

Spis treści

| | |
|------------------------|----|
| Akronimy | 2 |
| Wprowadzenie | 4 |
| Bibliografia | 10 |

Akronimy

| | |
|---------------|-----------------------------------------------|
| AGLM | Augmented Generalized Linear Model |
| CGLM | Conjugate Generalized Linear Model |
| DGLM | Double Generalized Linear Model |
| EDM | Exponential Dispersion Model |
| GLM | Generalized Linear Model |
| GLMM | Generalized Linear Mixed Model |
| GP | Generalized Poisson |
| HGLM | Hierarchical Generalized Linear Model |
| H-IWSL | Hierarchical Iterative Weighted Least Squares |
| IWSL | Iterative Weighted Least Squares |
| LMM | Linear Mixed Model |
| MBP | Minimum Bias Procedure |
| MML | Marginal Maximum Likelihood |
| MLE | Maximum Likelihood Estimation |
| NB | Negative-Binomial |
| NLM | Non-Linear Model |
| NLMM | Non-Linear Mixed Model |
| ZIGP | Zero-Inflation Generalized Poisson |
| ZINB | Zero Inflation Negative-Binomial |
| ZIP | Zero Inflation Poisson |

Wprowadzenie

Większość z nas korzysta z usług zakładów ubezpieczeń majątkowych ubezpieczając samochód, mieszkanie czy sprzęt elektroniczny (przedmioty ubezpieczenia), stając się ubezpieczonym (podmiotem ubezpieczenia). Każdy ubezpieczony musiał więc w tym celu odbyć spotkanie z agentem ubezpieczeniowym, a mówiąc inaczej każdy ubezpieczony uczestniczył w procesie, który można nazwać specjalistycznie procesem wyceny ryzyka, czyli procesem wyznaczenia wartości składki. Efektem procesu jest zawarta polisa ubezpieczeniowa, dostarczająca zakładowi ubezpieczeń informacji niezbędnych do wyznaczenia wartości składki, możliwego dzięki założeniu, iż dana polisa należy do grupy polis, w której zakłada się w sposób uproszczony jednorodność. Jednorodność ta jest rozumiana w ten sposób, iż każdej polisie w grupie jest przyporządkowywana taka sama wartość składki. Grupę polis nazywamy dalej portfelem. Istotnym problemem zakładu jest więc podział portfela polis zawierającego polisy odnośnie do jednego produktu (produkt to np. ubezpieczenie komunikacyjne OC, ubezpieczenie nieruchomości itd.) na odpowiednie grupy. Proces wyodrębniania tych grup i wyznaczania w każdej grupie wartości składki jest ogólnie nazywany *taryfikacją* (ang. *ratemaking*). Oczywiście w najprostszym przypadku portfel wszystkich polis dla danego produktu można potraktować jako portfel jednorodny i przypisać każdej polisie jednakową składkę. Czy kierowca pojazdu klasy ekonomicznej jest jednak skłonny w ubezpieczeniu AC zapłacić taką wartość składki, która pokryje wartość wypadku samochodu luksusowego? Czy kierowca jeżdżący

ostrożnie jest skłonny w ubezpieczeniu OC zapłacić taką samą składkę jak kierowca brawurowy? Czy właściciel domu stojącego na chronionym osiedlu zamkniętym jest skłonny zapłacić taką samą składkę jak właściciel domu znajdującego się w niebezpiecznej dzielnicy? Odpowiedź na powyższe pytania brzmi: raczej nie, dlatego obecnie zakłady dążą do tego, aby proces wyceny ryzyka był jak najbardziej spersonalizowany. Efektem tego jest podział portfela na coraz większą liczbę grup o tej samej składce. W szczególności może to doprowadzić do sytuacji, iż w ramach jednego produktu każda polisa będzie generowała inną wartość składki. Przykładem jest polisa w ubezpieczeniu komunikacyjnym OC/AC zawierana na zasadzie *Usage-based*. Oznacza to, iż ubezpieczony godzi się na instalację urządzenia elektronicznego w samochodzie pozwalającego na rzeczywisty pomiar różnego rodzaju czynników pomocnych do spersonalizowanej wyceny ryzyka. Czynniki te to przykładowo faktyczna liczba przejechanych kilometrów czy obszar, po którym porusza się pojazd. Zakład ubezpieczeń wyznacza wtedy wartość składki na zasadzie PAYD, inaczej *Pay-As-You-Drive* (zob. [Desyllas i Sako, 2013]). System zbierania danych jest nazywany *telematyką* (ang. *telematics*). Na rynku polskim jednym z przykładów jest rozwiązanie *Mobileye* (zob. <http://www.mobileye.com/pl-pl/rynki/firmy-ubezpieczeniowe/>).

Dążenie do coraz większej personalizacji składki ubezpieczeniowej powoduje, iż zakłady ubezpieczeń posiadają coraz więcej informacji, które mogą być brane pod uwagę w procesie wyceny ryzyka. Rodzi się zatem problem w jaki sposób wykorzystać zebrane dane. W obecnej praktyce, w masowych portfelach polis, wyznaczenie wartości składki dla polisy opiera się na modelu statystycznym, w którym estymuje się jego parametry czerpiąc informacje z danych zawartych w portfelu polis. Ujmując to językiem statystycznym, czerpie się informacje z próby. Modele te są modelami regresyjnymi klasy GLM (ozn. Generalized Linear Model (GLM)), czyli uogólnionymi modelami liniowymi. Wraz z coraz bardziej szczegółowymi danymi, jak również rozwojem algorytmów numerycznych, uzasadnione jest modyfikowanie stosowanych obecnie modeli statystycznych.

Celem niniejszej pracy jest zatem zaproponowanie modeli statystycznych do taryfikacji uwzględniających obecną specyfikę masowych portfeli ubezpieczeniowych traktowanych jako zbiory obserwacji – próby statystyczne. Formalnie proponowane modele są regresyjnymi modelami mieszanymi klasy HGLM (ozn. Hierarchical Generalized Linear Model (HGLM)) oraz NLMM (ozn. Non-Linear Mixed Model (NLMM)). Każdy model prezentowany jest w następujący sposób:

1. Określenie postaci modelu.
2. Wyznaczenie charakterystyk zmiennej losowej występującej w modelu.
3. Określenie metody estymacji parametrów modelu.
4. Scharakteryzowanie implementacji komputerowej modelu.

W nadawaniu nazwy konkretnego modelu powszechną praktyką w modelowaniu statystycznym jest używanie akronimów, które w miarę jasno charakteryzują postać modelu. Najczęściej z nazwy wynika również metoda estymacji parametrów modelu. Taka praktyka powoli przyjmuje się również w zagadnieniu taryfikacji (por. [Frees i in., 2014]). W pracy posłużono się więc następującą nomenklaturą dotyczącą nazwy modelu: *Klasa_modelu* – *Klasa_rozkadu_zmiennej_objasnianej* oraz użyto akronimów angielskich (por. spis akronimów).

Istotnym wkładem w powstanie monografii są rozwiązania informatyczne pozwalające na bezpośrednią aplikację skonstruowanych modeli w praktyce ubezpieczeniowej jako rozwinięcie modeli stosowanych obecnie. Do większości obliczeń został zastosowany program komputerowy R [R Core Team, 2013]. Wybór tego oprogramowania wyniknął z jego szerokich możliwości obliczeniowych w obszarze modelowania statystycznego, jak również powszechnej darmowej dostępności w ramach publicznej licencji GNU/GPL. W większości przypadków zostały wykorzystane standardowe procedury oraz funkcje zawarte w obu programach. Pewne nowe zastosowania wymagały jednak przeprowadzenia implementacji komputerowej algorytmów obliczeniowych, powodując, iż wyniki prac badawczych są obecnie dostępne szerokiemu gronu odbiorców.

Monografia składa się z 5 rozdziałów. Pierwszy jest rozdziałem opisowym, przedstawiającym koncepcję rozszerzenia obecnego podejścia modeli GLM w taryfikacji do stosowania modeli mieszanych. W pierwszej kolejności opisano dwa zagadnienia poruszane w monografii: taryfikację oraz wyznaczanie wartości składki czystej dla pojedynczego ryzyka. Uwagę skupiono na zdefiniowaniu pojęć ubezpieczeniowych oraz odpowiadających im pojęć statystycznych w celu ujednoczenia procesu taryfikacji proponowanego w pracy. Cały rozdział jest poświęcony poprawnemu zdefiniowaniu procesu taryfikacji, gdyż w literaturze przedmiotu autorzy stosują jego różną nomenklaturę. Aby to osiągnąć, omówiono różnorodne aspekty procesu taryfikacji związane w sensie statystycznym głównie z modelami regresyjnymi. Dalej podjęto temat celowości zastosowania modeli mieszanych w procesie taryfikacji. W tym celu przybliżono pojęcie modelu mieszanego w kontekście taryfikacji. Rozdział kończy charakterystyka danych zawartych w portfelach ubezpieczeniowych. Specyfika tych danych bezpośrednio wpływa na aplikacyjność różnych klas modeli mieszanych. W 2 oraz 3 rozdziale pracy scharakteryzowano podstawy teoretyczne dotyczące konstrukcji oraz estymacji parametrów modeli GLM oraz modeli mieszanych. Oba rozdziały mają charakter metodologiczny, jednak zawierają jedynie ten fragment metodologii, który znajduje zastosowanie w taryfikacji. Rozdział 2 zawiera opis rozkładów prawdopodobieństwa zmiennych losowych mających zastosowanie w procesie taryfikacji. Opisano w nim głównie dyspersyjną rodzinę rozkładów wykładniczych, podrodzinę rozkładów Tweedie, uogólniony rozkład Poissona oraz złożony rozkład Poissona. W rozdziale 3 podjęto tematykę dotyczącą samej metodologii związanej z modelami mieszanymi. Uwagę skupiono na modelach nieliniowych, gdyż de facto tylko taka klasa modeli znajduje zastosowanie w poruszanych zagadnieniach ubezpieczeniowych. Popularne modele liniowe nie są prawidłowe, chociażby ze względu na multiplikatywny charakter procesu taryfikacji. Przedstawiono w kolejności poszczególne klasy modeli regresyjnych począwszy od modelu najprostszego, jakim jest klasyczny model nieliniowy z efektami stałymi, dalej przechodząc do modeli

uogólnionych i ostatecznie wprowadzając efekty losowe, uzyskując modele mieszane. Opis każdego modelu zawiera: założenia, postać modelu oraz metodę estymacji parametrów modelu. Rozdziały 4 i 5 poświęcono w całości modelowaniu w procesie taryfikacji. Punktem wyjścia jest prezentacja obecnej praktyki ubezpieczeniowej, jaką są modele GLM. Dalej wprowadzono modele należące do szerszej klasy modeli, a uwzględniające efekty występujące w portfelach ubezpieczeniowych nieuwzględniane w modelach GLM. Dotyczy to głównie występowania dużej liczby polis, dla których w okresie ubezpieczenia nie zaszła żadna szkoda. Konstruowane są więc modele, które nie są formalnie modelami GLM, ale uwzględniają wskazane w rozdziale 1 efekty pojawiające w danych ubezpieczeniowych. Następnie wprowadzono modele mieszane jako rozszerzenie modeli GLM. Istotą stosowania tych modeli jest uwzględnienie innego efektu występującego w portfelach polis, jakim są różnego rodzaju zależności. Modele mieszane pozwalają na modelowanie tych zależności. Każdy podrozdział obu rozdziałów zawiera w pierwszej części teoretyczny opis modelu statystycznego, w którym zdefiniowano założenia komponentów danego modelu, głównie rozkłady prawdopodobieństwa zmiennych losowych. Następnie wyprowadzono odpowiednie wielkości istotne w taryfikacji, takie jak wartość oczekiwana, wariancja czy kowariancja zmiennych losowych. Na końcu rozważono estymację parametrów modelu oraz aspekty programistyczne związane z jej praktycznym przeprowadzeniem. W drugiej części każdego podrozdziału przedstawiono przykład empiryczny, który ma na celu zobrazowanie jak działa konkretny model statystyczny. W efekcie uzyskano oszacowane wartości parametrów rozważanego modelu. W przykładach wykorzystano dwa portfele polis pochodzące z polskiego rynku ubezpieczeniowego, oznaczane jednolicie w całej pracy jako `danePL` oraz `danePLpanel`. Jeden portfel zawiera dane w układzie przekrojowym, natomiast drugi dane w układzie przekrojowo-czasowym.

Suplementem do pracy jest strona internetowa¹, na której są dostępne kody procedur powstałych na użytek części empirycznej pracy. We wszystkich przykładach empirycznych wykorzystano dane z działającego obecnie zakładu ubezpieczeń majątkowych. Ze względu na klauzulę poufności danych, baza danych, zawierająca szczegółowe informacje z portfela polis komunikacyjnych AC, nie może jednak być udostępniona. Kody procedur są więc rozszerzone o przypadki portfeli ogólnodostępnych w literaturze przedmiotu. W dodatku C pracy przedstawiono więc dodatkowo dwa portfele, nazwane `dataOhlsson` oraz `motorins`. W celu zwiększenia użyteczności sfery informatycznej, na potrzeby pracy powstał także pakiet `insuranceData` do pobrania z serwera CRAN². Dzięki temu czytelnik łatwo może prześledzić działanie algorytmów, analizować wyniki czy porównywać modele. Dla pewnej klasy modeli wykorzystano dodatkowo komercyjny program komputerowy SAS. Wynikało to z faktu, iż program ten posiada szeroko rozbudowaną procedurę o nazwie NLMIXED, pozwalającą na elastyczne definiowanie pewnej klasy modeli.

Autorka składa gorące podziękowania recenzentowi Panu dr hab. Stanisławowi Wanatowi, który poprzez wnikliwą analizę zawartej w monografii tematyki przyczynił się do poprawy jakości pracy. Dodatkowo podziękowania kieruje w stronę Kierownika Katedry Metod Statystyczno-Matematycznych w Ekonomii Uniwersytetu Ekonomicznego w Katowicach Pana prof. dr hab. Włodzimierza Szkutnika oraz pozostałych pracowników za merytoryczne uwagi wygłaszane na seminariach naukowych.

Projekt został sfinansowany ze środków Narodowego Centrum Nauki.

¹ <http://web.ue.katowice.pl/woali>

² <http://cran.r-project.org/web/packages/insuranceData/index.html>

Bibliografia

- Antonio K., Beirlant J. (2007), *Actuarial statistics with generalized linear mixed models*, „Insurance: Mathematics and Economics”, 40(1), s. 58–76 /-/.
- Antonio K., Beirlant J., Hoedemakers T., Verlaak R. (2006), *Lognormal mixed models for reported claims reserves*, „North American Actuarial Journal”, 10(1), s. 30–48 /-/.
- Antonio K., Valdez E.A. (2012), *Statistical concepts of a priori and a posteriori risk classification in insurance*, „ASTA Advances in Statistical Analysis”, 96(2), s. 187–224 /-/.
- Beirlant J., Goegebeur Y., Segers J., Teugels J. (2006), *Statistics of extremes: theory and applications*, John Wiley & Sons /-/.
- Biecek P. (2011), *Analiza danych z programem R: modele liniowe z efektami stałymi, losowymi i mieszanymi*, Wydawnictwo Naukowe PWN /-/.
- Bjørnstad J.F. (1996), *On the generalization of the likelihood function and the likelihood principle*, „Journal of the American Statistical Association”, 91(434), s. 791–806 /-/.
- Booth J.G., Hobert J.P. (1999), *Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm*, „Journal of the Royal Statistical Society: Series B (Statistical Methodology)”, 61(1), s. 265–285 /-/.
- Boucher J., Denuit M. (2006), *Fixed versus random effects in Poisson regression models for claim counts: A case study with motor insurance*, „Astin

- Bulletin”, 36(1), s. 285 /-/.
- Boucher J.P., Denuit M., Guillén M. (2007), *Risk Classification for Claim Counts: A Comparative Analysis of Various Zeroinflated Mixed Poisson and Hurdle Models*, „North American Actuarial Journal”, 11(4), s. 110–131 /-/.
- Boucher J.P., Denuit M., Guillén M. (2009), *Number of Accidents or Number of Claims? An Approach with Zero-Inflated Poisson Models for Panel Data*, „Journal of Risk and Insurance”, 76(4), s. 821–846 /-/.
- Boucher J.P., Guillén M. (2009), *A survey on models for panel count data with applications to insurance*, „RACSAM-Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas”, 103(2), s. 277–294 /-/.
- Boucher J.P., Inoussa R. (2014), *A posteriori ratemaking with panel data*, „ASTIN Bulletin”, DOI: <http://dx.doi.org/10.1017/asb.2014.11> (About DOI), Published online: 23 April 2014, s. 1–26 /-/.
- Breslow N.E., Clayton D.G. (1993), *Approximate inference in generalized linear mixed models*, „Journal of the American Statistical Association”, 88(421), s. 9–25 /-/.
- Van den Broek J. (1995), *A score test for zero inflation in a Poisson distribution*, „Biometrics”, s. 738–743 /-/.
- Broström G., Holmberg H. (2011), *Generalized linear models with clustered data: Fixed and random effects models*, „Computational Statistics & Data Analysis”, 55(12), s. 3123–3134 /-/.
- Brown H., Prescott R. (2006), *Applied mixed models in medicine*, John Wiley & Sons /-/.
- Bühlmann H., Gisler A. (2005), *A course in credibility theory and its applications*, Springer /-/.
- Cieślak B. (2004), *System bonus-malus jako narzędzie konkurencji na rynku ubezpieczeń komunikacyjnych, w: Metody ilościowe w naukach ekonomicznych*, „Oficyna Wydawnicza SGH, Warszawa”, s. 31–51 /-/.
- Cizek P., Härdle W.K., Weron R. (2005), *Statistical tools for finance and*

- insurance*, Springer /-/.
- Consul P., Famoye F. (1992), *Generalized Poisson regression model*, „Communications in Statistics-Theory and Methods”, 21(1), s. 89–109 /-/.
- Consul P.C., Jain G.C. (1973), *A generalization of the Poisson distribution*, „Technometrics”, 15(4), s. 791–799 /-/.
- Czado C., Erhardt V., Min A., Wagner S. (2007), *Zero-inflated generalized Poisson models with regression effects on the mean, dispersion and zero-inflation level applied to patent outsourcing rates*, „Statistical Modelling”, 7(2), s. 125–153 /-/.
- Davidian M., Giltinan D.M. (1995), *Nonlinear models for repeated measurement data*, tom 62, CRC Press /-/.
- De Jong P., Heller G.Z. (2008), *Generalized linear models for insurance data*, Cambridge University Press /-/.
- Dean C., Lawless J., Willmot G. (1989), *A mixed poisson-inverse-gaussian regression model*, „Canadian Journal of Statistics”, 17(2), s. 171–181 /-/.
- Demidenko E. (2005), *Mixed models: theory and applications*, tom 518, John Wiley & Sons /-/.
- Denuit M., Maréchal X., Pitrebois S., Walhin J.F. (2007), *Actuarial modelling of claim counts: Risk classification, credibility and bonus-malus systems*, Wiley. com /-/.
- Desyllas P., Sako M. (2013), *Profiting from business model innovation: Evidence from Pay-As-You-Drive auto insurance*, „Research Policy”, 42(1), s. 101–116 /5/.
- Dimakos X.K., Di Rattalma A.F. (2002), *Bayesian premium rating with latent structure*, „Scandinavian Actuarial Journal”, 2002(3), s. 162–184 /-/.
- Dionne G., Vanasse C. (1989), *A generalization of automobile insurance rating models: the negative binomial distribution with a regression component*, „Astin Bulletin”, 19(2), s. 199–212 /-/.
- Dionne G., Vanasse C. (1992), *Automobile insurance ratemaking in the presence of asymmetrical information*, „Journal of Applied Econometrics”,

- 7(2), s. 149–165 /-/.
- Dunn P.K., Smyth G.K. (2005), *Series evaluation of Tweedie exponential dispersion model densities*, „Statistics and Computing”, 15(4), s. 267–280 /-/.
- Dunn P.K., Smyth G.K. (2008), *Evaluation of Tweedie exponential dispersion model densities by Fourier inversion*, „Statistics and Computing”, 18(1), s. 73–86 /-/.
- Famoye F., Singh K.P. (2006), *Zero-inflated generalized Poisson regression model with an application to domestic violence data*, „Journal of Data Science”, 4(1), s. 117–130 /-/.
- Famoye F., Wulu J., Singh K.P. (2004), *On the generalized Poisson regression model with an application to accident data*, „Journal of Data Science”, 2(2004), s. 287–295 /-/.
- Frątczak E. (2013), *Zaawansowane modele analiz statystycznych*, Oficyna Wydawnicza SGH /-/.
- Frees E.W. (2009), *Regression modeling with actuarial and financial applications*, Cambridge University Press /-/.
- Frees E.W., Derrig R.A., Meyers G. (2014), *Predictive modeling applications in actuarial science*, Cambridge University Press /6/.
- Gnot S. (1991), *Estymacja komponentów wariancyjnych w modelach liniowych*, WNT, Warszawa /-/.
- Gray R.J., Pitts S.M. (2012), *Risk Modelling in General Insurance: From Principles to Practice*, Cambridge University Press /-/.
- Greenwood M., Yule G.U. (1920), *An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents*, „Journal of the Royal Statistical Society”, s. 255–279 /-/.
- Gumpertz M.L., Pantula S.G. (1992), *Nonlinear regression with variance components*, „Journal of the American Statistical Association”, 87(417), s. 201–209 /-/.
- Haberman S., Renshaw A.E. (1996), *Generalized linear models and actuarial*

- science*, „The Statistician”, s. 407–436 /-/.
- Joe H., Zhu R. (2005), *Generalized Poisson distribution: the property of mixture of Poisson and comparison with negative binomial distribution*, „Biometrical Journal”, 47(2), s. 219–229 /-/.
- Jørgensen B. (1986), *Some properties of exponential dispersion models*, „Scandinavian Journal of Statistics”, s. 187–197 /-/.
- Jørgensen B. (1987), *Exponential dispersion models*, „Journal of the Royal Statistical Society. Series B (Methodological)”, s. 127–162 /-/.
- Jørgensen B., Paes De Souza M.C. (1994), *Fitting Tweedie’s compound Poisson model to insurance claims data*, „Scandinavian Actuarial Journal”, 1994(1), s. 69–93 /-/.
- Karlis D. (2005), *EM algorithm for mixed Poisson and other discrete distributions*, „Astin bulletin”, 35(1), s. 3–24 /-/.
- Karlis D., Xekalaki E. (2005), *Mixed poisson distributions*, „International Statistical Review”, 73(1), s. 35–58 /-/.
- Kowalczyk P., Poprawska E., Ronka-Chmielowiec W. (2006), *Metody aktuarialne*, Wydawnictwo Naukowe PWN, Warszawa /-/.
- Laird N.M., Ware J.H. (1982), *Random-effects models for longitudinal data*, „Biometrics”, s. 963–974 /-/.
- Lambert D. (1992), *Zero-inflated Poisson regression, with an application to defects in manufacturing*, „Technometrics”, 34(1), s. 1–14 /-/.
- Lawless J.F. (1987), *Negative binomial and mixed Poisson regression*, „Canadian Journal of Statistics”, 15(3), s. 209–225 /-/.
- Lee Y., Nelder J.A. (1996), *Hierarchical generalized linear models*, „Journal of the Royal Statistical Society. Series B (Methodological)”, s. 619–678 /-/.
- Lee Y., Nelder J.A. (2001), *Hierarchical generalised linear models: a synthesis of generalised linear models, random-effect models and structured dispersions*, „Biometrika”, 88(4), s. 987–1006 /-/.
- Lee Y., Nelder J.A., Pawitan Y. (2006), *Generalized linear models with random effects: unified analysis via H-likelihood*, CRC Press /-/.

- Lemaire J. (1995), *Bonus-malus systems in automobile insurance*, tom 19, Springer /-/.
- Lindsey J.K. (1997), *Applying generalized linear models*, Springer /-/.
- Littell R.C. (2006), *SAS for mixed models*, SAS institute /-/.
- Marzec J. (2008), *Bayesowskie modele zmiennych jakościowych i ograniczonych w badaniach niespłacalności kredytów*, „Zeszyty Naukowe/Uniwersytet Ekonomiczny w Krakowie. Seria Specjalna, Monografie”, (188) /-/.
- McCullagh P., Nelder J. (1989), *Generalized linear models. Second Edition*, „Monographs on statistics and applied probability.” /-/.
- McCulloch C.E. (2006), *Generalized linear mixed models*, Wiley Online Library /-/.
- Mildenhall S.J. (1999), *A systematic relationship between minimum bias and generalized linear models*, w: *Proceedings of the Casualty Actuarial Society*, tom 86, s. 393–487 /-/.
- Molenberghs G., Verbeke G. (2005), *Models for discrete longitudinal data*, Springer /-/.
- Murphy K.P., Brockman M.J., Lee P.K. (2000), *Using generalized linear models to build dynamic pricing systems*, w: *Casualty Actuarial Society Forum, Winter*, s. 107–139 /-/.
- Nelder J., Verrall R. (1997), *Credibility theory and generalized linear models*, „Astin Bulletin”, 27, s. 71–82 /-/.
- Nelder J., Wedderburn W. (1972), *Generalized Linear Models*, „Journal of the Royal Statistical Society. Series A (General)”, 135(3), s. 370-384 /-/.
- Nelder J.A., Mead R. (1965), *A simplex method for function minimization*, „The Computer Journal”, 7(4), s. 308–313 /-/.
- Ohlsson E., Johansson B. (2010), *Non-life insurance pricing with generalized linear models*, Springer /-/.
- Ostasiewicz W., Dębicka J. (2004), *Składki i ryzyko ubezpieczeniowe: modelowanie stochastyczne*, Wydawnictwo Akademii Ekonomicznej, Wrocław /-/.

- Otto W. (2013), *Ubezpieczenia majątkowe—Część I—Teoria ryzyka*, wydanie drugie /-/.
- Panjer H.H. (1981), *Recursive evaluation of a family of compound distributions*, „Astin Bulletin”, 12(1), s. 22–26 /-/.
- Pawitan Y. (2001), *In all likelihood: statistical modelling and inference using likelihood*, Oxford University Press /-/.
- Pinheiro J.C., Bates D.M. (1995), *Approximations to the log-likelihood function in the nonlinear mixed-effects model*, „Journal of Computational and Graphical Statistics”, 4(1), s. 12–35 /-/.
- Pinheiro J.C., Bates D.M. (2000), *Linear mixed-effects models: basic concepts and examples*, Springer /-/.
- Pinquet J. (2000), *Experience rating through heterogeneous models*, w: *Handbook of Insurance*, Springer, s. 459–500 /-/.
- Pitkanen P. (1975), *Tariff theory*, „International Journal for Actuarial Studies in Non-life Insurance and Risk Theory”, s. 204 /-/.
- Podgórska M. (2008), *System bonus-malus sprawiedliwy w sensie przejść między klasami*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, (1 (1201)), s. 206–231 /-/.
- R Core Team (2013), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria /6/.
- Ridout M., Demétrio C.G., Hinde J. (1998), *Models for count data with many zeros*, w: *Proceedings of the XIXth International Biometric Conference*, tom 19, s. 179–192 /-/.
- Rigby R., Stasinopoulos D. (2005), *Generalized additive models for location, scale and shape*, „Journal of the Royal Statistical Society: Series C (Applied Statistics)”, 54(3), s. 507–554 /-/.
- Ronka-Chmielowiec W. (1997), *Ryzyko w ubezpieczeniach—metody oceny*, Wydawnictwo Akademii Ekonomicznej, Wrocław /-/.
- Sarma K.S. (2013), *Predictive modeling with SAS Enterprise Miner: practical solutions for business applications*, SAS Institute /-/.
- Sichel H.S. (1975), *On a distribution law for word frequencies*, „Journal of

- the American Statistical Association”, 70(351a), s. 542–547 /-/.
- Smyth G.K. (1989), *Generalized linear models with varying dispersion*, „Journal of the Royal Statistical Society. Series B (Methodological)”, s. 47–60 /-/.
- Smyth G.K., Jørgensen B. (2002), *Fitting Tweedie’s compound Poisson model to insurance claims data: dispersion modelling*, „Astin Bulletin”, 32(1), s. 143–157 /-/.
- Sobiecki D. (2013), *Dwustopniowe modelowanie składki za ubezpieczenie komunikacyjne OC*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, (312), s. 116–134 /-/.
- Straub E. (1988), *Non-Life Insurance Mathematics*, Springer /-/.
- Szymańska A. (2008), *Wybrane miary efektywności systemów bonus-malus ubezpieczeń komunikacyjnych OC*, „Prace Naukowe Akademii Ekonomicznej we Wrocławiu”, (1197), s. 428–435 /-/.
- Tweedie M. (1984), *An index which distinguishes between some important exponential families*, w: *Statistics: Applications and New Directions: Proc. Indian Statistical Institute Golden Jubilee International Conference*, s. 579–604 /-/.
- Venables W.N., Ripley B.D. (2002), *Modern applied statistics with S*, Springer /-/.
- Verbeke G., Lesaffre E. (1996), *A linear mixed-effects model with heterogeneity in the random-effects population*, „Journal of the American Statistical Association”, 91(433), s. 217–221 /-/.
- Vonesh E.F., Chinchilli V.M. (1997), *Linear and nonlinear models for the analysis of repeated measurements*, tom 1, CRC press /-/.
- Wanat S. (2012), *Modele zależności w agregacji ryzyka ubezpieczyciela*, „Zeszyty Naukowe/Uniwersytet Ekonomiczny w Krakowie. Seria Specjalna, Monografie”, (211) /-/.
- Werner G., Modlin C. (2010), *Basic ratemaking*, w: *Casualty Actuarial Society* /-/.
- West B.T., Welch K.B., Galecki A.T. (2006), *Linear mixed models: a prac-*

-
- tical guide using statistical software*, CRC Press /-/.
- Winkelmann R. (2003), *Econometric analysis of count data*, Springer /-/.
- Wolfinger R. (1993), *Laplace's approximation for nonlinear mixed models*, „*Biometrika*”, 80(4), s. 791–795 /-/.
- Wolny-Dominiak A. (2012), *Modeling of claim counts using data mining procedures in R CRAN*, w: *Proceedings of 30th International Conference Mathematical Methods in Economics*, s. 980–985 /-/.
- Wolny-Dominiak A. (2013), *Zero-inflated claim count modeling and testing—a case study*, „*Ekonometria*”, (1 (39)), s. 144–151 /-/.
- Yang Z., Hardin J.W., Addy C.L. (2009), *Testing overdispersion in the zero-inflated Poisson model*, „*Journal of Statistical Planning and Inference*”, 139(9), s. 3340–3353 /-/.
- Yip K.C., Yau K.K. (2005), *On modeling claim frequency data in general insurance with extra zeros*, „*Insurance: Mathematics and Economics*”, 36(2), s. 153–163 /-/.
- Żądło T. (2014), *On the Prediction of the Subpopulation Total Based on Spatially Correlated Longitudinal Data*, „*Mathematical Population Studies*”, 21(1), s. 30–44 /-/.
- Zhang Y. (2013), *Likelihood-based and bayesian methods for tweedie compound poisson linear mixed models*, „*Statistics and Computing*”, 23(6), s. 743–757 /-/.