

Anna Nikodem

Akademia Ekonomiczna we Wrocławiu

METODA PUNKTU SIODŁOWEGO APROKSYMACJI ROZKŁADÓW

1. Wstęp

Dla większości postaci rozkładów liczby wypłat oraz wielkości indywidualnych wypłat generowanych z portfela ubezpieczeniowego postać rozkładu zagregowanych szkód jest zbyt skomplikowana obliczeniowo. W takich przypadkach złożony rozkład przybliża się rozkładem o prostszej postaci. W artykule zostanie przedstawiona metoda punktu siodłowego (*saddlepoint approximation*) i jej zastosowanie do aproksymacji rozkładu zagregowanej wielkości wypłat, przy założeniu, że wielkości indywidualnych wypłat są niezależnymi zmiennymi losowymi o jednakowych rozkładach.

2. Model zagregowanych wypłat

Łączna wielkość wypłat S generowanych z portfela ubezpieczeniowego w jednostkowym przedziale czasowym przybiera postać:

$$S = \sum_{i=1}^N X_i, \quad (1)$$

gdzie zmienna losowa X_i jest wielkością i -tej szkody, a zmienna N jest liczbą wypłacanych roszczeń z portfela polis. Dla uproszczenia zakłada się, że odszkodowania $X_1, X_2, \dots, X_N$ są niezależnymi dodatnimi zmiennymi losowymi o jednakowym rozkładzie oraz że wielkości wypłat oraz ich liczba są wzajemnie niezależne.

3. Metoda punktu siodłowego

Dla zmiennej losowej S mającej dystrybuantę F_S funkcja charakterystyczna φ zmiennej S jest funkcją określoną dla rzeczywistych wartości t postaci:

$$\varphi_S(t) \stackrel{\text{def}}{=} E[e^{itS}] = \int_{-\infty}^{\infty} e^{its} dF_S(s). \quad (2)$$

W przypadku rozkładów złożonych, gdzie zmienna S jest postaci (1), funkcję charakterystyczną można zapisać jako złożenie funkcji tworzącej prawdopodobieństwa P_N zmiennej losowej N i funkcji charakterystycznej φ_X zmiennej losowej X , tzn.:

$$\varphi_S(t) = E[e^{itS}] = \sum_{n=0}^{\infty} E[e^{itX}]^n P(N=n) = P_N(\varphi_X(t)). \quad (3)$$

Z lematu Riemanna-Lebesgue'a (por. [Feller 1969]) wynika, że dla zmiennej losowej X_i funkcja charakterystyczna $\varphi_{X_i}(t)$ dąży do zera, gdy $|t| \rightarrow \infty$. Jednak dla rozkładów złożonych funkcja charakterystyczna zmiennej losowej S mająca postać:

$$\varphi_S(t) = \sum_{n=0}^{\infty} E[e^{itX}]^n P(N=n) \quad (4)$$

dąży do $p_0 = P(N=0)$, dla $|t| \rightarrow \infty$. Z tego powodu w dalszej części artykułu będzie rozpatrywana zmienna losowa $\tilde{S}$ określona w następujący sposób:

$$\tilde{S} = \sum_{i=1}^{\tilde{N}} X_i, \quad (5)$$

gdzie $P(\tilde{N}=n) = P(N=n)/(1-p_0)$ dla $n = 1, 2, \dots$. Wtedy dla tak ustalonej zmiennej losowej $\tilde{S}$ funkcja charakterystyczna mająca postać:

$$\varphi_{\tilde{S}}(t) = \frac{1}{1-p_0} (\varphi_S(t) - p_0) \quad (6)$$

dąży do zera dla $|t| \rightarrow \infty$. Zachodzi także następujący związek między rozkładami zmiennych losowych S i $\tilde{S}$, określonych odpowiednio w (1) i (5):

$$f_S(s) = (1-p_0)f_{\tilde{S}}(s). \quad (7)$$

Z funkcji charakterystycznej zmiennej losowej $\tilde{S}$ można wyznaczyć rozkład tej zmiennej, korzystając z odwrotnej transformaty Fouriera danej wzorem

$$f_{\tilde{S}}(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-its} \varphi_{\tilde{S}}(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-its} m_{\tilde{S}}(it) dt, \quad (8)$$

gdzie $m_{\tilde{S}}(z) \stackrel{\text{def}}{=} E[e^{z\tilde{S}}]$ jest funkcją tworzącą momenty zmiennej losowej $\tilde{S}$.

Wstawiając do wzoru (8) nową zmienną $T = it$, całkę można zapisać następująco:

$$\begin{aligned} f_{\tilde{S}}(s) &= \frac{1}{2\pi i} \int_{r-i\infty}^{r+i\infty} e^{-Ts} m_{\tilde{S}}(T) dT = \frac{1}{2\pi i} \int_{r-i\infty}^{r+i\infty} e^{-Ts} e^{\kappa_{\tilde{S}}(T)} dT = \\ &= f_{\tilde{S}}(s) = \frac{1}{2\pi i} \int_{r-i\infty}^{r+i\infty} e^{\kappa_{\tilde{S}}(T) - Ts} dT, \end{aligned} \quad (9)$$

gdzie $\kappa_{\tilde{S}}(z) \stackrel{\text{def}}{=} \log m_{\tilde{S}}(z)$ jest funkcją tworzącą kumulanty zmiennej losowej $\tilde{S}$.

Wartość całki (9) jest niezmiennicza, jeśli będzie się całkować wzdłuż linii równoległych do osi urojonej. W metodzie tej drogę całkowania wybiera się tak, by przechodziła przez punkt siodłowy. Wielkość r w wyrażeniu (9) jest zatem tak wyznaczona, aby funkcja podcałkowa osiągała minimum dla rzeczywistych T . W przestrzeni zespolonej punkt ten jest punktem siodłowym i istnieje w punkcie r_0 , dla którego pochodna funkcji podcałkowej jest równa zero, tzn. gdy

$$\kappa'_{\tilde{S}}(T) = s. \quad (10)$$

Wstawiając do funkcji podcałkowej we wzorze (9) za T wielkość $r_0 + iy$, otrzymuje się:

$$f_{\tilde{S}}(s) = \frac{1}{2\pi i} \int_{r_0-i\infty}^{r_0+i\infty} e^{\kappa_{\tilde{S}}(r_0+iy) - (r_0+iy)s} d(r_0 + iy) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\kappa_{\tilde{S}}(r_0+iy) - (r_0+iy)s} dy. \quad (11)$$

Przy obliczaniu wartości całki (11) funkcję $\kappa_{\tilde{S}}(r_0 + iy) - (r_0 + iy)s$ rozwija się w szereg Taylora w punkcie $r_0 + iy$, tzn.:

$$\begin{aligned} \kappa_{\tilde{S}}(r_0 + iy) - (r_0 + iy)s &= \kappa_{\tilde{S}}(r_0) - r_0 s + i(\kappa'_{\tilde{S}}(r_0) - s)y + \\ &\quad - \frac{1}{2} \kappa''_{\tilde{S}}(r_0) y^2 - \frac{1}{6} \kappa'''_{\tilde{S}}(r_0) i y^3 + \dots \end{aligned} \quad (12)$$

Uwzględniając pierwsze cztery składniki w (12) oraz warunek (10), tzn. $\kappa'_{\tilde{S}}(r_0) = s$, rozkład zmiennej $\tilde{S}$ można przybliżyć następującym wyrażeniem:

$$\begin{aligned} f_{\tilde{S}}(s) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\kappa_{\tilde{S}}(r_0+iy) - (r_0+iy)s} dy \approx \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\kappa_{\tilde{S}}(r_0) - r_0 s - \frac{1}{2} \kappa''_{\tilde{S}}(r_0) y^2} dy = \\ &= \frac{1}{2\pi} e^{\kappa_{\tilde{S}}(r_0) - r_0 s} \int_{-\infty}^{\infty} e^{-\frac{1}{2} \kappa''_{\tilde{S}}(r_0) y^2} dy. \end{aligned} \quad (13)$$

Ponieważ $\int_{-\infty}^{\infty} e^{-\frac{1}{2}\kappa_S''(r_0)y^2} dy = \sqrt{\frac{2\pi}{\kappa_S''(r_0)}}$, więc przybliżenie dla zmiennej $\tilde{S}$ przybiera postać:

$$f_{\tilde{S}}(s) \approx \frac{e^{\kappa_{\tilde{S}}(r_0) - r_0 s}}{\sqrt{2\pi\kappa_{\tilde{S}}''(r_0)}}. \quad (14)$$

Zatem ostatecznie z (6) i (7) dla zmiennej losowej S przybliżenie jest następujące:

$$f_S(s) \approx \frac{(e^{\kappa_S(r_0)} - p_0)e^{-r_0 s}}{\sqrt{2\pi\kappa_S''(r_0)}}, \quad (15)$$

gdzie r_0 jest punktem siodłowym wyznaczonym z równania $\kappa_S'(r) = s$.

4. Zastosowanie metody punktu siodłowego do aproksymacji rozkładów złożonych Poissona i ujemnie dwumianowego

Jeśli wielkość zagregowanych wypłat generowanych z portfela ma złożony rozkład Poissona z parametrem λ , wtedy punkt siodłowy r_0 wyznacza się z równania $\kappa_S'(r) = \lambda m_X'(r)$. Pozostałe wielkości w wyrażeniu (15) są równe:

- $p_0 = e^{-\lambda}$,
- $e^{\kappa_S(r_0)} = e^{\lambda(m_X(r_0)-1)}$,
- $\kappa_S''(r_0) = \lambda m_X''(r_0)$.

Gdy łączna wielkość świadczeń z portfela polis ma złożony rozkład ujemnie dwumianowy z parametrami α i p , punkt siodłowy r_0 jest rozwiązaniem równania

$\kappa_S'(r) = \frac{\alpha(1-p)m_X'(r)}{1-(1-p)m_X(r)}$, a pozostałe elementy wyrażenia (15) wynoszą:

- $p_0 = p^\alpha$,
- $e^{\kappa_S(r_0)} = \left(\frac{p}{1-m_X(r_0)(1-p)} \right)^\alpha$,
- $\kappa_S''(r_0) = \frac{\alpha(1-p)^2 m_X''(r_0)^2}{(1-(1-p)m_X(r_0))^2} + \frac{\alpha(1-p)m_X''(r_0)}{1-(1-p)m_X(r_0)}$.

Przykład 1

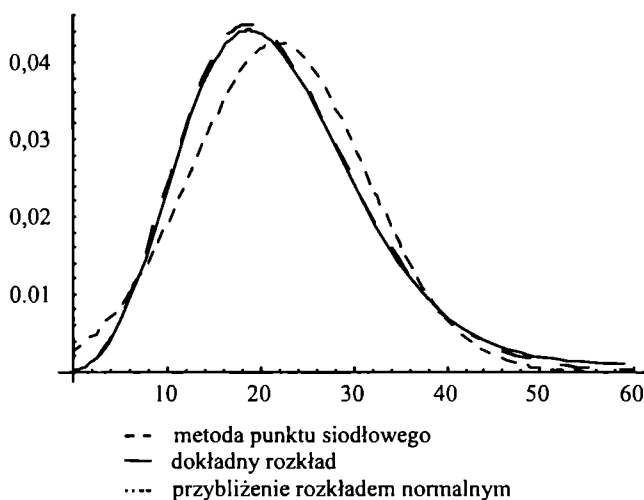
Niech liczba wypłat generowanych z portfela ma rozkład Poissona z wartością oczekiwaną 11, a wielkości wypłacanych odszkodowań mają rozkład wykładniczy o parametrze teta równym 0,5. W tabeli 1 przedstawiono wartości wyznaczone z

dokładnego rozkładu zagregowanej wielkości wypłat (oznaczone w tab. 1 przez D), za pomocą metody punktu siodłowego (MPS), oraz dla porównania wartości po zastosowaniu przybliżenia rozkładem normalnym (RN). Dokładny rozkład prawdopodobieństwa łącznych wypłat generowanych z tego portfela i zastosowane metody przedstawiono na rys. 1.

Tabela 1. Wartości funkcji w zależności od zastosowanej metody

s	10	20	30	40	50	60
D	0,0232824	0,0437935	0,024364	0,0070548	0,0013353	0,0001861
MPS	0,0238859	0,0446021	0,024729	0,0071461	0,0013507	0,0001881
RN	0,0187645	0,0415718	0,029562	0,0067479	0,0004944	0,0000116

Źródło: opracowanie własne.



Rys. 1. Dokładna, wyznaczona metodą punktu siodłowego oraz przybliżona rozkładem normalnym, funkcja gęstości złożonego rozkładu Poissona z $\lambda = 11$ i $X \sim \text{Exp}(0,5)$

Źródło: opracowanie własne.

Przykład 2

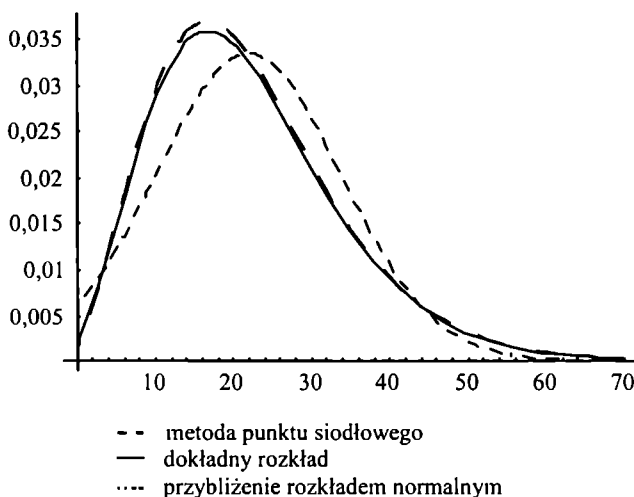
Niech liczba szkód generowanych z portfela ma rozkład ujemnie dwumianowy z parametrami $\alpha = 9$ i $p = 9/20$, a wielkości świadczeń mają rozkład wykładniczy o parametrze równym 0.5. W tabeli 2 przedstawiono wartości wyznaczone z do-

Tabela 2. Wartości funkcji w zależności od zastosowanej metody

s	20	30	40	50	60	70
D	0,0346171	0,02119	0,003854	0,003059	0,000882	0,0002254
MPS	0,0355379	0,02166	0,003923	0,003112	0,000896	0,0002288
RN	0,0330354	0,02673	0,003088	0,002110	0,000205	0,0000099

Źródło: opracowanie własne.

kładnego rozkładu (D), za pomocą metody punktu siodłowego (MPS), oraz wartości po zastosowaniu przybliżenia rozkładem normalnym (RN). Dokładny rozkład prawdopodobieństwa łącznych wypłat generowanych z tego portfela i zastosowane metody przedstawiono na rys. 2.



Rys. 2. Dokładna, wyznaczona metodą punktu siodłowego oraz przybliżona rozkładem normalnym, funkcja gęstości złożonego rozkładu ujemnie dwumianowego z parametrami $r = 9$ i $p = 9/20$ oraz $X \sim \text{Exp}(0,5)$

Źródło: opracowanie własne.

W obu przykładach wartości rozkładu prawdopodobieństwa uzyskane za pomocą metody punktu siodłowego aproksymacji rozkładu złożonego Poissona i ujemnie dwumianowego są bliższe wartościom dokładnego rozkładu prawdopodobieństwa zagregowanej wielkości wypłat generowanych z portfela polis niż wartości otrzymane przy przybliżaniu rozkładem symetrycznym. Jest to związane z tym, że w praktyce rozkład zagregowanych wielkości roszczeń wypłacanych przez firmy ubezpieczeniowe jest asymetryczny i przybliżanie go rozkładem normalnym nie jest słuszne.

Literatura

- Daniels H.E., *Exact Saddlepoint Approximation*, „Biometrika” 1980, vol. 67, no. 1, 59-63.
 Daniels H.E., *Saddlepoint Approximation in Statistics*, „The Annals of Mathematical Statistics” 1954 vol. 25, no. 4, 631-650.
 Feller W., *Wstęp do rachunku prawdopodobieństwa. Tom II*, PWN, Warszawa 1969.
 Huzurbazar S., *Practical Saddlepoint Approximation*, „The American Statistician” 1999 vol. 53, no. 3, 225-232.

Jensen J.L., *Uniform Saddlepoint Approximation*, Adv. Appl. Prob. 1988, 20, 622-634.

Lugannani R., Stephen R., *Saddle Point Approximation for the Distribution of the Sum of Independent Random Variables*, Adv. Appl. Prob. 1980, 12, 475-490.

SADDLEPOINT METHOD FOR PROBABILITY DISTRIBUTION APPROXIMATION

Summary

When the exact distribution cannot be conveniently obtained, we can use an approximation. In the paper the saddlepoint approximation is described. This method is used for approximating density of the aggregated amount of claims, where the individual claim amounts generated from the insurance portfolio are independently and identically distributed.