymi są:

[5] Zmienną losową jednostajną interpretowano po prostu jako prawdopodobieństwo.

1. Prawdopodobieństwo niewypłacalności.
   Niech dla danej symulacji $t$ oznacza liczbę tych iteracji spośród 100 000, w których nastąpiła niewypłacalność zakładu, czyli łączna wielkość szkód przekroczyła zebrane składki; prawdopodobieństwo niewypłacalności zarejestrowane dla ustalonej symulacji to wielkość:
$$\overline{q} = \frac{t}{100\,000}.$$
2. Współczynnik $T$.
   Jako współczynnik $T$ oznaczamy dla ustalonej symulacji stosunek dwóch wielkości. W liczniku umieszczamy różnicę między wielkością otrzymanego prawdopodobieństwa niewypłacalności a zakładanym przez towarzystwo maksymalnym prawdopodobieństwem niewypłacalności. W mianowniku zaś umieszczamy wartość maksymalnego, akceptowanego przez zakład prawdopodobieństwa niewypłacalności. Stosunek ten wyrażamy procentowo. Korzystając z wprowadzonych oznaczeń, możemy zapisać:
$$T = \frac{\overline{q} - (1 - B)}{1 - B} \cdot 100\%.$$
   Im wyższa jest wartość $T$, tym składka jest mniej odporna na odpowiedni typ zaburzenia. Przyjęcie przez współczynnik $T$ wartości równej zero interpretujemy jako fakt zarejestrowania maksymalnego dopuszczanego przez towarzystwo prawdopodobieństwa niewypłacalności. Przyjęcie przez współczynnik $T$ wartości mniejszej od 0 oznacza, że cel ubezpieczyciela w zakresie utrzymania prawdopodobieństwa niewypłacalności na odpowiednio niskim poziomie został zrealizowany.
3. Liczba stanów niewypłacalności.
   Interpretacja prawdopodobieństwa niewypłacalności może być nieco skomplikowana, gdyż fakt, że przy założonym poziomie bezpieczeństwa 99% (czyli zakład godzi się na 1-procentowe prawdopodobieństwo niewypłacalności, co odpowiada wystąpieniu niewypłacalności 1000 na 100 000 iteracji) niewypłacalność nastąpiła w 1010 przypadkach, z jednej strony może oznaczać niewłaściwe skalkulowanie składki, a z drugiej – może wynikać z losowości procesu szkodowego. Aby uwzględnić w analizie tę drugą możliwość, skorzystano z teorii testów statystycznych, by zbadać to, czy liczba iteracji, w których nastąpiła niewypłacalność zakładu, jest istotnie statystycznie wyższa od liczby, którą akceptuje zakład ubezpieczeń przy danym poziomie bezpieczeństwa. W przypadku istotnej różnicy dla danej symulacji mówimy o wystąpieniu stanu niewypłacalności. W toku analizy wyników informacja o liczbie stanów niewypłacalności będzie zawsze dotyczyła liczby tak stwierdzonych statystycznie istotnych stanów wypłacalności. Badanie wystąpienia stanu niewypłacalności jest

przeprowadzane następująco. Niech $H$ oznacza wielkość składki dla ustalonej symulacji, a $W_{sym}$ oznacza wielkość zagregowanych szkód zarejestrowanych dla pojedynczej iteracji. Jako $t$ oznaczamy liczbę iteracji (spośród 100 000 przeprowadzonych), dla których zachodzi $W_{sym} > H$. Jeżeli zachodzi $\overline{q} > \overline{q}*$, zaobserwowana frakcja niewypłacalności należy do zdefiniowanego obszaru krytycznego $O$, to stwierdzamy statystycznie istotne wystąpienie stanu niewypłacalności. Lewy kraniec obszaru krytycznego, oznaczony jako $\overline{q}*$, można obliczyć w następujący sposób[6]:

$$\overline{q}* = \frac{(1-B)\cdot 100\,000 + u_{1-\alpha}\cdot\sqrt{100\,000\cdot B\cdot(1-B)}}{100\,000},$$

gdzie $u_{1-\alpha}$ oznacza kwantyl standardowego rozkładu normalnego, a $\alpha$ jest poziomem istotności użytym do testu. Wartość $\overline{q}*$ dla poziomu istotności równego 1% wynosi przykładowo: 0,55% dla poziomu bezpieczeństwa 99,5%; 1,07% dla poziomu bezpieczeństwa 99% i 2,10% dla poziomu bezpieczeństwa 98%.

4. Średnia wartość nadwyżki szkód.
Miernik ten oznacza średnią wartość nadwyżki zgłoszonych szkód $W$ ponad zebraną składkę $H$. Wielkość ta dla ustalonej symulacji będzie oznaczana $e_n(H)$ i dana jest wzorem

$$e_n(H) = \frac{\sum_{i=1}^{N} z_i}{N},$$

gdzie $N$ oznacza liczbę iteracji, w których nastąpiła niewypłacalność zakładu, a $z_i$ kolejne realizacje nadwyżki łącznej wartości szkód ponad wartość zgromadzonych składek w przypadku, gdy nastąpiła niewypłacalność.

5. Współczynnik $R$.
Wielkość $e_n(H)$ jest miernikiem wyrażonym w sposób absolutny. W praktyce ubezpieczeniowej równie istotne może być stwierdzenie, jak wielki był niedobór zebranej składki w stosunku do zgłoszonej sumy szkód. Współczynnik $R$ dla ustalonej symulacji definiujemy zatem jako stosunek średniej wartości nadwyżki szkód do wielkości pobranej składki:

$$R = \frac{e_n(H)}{H}\cdot 100\%.$$

[6] Wzór wynika stąd, że stan niewypłacalności zakładu potraktowano jako realizację dwustanowej zmiennej losowej $Q$, która z pewnym prawdopodobieństwem $q$ przyjmuje wartość 1 odpowiadającą wystąpieniu niewypłacalności, z prawdopodobieństwem $1-q$ przyjmuje zaś wartość 0, co odpowiada wypłacalności. Aby zweryfikować hipotezę, że parametr $q$ jest wyższy od poziomu akceptowanego przez zakład $1-B$, należy zweryfikować odpowiednią hipotezę, posiłkując się estymatorem prawdopodobieństwa zakładu, otrzymanym na podstawie 100 000 iteracji, jako statystyką testującą. Ze względu na dużą liczbę niezależnych realizacji rozkład $\overline{q}$ można przybliżyć rozkładem normalnym o wartości oczekiwanej $q$ i wariancji $\frac{q(1-q)}{100000}$.

Badano również zmienność nadwyżki szkód ponad zgromadzoną składkę, jednak analiza ta nie przyniosła interesujących wniosków.

## 5. Analiza odporności składki

### 5.1. Symulacje pozbawione zaburzeń

Pierwszym etapem analizy było przeprowadzenie symulacji bez zaburzeń procesu szkodowego, które miało pomóc w postawieniu wstępnych hipotez badawczych do dalszej weryfikacji oraz służyć jako punkt odniesienia przy badaniu symulacji z zaburzeniami. Przeprowadzono 81 symulacji dla trzech różnych: wielkości portfela, wartości poziomu bezpieczeństwa, rozkładów wielkości pojedynczej szkody oraz metod aproksymacji.

Badanie liczby stanów niewypłacalności jednoznacznie wykazało słabość aproksymacji normalnej. Istotny stan niewypłacalności wystąpił we wszystkich 27 symulacjach przeprowadzonych przy tej metodzie aproksymacji (dla porównania dla aproksymacji gamma i NP było to odpowiednio 12 i 9 stanów). Słabość aproksymacji normalnej potwierdza również porównanie współczynnika $T$, przedstawione w tabeli 4.

**Tabela 4. Współczynnik T dla symulacji bez zaburzeń**

| Współczynnik T dla symulacji bez zaburzeń | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Aproksymacja | Rozkład główny | Poziom bezpieczeństwa i wielkość portfela | | | | | | | | |
| | | 98,00% | | | 99,00% | | | 99,50% | | |
| | | 500 | 1000 | 5000 | 500 | 1000 | 5000 | 500 | 1000 | 5000 |
| Normalna | gamma | 54% | 39% | 21% | 74% | 60% | 28% | 132% | 93% | 31% |
| | LN | 51% | 37% | 16% | 76% | 53% | 27% | 128% | 89% | 43% |
| | Weibulla | 48% | 35% | 17% | 73% | 66% | 31% | 130% | 88% | 30% |
| Gamma | gamma | –1% | 1% | –1% | –1% | 2% | 1% | –9% | –5% | 3% |
| | LN | –2% | 5% | 1% | 7% | 3% | 1% | –7% | –11% | 8% |
| | Weibulla | –2% | –1% | –1% | 7% | –3% | –1% | –6% | 1% | –3% |
| NP | gamma | –8% | –7% | 0% | –1% | –3% | –3% | 2% | –5% | 1% |
| | LN | –7% | 1% | 0% | 1% | –5% | 3% | 2% | –8% | 1% |
| | Weibulla | –5% | –1% | –1% | 0% | –8% | –1% | –11% | –4% | –2% |

Źródło: opracowanie własne.

Przykładowo, wartość 54% w lewym górnym rogu tabeli oznacza, że w przypadku symulacji bez zaburzeń, gdy rozkładem głównym jest rozkład gamma, portfel zawiera 500 polis, a towarzystwo ubezpieczeniowe założyło poziom bezpieczeństwa równy 98% i stosuje aproksymację normalną, otrzymana wartość estymatora prawdopodobieństwa niewypłacalności jest o 54% wyższa od zakładanej przez towarzystwo[7]. Analiza danych z tabeli 4 pozwala również stwierdzić, że jakość aproksymacji normalnej istotnie zwiększa się w miarę wzrostu wielkości portfela (co ma swoje uzasadnienie w centralnym twierdzeniu granicznym), jak również istotnie zmniejsza się w miarę wzrostu poziomu bezpieczeństwa (w przypadku modelowania zdarzeń coraz rzadszych liczebność próby niezbędna do tego, by można było zastosować centralne twierdzenie graniczne, jest coraz większa). Warto zwrócić uwagę na fakt, że podobnych zależności nie zaobserwowano w przypadku dwóch pozostałych metod aproksymacji.

## 5.2. Zaburzenie typu I – rozkłady główne gamma, LN i Weibulla

W przypadku badania zaburzenia typu I dla rozkładów gamma, LN i Weibulla ogólna liczba przeprowadzonych symulacji wynosi 729[8]. Analiza liczby stanów niewypłacalności potwierdza niską jakość aproksymacji normalnej zaobserwowaną przy badaniu symulacji bez zaburzeń. Stan niewypłacalności został zarejestrowany w 583 przypadkach, to jest w 80% ogółu przeprowadzonych symulacji. Na 243 symulacje przypadające na każdą z metod aproksymacji składka okazała się nieodporna w 241 symulacjach z wykorzystaniem aproksymacji normalnej, 202 przypadkach z wykorzystaniem aproksymacji gamma i 140 przypadkach z wykorzystaniem aproksymacji NP. Nie zaobserwowano istotnych różnic pomiędzy symulacjami z zastosowaniem różnych rozkładów zaburzających. Występuje jednak dość wyraźna dodatnia zależność pomiędzy siłą zaburzenia a liczbą zaobserwowanych nieodpornych składek. Dla siły zaburzenia równej 0,01 stan niewypłacalności zaobserwowano w 136 przypadkach, dla siły zaburzenia 0,05 w 220 przypadkach, a dla siły zaburzenia 0,1 w 227 przypadkach na 243 wszystkie symulacje.

Analiza prawdopodobieństwa niewypłacalności potwierdza uzyskane wyżej rezultaty, pozwala również na sformułowanie kolejnego ciekawego wniosku, że na niewypłacalność składki duży wpływ ma skośność rozkładu zaburzającego. Przypomnijmy, że rolę rozkładów zaburzających pełnią rozkład Burra (ze skośnością 5,303), rozkład Pareto (ze skośnością 7,071) oraz rozkład GP (ze skośnością 8,839). Przy okazji porównywania wyników trzech symulacji dla przykładowego portfela 500 polis różniących się tylko rozkładem zaburzającym zaobserwowano, że w 52% przypadków największa wartość prawdopodobieństwa niewypłacalności przyjmowana jest dla symulacji z zaburzeniem rozkładem Burra, a tylko w 17%

[7] Towarzystwo, wybierając 98-procentowy poziom bezpieczeństwa, zakłada prawdopodobieństwo niewypłacalności na poziomie 2%, estymator tego prawdopodobieństwa wynosi zaś 3,08%.

[8] Założono możliwość wystąpienia trzech rozkładów głównych i trzech rozkładów zaburzających występujących z trzema siłami zaburzenia. Składkę wyliczano deterministycznie, korzystając z trzech metod aproksymacji i dla trzech poziomów bezpieczeństwa. Rozmiar portfela mógł przyjąć jedną z trzech wartości, tak więc łączna liczba możliwych kombinacji parametrów definiujących wynosi $3^6 = 729$ kombinacji.

dla symulacji z zaburzeniem rozkładem GP. Najmniejsze wartości zarejestrowano w 42% przypadków dla rozkładu GP, a w 22% przypadków dla rozkładu Burra. Do analogicznych wniosków prowadzi również analiza wielkości współczynnika $T$. Pozwala to twierdzić, że wzrost skośności rozkładu zaburzającego przekłada się na wyższą odporność składki. Wyjaśnieniem tej zależności może być fakt, że w miarę wzrostu skośności częściej pojawiają się mniejsze szkody, które są z łatwością pokrywane przez składkę, w przypadku zaś zajścia szkód ekstremalnie dużych wielkość szkody nie jest już taka ważna, gdyż składka i tak okazuje się niewystarczająca.

Badanie współczynnika $T$ pozwala na przedstawienie także innego wniosku. Wraz ze wzrostem siły zaburzenia, bez względu na metodę aproksymacji, wzrasta wartość współczynnika $T$, co jest równoznaczne ze wzrostem prawdopodobieństwa niewypłacalności. Potwierdza się zaobserwowana dla symulacji bez zaburzeń zależność, że jakość aproksymacji normalnej poprawia się wraz ze wzrostem wielkości portfela i pogarsza się wraz ze wzrostem poziomu bezpieczeństwa. Zależność ta jest widoczna we wszystkich przeprowadzonych symulacjach.

### 5.3. Zaburzenie typu I – analiza rodziny rozkładów Pareto

Dla zaburzenia typu I odrębnej analizie poddano rodzinę rozkładów Pareto – rozkład Pareto potraktowano jako rozkład główny, natomiast rozkłady Burra i GP jako rozkłady zaburzające. Podstawowym celem takiej analizy było głębsze zbadanie zależności pomiędzy skośnością rozkładu a odpornością składki. Ponieważ dwa pierwsze momenty rozkładów Pareto, Burra i GP są identyczne, istotne rozbieżności pomiędzy wynikami symulacji pozwolą lepiej ocenić wpływ różnic wartości momentów wyższych rzędów na odporność składki, przede wszystkim skośności. Przy niezmienionym procesie szacowania parametrów procesu szkodowego[9] składki dla procesu szkodowego niezaburzonego i procesów zaburzonych będą takie same. Dla rozkładów z rodziny Pareto przeprowadzono łącznie 189 symulacji – 162 z zaburzeniem i 27 bez zaburzenia[10].

W przypadku symulacji bez zaburzeń stan niewypłacalności wystąpił w dziewięciu przypadkach – są to wszystkie przypadki zastosowania aproksymacji normalnej. Słabość aproksymacji normalnej w przypadku rozkładów Pareto jest tym większa, ponieważ nie uwzględnia ona w ogóle skośności, która jest znacznie wyższa niż we wcześniej rozpatrywanych rozkładach głównych. Właśnie skośność ma

[9] Szacowanie parametrów przez zakład ubezpieczeń na podstawie prawidłowych informacji o wartości oczekiwanej i wariancji.

[10] Dla przypadków bez zaburzenia liczba kombinacji powstała jako iloczyn trzech wielkości portfela, trzech poziomów bezpieczeństwa i trzech metod aproksymacji. Dla przypadków z zaburzeniem liczba kombinacji powstała przez uwzględnienie trzech wielkości portfela, trzech poziomów bezpieczeństwa, trzech metod aproksymacji, dwóch rozkładów zaburzających i trzech sił zaburzenia.

istotny wpływ na kalkulację składki w przypadku aproksymacji gamma i NP, co pokazuje tabela 5.

**Tabela 5. Składki dla portfela 1000 polis przy niezaburzonym rozkładzie głównym**

| Aproksymacja | Poziom bezpieczeństwa | Rozkład główny | | | | |
|---|---|---|---|---|---|---|
| | | gamma | LN | Weibulla | Pareto | gamma1 |
| Normalna | 98,0% | 330 538 | 330 538 | 330 538 | 424 977 | 424 977 |
| | 99,0% | 347 865 | 347 865 | 347 865 | 454 839 | 454 839 |
| | 99,5% | 363 722 | 363 722 | 363 722 | 482 168 | 482 168 |
| Gamma | 98,0% | 341 199 | 341 201 | 341 190 | 506 171 | 474 281 |
| | 99,0% | 362 665 | 362 667 | 362 652 | 573 824 | 524 888 |
| | 99,5% | 382 791 | 382 792 | 382 774 | 640 859 | 573 814 |
| NP | 98,0% | 341 479 | 341 480 | 341 469 | 521 513 | 478 608 |
| | 99,0% | 362 865 | 362 867 | 362 852 | 587 195 | 528 370 |
| | 99,5% | 382 881 | 382 883 | 382 864 | 651 215 | 576 083 |

Źródło: opracowanie własne.

W tabeli 5 na szczególną uwagę zasługuje porównanie składek skalkulowanych przy założeniu, że rozkładem szkody jest Pareto (przedostatnia kolumna) i rozkład gamma, ale o takiej samej wartości oczekiwanej i wariancji jak rozkład Pareto (ostatnia kolumna). Rozkłady te różnią się skośnością, która dla rozkładu Pareto wynosi 7,07, a dla rozkładu gamma 2,83. Przy aproksymacji normalnej nie ma różnic pomiędzy wielkościami składek w obu sytuacjach, gdyż metoda ta uwzględnia tylko dwa pierwsze momenty, które są identyczne dla obu rozkładów. Istotne różnice można jednak zauważyć dla dwóch pozostałych metod aproksymacji. Założenie, że szkody pochodzą z rozkładu Pareto, prowadzi do kalkulacji składki o ok. 10% wyższej w stosunku do założenia, że za szkody odpowiada rozkład gamma. Wyższa składka implikuje zaś mniejsze prawdopodobieństwo niewypłacalności. Przykład ten pokazuje, że w przypadku wyboru do kalkulacji składki jednego rozkładu spośród kilku istotnie różniących się skośnością wybór metody aproksymacji ma bardzo duże znaczenie z punktu widzenia niewypłacalności.

Badanie symulacji bez zaburzeń dla rozkładu Pareto potwierdziło zaobserwowane wcześniej wnioski dotyczące zależności pomiędzy jakością aproksymacji normalnej a wielkością portfela i poziomem bezpieczeństwa. W przypadku symulacji z zaburzeniami stan niewypłacalności wystąpił we wszystkich przypadkach stosowania aproksymacji normalnej (54) i w jednym przypadku aproksymacji gamma[11]. Analiza współczynnika $T$ pozwala stwierdzić, że wprowadzenie zaburzenia i zwiększanie jego natężenia nie wpływa w sposób jednoznaczny na niewypłacalność zakładu.

[11] Zaburzenie rozkładem Burra o sile 5% dla portfela wielkości 5000 polis przy poziomie bezpieczeństwa 99,5%.

**Tabela 6. Wielkość współczynnika $T$ dla symulacji z zaburzeniami – rodzina rozkładów Pareto, portfel 1000 polis**

| Aproksymacja | Poziom bezpieczeństwa | Rozkład zaburzający i siła zaburzenia | | | | | |
|---|---|---|---|---|---|---|---|
| | | Burra | | | GP | | |
| | | 0,01 | 0,05 | 0,1 | 0,01 | 0,05 | 0,1 |
| Normalna | 98,0% | 94,5% | 87,0% | 89,0% | 88,0% | 90,0% | 83,5% |
| | 99,0% | 165,0% | 171,0% | 174,0% | 165,0% | 165,0% | 173,0% |
| | 99,5% | 294,0% | 286,0% | 292,0% | 288,0% | 276,0% | 296,0% |
| Gamma | 98,0% | –20,5% | –20,0% | –24,0% | –24,5% | –25,5% | –24,0% |
| | 99,0% | –19,0% | –18,0% | –23,0% | –25,0% | –30,0% | –28,0% |
| | 99,5% | –12,0% | –12,0% | –20,0% | –20,0% | –28,0% | –26,0% |
| NP | 98,0% | –31,5% | –35,5% | –33,5% | –38,0% | –36,5% | –39,5% |
| | 99,0% | –37,0% | –41,0% | –33,0% | –37,0% | –36,0% | –34,0% |
| | 99,5% | –24,0% | –32,0% | –32,0% | –28,0% | –30,0% | –30,0% |

Źródło: opracowanie własne.

Na podstawie tabeli 6 możemy porównać wartości współczynnika $T$ dla przykładowego portfela 1000 polis w sytuacji, gdy rozkładem zaburzającym jest rozkład Burra i rozkład GP. Na 27 przypadków przedstawionych powyżej w 20 z nich współczynnik $T$ przyjął wyższą wartość dla zaburzenia rozkładem Burra, a tylko w czterech dla rozkładu GP. Wzrost współczynnika $T$ wpływa niekorzystnie na niewypłacalność zakładu, na podstawie powyższej analizy możemy więc potwierdzić zaobserwowany wcześniej wniosek, że wzrost skośności rozkładu zaburzającego zwiększa odporność składki[12].

## 5.4. Zaburzenie typu II

W przypadku zaburzenia typu II nie występują rozkłady zaburzające, szkody pochodzą z jednego rozkładu, jest to jednak inny rozkład niż ten, który zakłada przy kalkulacji składki towarzystwo ubezpieczeniowe. Ogólna liczba przeprowadzonych symulacji wynosi 162[13].

Z jednej strony, wydaje się, iż wykorzystanie do kalkulacji składki rozkładu o zupełnie innej postaci funkcyjnej niż ten, z którego pochodzą szkody, powinno mieć negatywny wpływ na odporność składki. Z drugiej zaś, fakt niewystępowania rozkładów zaburzających, cechujących się relatywnie większą wariancją w stosunku do rozkładów głównych, powinien sprzyjać mniejszemu prawdopodobieństwu niewypłacalności, gdyż są mniejsze szanse wystąpienia szkód o bardzo

[12] Skośność rozkładu GP wynosi 8,839 i jest wyższa od skośności rozkładu Burra, wynoszącej 5,303.

[13] Pojedyncza symulacja jest zdeterminowana przez jedną z trzech metod aproksymacji, jeden z trzech poziomów bezpieczeństwa, jedną z trzech wielkości portfela, jeden z trzech rozkładów, jaki towarzystwo zakłada, że opisuje szkody, i jeden z dwóch rozkładów, jaki w rzeczywistości cechuje pojawiające się szkody. Łącznie są to: $3^4 \cdot 2 = 162$ symulacje.

dużej wartości, które mogą istotnie wpłynąć na niewypłacalność zakładu. Analiza pozwoli ocenić, który z tych czynników ma większe znaczenie.

Łączna liczba przypadków niewypłacalności na 162 symulacje wyniosła 56 (w tym wszystkie symulacje z aproksymacją normalną – 54), co stanowi 34,6% wszystkich przypadków. Jest to wynik lepszy w porównaniu z zaburzeniem typu I (spowodowało ono wystąpienie niewypłacalności w 80% symulacji). Odporność składki może być wyższa niż w przypadku zaburzenia typu I, gdyż w zaburzeniu typu I, w przeciwieństwie do zaburzenia typu II, z ustalonym dodatnim prawdopodobieństwem szkody pochodzą z rozkładu zaburzającego o znacznie wyższej wariancji w stosunku do rozkładów głównych, co sprzyja powstawaniu szkód ekstremalnych o bardzo dużej wartości i może przekładać się na większe prawdopodobieństwo niewypłacalności. Składka w zaburzeniu typu I nie uwzględnia zwiększonej wariancji rozkładu wielkości pojedynczej szkody, stan ten nie występuje w przypadku zaburzenia typu II.

**Tabela 7. Współczynnik $T$ dla zaburzenia typu II i wariantu bez zaburzeń (portfel 500 polis)**

| Aproksymacja | Poziom bezpieczeństwa | Rozkład zakładany i rzeczywisty | | | | | |
|---|---|---|---|---|---|---|---|
| | | Gamma | | | LN | | |
| | | gamma | LA | Weibulla | LN | gamma | Weibulla |
| Normalna | 98,00% | 54% | 54% | 54% | 51% | 48% | 51% |
| | 99,00% | 74% | 79% | 73% | 76% | 73% | 80% |
| | 99,50% | 132% | 130% | 136% | 128% | 142% | 127% |
| Gamma | 98,00% | –1% | –1% | 0% | –2% | 0% | –1% |
| | 99,00% | –1% | 1% | 5% | 7% | –1% | 7% |
| | 99,50% | –9% | –4% | –1% | –7% | 3% | 3% |
| NP | 98,00% | –8% | –2% | –1% | –7% | –2% | –3% |
| | 99,00% | –1% | –2% | –2% | 1% | –1% | –5% |
| | 99,50% | 2% | –11% | 0% | 2% | –3% | –12% |

Źródło: opracowanie własne.

Analiza wartości współczynnika $T$ w przypadku wystąpienia zaburzenia typu II w porównaniu do wariantu bez zaburzeń uwidacznia brak istotnych różnic, co pokazuje tabela 7. W tabeli tej w kolumnach pierwszej i czwartej (kolumny białe) rozkład zakładany jest tożsamy z rzeczywistym, co oznacza brak zaburzenia; w pozostałych kolumnach rozkłady te są różne. Widać, że wystąpienie zaburzenia typu II nie pogarsza w istotny sposób wypłacalności zakładu ubezpieczeń. Wydaje się, że błąd, jaki popełnia towarzystwo ubezpieczeniowe, zakładając przy kalkulacji składki inną od rzeczywistej postać funkcyjną rozkładu wielkości pojedynczej szkody, nie pociąga za sobą poważnych konsekwencji, o ile momenty rozkładu rzeczywistego i zakładanego są do siebie zbliżone. W przypadku, gdy szkody pochodzą z jednego z trzech rozpatrywanych rozkładów głównych, a zakład dokona kalkulacji składki na podstawie innego z tych trzech rozkładów, nie będzie to miało wyraźnego wpływu na odporność tej składki. Może to wiązać się ze zbliżonymi współczynnikami skośności wszystkich trzech rozkładów przyjętych w niniejszej pracy jako główne.

## 5.5. Zaburzenie typu III

W zaburzeniu typu III proces szkodowy wygląda identycznie jak w zaburzeniu typu I. W zaburzeniu typu I zakład ubezpieczeń mylił się jednak zarówno co do homogeniczności polis, jak i co do wartości dwóch pierwszych momentów, w zaburzeniu typu III zaś zakład wprawdzie wciąż popełnia błąd na etapie identyfikacji rozkładu szkód, zna jednak rzeczywiste wartości średniej i wariancji rozkładu mieszanego. Zaburzenie typu III ma dwojaki wpływ na wystąpienie niewypłacalności zakładu. Z jednej strony, z dodatnim prawdopodobieństwem pojawia się rozkład zaburzający o wyraźnie większej wariancji, który identycznie jak w zaburzeniu typu I sprzyja większemu prawdopodobieństwu niewypłacalności zakładu. Z drugiej jednak, wiedza towarzystwa ubezpieczeniowego na temat rzeczywistych momentów rozkładu pojedynczej szkody skutkuje kalkulacją wyższej składki[14].

Bezpośrednia porównywalność zaburzenia typu I i typu III (identyczny rozkład szkodowy i wyższa składka w przypadku zaburzenia typu III) pozwala twierdzić, że w przypadku zaburzenia typu III składka okaże się bardziej odporna. Znajduje to swoje uzasadnienie w analizie liczby stanów niewypłacalności, co widać w tabeli 8.

**Tabela 8. Porównanie liczby stanów niewypłacalności dla zaburzenia typu I i typu III**

| Aproksymacja | Typ I | | | | Typ III | | | |
|---|---|---|---|---|---|---|---|---|
| | siła zaburzenia | | | | siła zaburzenia | | | |
| | 0,01 | 0,05 | 0,1 | suma | 0,01 | 0,05 | 0,1 | suma |
| Normalna | 79 | 81 | 81 | **241** | 81 | 81 | 81 | **243** |
| Gamma | 43 | 81 | 78 | **202** | 1 | 52 | 67 | **120** |
| NP | 14 | 58 | 68 | **140** | 4 | 43 | 63 | **110** |
| Suma | **136** | **220** | **227** | **583** | **86** | **176** | **211** | **473** |

Źródło: opracowanie własne.

Stan niewypłacalności w przypadku zaburzenia typu III nastąpił w 473 przypadkach na 729 symulacji, czyli w 64,9% przypadków. W przypadku zaburzenia typu I niewypłacalność zaś nastąpiła w 80% przypadków. Dane z tabeli 8 potwierdzają również obserwowaną wcześniej niską jakość aproksymacji normalnej. Ciekawe wnioski wyciągnięto na podstawie porównania liczby stanów niewypła-

[14] Pojawienie się zaburzenia o sile 1% powoduje przyrost składki o mniej niż 1%.

calności w przypadku zaburzenia typu III i wariantu bez zaburzeń, co widać w tabeli 9.

**Tabela 9. Porównanie liczby stanów niewypłacalności dla zaburzenia typu III i wariantu bez zaburzeń**

| Rozkład główny | Siła zaburzenia | | | |
|---|---|---|---|---|
| | 0 | 0,01 | 0,05 | 0,1 |
| Gamma | 16 | 11 | 20 | 24 |
| LN | 21 | 10 | 19 | 24 |
| Weibulla | 11 | 9 | 18 | 24 |
| **Suma** | **48** | **30** | **57** | **72** |

Źródło: opracowanie własne.

W tabeli 9 pierwsza kolumna, oznaczona siłą zaburzenia 0, odpowiada wariantowi bez zaburzeń. Widać, że wystąpienie zaburzenia z częstością 1% ogółu szkód prowadzi do skalkulowania składki odporniejszej niż w sytuacji braku jego wystąpienia. Wydaje się to być spowodowane tym, iż efekt wzrostu składki wywołanego użyciem innych momentów do jej kalkulacji przeważył nad efektem wzrostu szkód wywołanego wystąpieniem rozkładu zaburzającego o istotnie wyższej wariancji. Dla wyższych sił zaburzenia (5% i 10%) zależność ta nie występuje.

## 5.6. Analiza średniej nadwyżki szkód ponad składkę

Analiza średniej wartości nadwyżki szkód pozwoliła stwierdzić, że stosowana metoda aproksymacji ma niewielki wpływ na wielkość współczynnika $R$. Zwiększanie poziomu bezpieczeństwa i rozmiaru portfela powoduje spadek wielkości współczynnika $R$ i średniej wartości nadwyżki szkód[15].

W przypadku wystąpienia zaburzenia typu I średnie wartości nadwyżki szkód są wyższe niż dla symulacji bez zaburzeń. Oznacza to, że wprowadzenie zaburzeń rozkładami o dużej wariancji i skośności prowadzi do pogorszenia sytuacji ubezpieczyciela, jeżeli nie występuje podwyższenie składki[16]. Można również stwierdzić dodatnią zależność pomiędzy wielkością siły zaburzenia a średnią wartością nadwyżki szkód. Przyczyny takiego zjawiska należy szukać w rozkładzie zaburzającym, który ma istotnie wyższą wariancję od rozkładu głównego i może powodować występowanie ekstremalnie dużych szkód, co przekłada się na wyższe wartości średniej nadwyżki szkód i współczynnika $R$. Podobne wnioski można zaobserwować w przypadku zaburzenia typu III.

[15] W wartościach absolutnych tylko dla poziomu bezpieczeństwa przy zmianie wielkości portfela wartości absolutne są nieporównywalne.

[16] Charakter zaburzenia typu I powoduje, że składki kalkulowane przez ubezpieczyciela są takie same jak dla wariantu bez zaburzeń.

Analiza rodziny rozkładów Pareto pokazała jeszcze jedną ciekawą zależność. Gruboogonowy charakter rozkładu Pareto powoduje, iż warunkowa wartość oczekiwana niedoboru jest rosnącą funkcją wielkości skalkulowanej składki. Średnia wartość nadwyżki szkód rośnie więc wraz ze wzrostem wielkości portfela i poziomu bezpieczeństwa.

We wszystkich wariantach widoczna była ujemna zależność pomiędzy poziomem bezpieczeństwa i rozmiarem portfela a średnią wartością nadwyżki szkód i współczynnikiem $R$.

## 6. Dalsze potencjalne obszary badań

Model użyty w badaniu może zostać poszerzony o dodatkowe elementy. Wśród najistotniejszych modyfikacji, które mogłyby być zaimplementowane, należy wymienić:

1. Wprowadzenie zależności pomiędzy poszczególnymi polisami – w aspekcie zarówno częstości zgłaszanych szkód, jak i ich wielkości.
2. Osobne modelowanie i symulacja najważniejszych procesów związanych z działalnością ubezpieczeniową – procesu powstawania szkód, procesu zgłaszania szkód, procesu ich weryfikacji i wypłaty i procesu inwestycyjnego. Każdy z tych procesów może być zależny od różnych zestawów zmiennych mikro- i makroekonomicznych.
3. Wprowadzenie analizy wieloookresowej. Wynik jednego okresu (różnica między zebraną składką i wypłaconymi szkodami) stanowiłby punkt startowy dla okresu kolejnego. W takim wypadku celowe byłoby uwzględnienie możliwości „uczenia się” ubezpieczyciela – podwyższania/obniżania składek w zależności od informacji zbieranych we wcześniejszych okresach szkodowych.
4. Uwzględnienie w modelowaniu wypłat takich powszechnie stosowanych rozwiązań, jak franszyza i maksymalne poziomy odpowiedzialności ubezpieczyciela.
5. Zastosowanie innych metod aproksymacji rozkładu łącznej wielkości szkód, które będą uwzględniały więcej charakterystyk rozkładu pojedynczej szkody, w szczególności kurtozę.

## Bibliografia

[1] Berger G. (1972), *Integration of the normal power approximation*, „ASTIN Bulletin”, vol. 7, s. 90–95.

[2] Chaubey Y., Garrido J., Trudeau S. (1998), *On the computation of aggregate claims distributions: some new approximations*, „Insurance: Mathematics and Economics”, vol. 23, s. 215–230.

[3] Iwanik J., Nowicka-Zagrajek J. (2005), *Premium in the individual and collective risk model*, w: *Statistical Tools for Finance and Insurance*, red. P. Čižek, W. Härdle, R. Weron, Springer, New York, s. 423–425.

[4] Kaas R., Goovaerts M., Dhaene J., Denuit M. (2001), *Modern Actuarial Risk Theory*, Kluwer Academic Publishers, Boston.

[5] Pentikäinen T. (1967), *On the solvency of insurance companies*, „ASTIN Bulletin", vol. 4, s. 236–247.

[6] Pentikäinen T. (1977), *On the approximation of the total amount of claims*, „ASTIN Bulletin", vol. 9, s. 281–289.

[7] Ramsay C.M. (1991), *A note on the Normal Power Approximation*, „ASTIN Bulletin", vol. 21, s. 147–150.

[8] Reijnen R., Albers W., Kallenberg C.M. (2005), *Approximations for stop-loss reinsurance premiums*, „Insurance: Mathematics and Economics", vol. 36, s. 237–250.

[9] Seal H.L. (1977), *Approximations to risk theory's $F(x,t)$ by means of the gamma distribution*, „ASTIN Bulletin", vol. 9, s. 213–218.

[10] Serfling R.J. (1991), *Twierdzenia graniczne statystyki matematycznej*, Wydawnictwo Naukowe PWN, Warszawa.

***

## Robustness of quantile premium with respect to variation of single loss amount distribution in collective risk model

**Abstract**

*In the article, we analyze the problem of a premium calculation by an insurance company using quantile method and collective risk model. In a given model, the premium calculation is influenced by approximation errors resulting from using one of many available approximation methods and premium robustness to different disturbances. In our analysis, we considered the possibility of errors in the premium calculation resulting from the fact, that an insurance company does not have the full knowledge about the claim process characteristics and is forced to take certain assumptions. We tested the sensitivity of the results with respect to the changes of the insurance portfolio size, probability distributions used to describe the amount of a single claim, approximation method, the level of disturbance and the quantile used for premium calculatio.*

**Autorzy**:

Arkadiusz Filip, Instytut Ekonometrii, Szkoła Główna Handlowa w Warszawie, Al. Niepodległości 162, 02-554 Warszawa,

e-mail: arekfilip@gmail.com

Marcin Wienke,
e-mail: marcin.eternauta@gmail.com