

Prof. dr hab. inż. Ewa Drabik

Warszawa 25.10.2021

Wydział Zarządzania

Politechnika Warszawska

Ul. Narbutta 85

02 – 524 Warszawa

**Recenzja pracy doktorskiej mgr Norberta Paska zatytułowanej
„Zastosowanie uogólnionych liniowych modeli mieszanych w taryfikacji
ubezpieczeń komunikacyjnych” dla Rady Kolegium Analiz Ekonomicznych
SGH**

Przedłożona mi do recenzji rozprawa doktorska dotyczy porównania metod stosowanych w ramach taryfikacji ubezpieczeń majątkowych służących do predykcji częstości szkód. W celach porównawczych wykorzystywane są liniowe oraz mieszane modele ekonometryczne uwzględniające problemy takie jak nadmierne rozproszenie danych, nadmierna liczba zer w rozkładzie zmiennej zależnej, czy też zależności pomiędzy zmiennymi określonego modelu. Jako wkład własny Doktorant deklaruje zaproponowanie uogólnionego modelu mieszanego z ujemnym rozkładem dwumianowym. Ponadto nowatorskie jest wprowadzenie do modelu losowych zmiennych uwzględniających działania agenta ubezpieczeniowego oraz menedżera co prowadzi do tzw. modelu wielopoziomowego. Głównym celem jest oprócz analizy porównawczej, „lepsze” dopasowanie parametrów modelu do danych empirycznych. Hipotezy badawcze weryfikowane w pracy pokrywają się, w zasadzie, z celem pracy. Dodać można jedynie, że Doktorant „wnioskuje”, iż błędy predykcji w modelu wielopoziomowym są niższe niż w przypadku modelu jednopozomowego.

Rozprawa wraz ze spisem treści, tabel i rysunków liczy 150 stron. Składa się z pięciu rozdziałów, w których autor przedstawia kolejno rys historyczny dotyczący rozwoju ubezpieczeń uwzględniających informacje o procesach taryfikacyjnych. Przede wszystkim, opisana jest historia ryzyka ubezpieczeń komunikacyjnych. Zostały również omówione znane

z literatury modele mieszane. Przedstawiono również pojęcia związane z transferami ryzyk, które wpływają na zmianę finansów przedsiębiorstw ubezpieczeniowych, a także zagadnień związanych z hazardem moralnym charakterystycznym dla rynku ubezpieczeniowego.

Przegląd prac na poruszany temat (aktuarialny) został zaprezentowany w rozdziale drugim. Zdefiniowano pojęcie rynku technicznego oraz podstawowe miary opisujące zyskowność. Opisane miary oraz rozkłady prawdopodobieństwa zostały wykorzystane w dalszej części pracy. Zaprezentowany został przykład empiryczny, którego celem była prezentacja pomysłu autora – wprowadzenie efektów losowych, zarówno działań pochodzących od agenta ubezpieczeniowego jak i menedżera.

W rozdziale trzecim przedstawiono charakterystykę danych empirycznych wykorzystywanych w badaniu. Duża próba – 99,7 tysięcy rocznych polis komunikacyjnych – robi wrażenie! Informacje dotyczą, m.in. pojazdu, właściciela, losowych szkód oraz liczby i wysokości wypłat. Omówiona została również metoda badawcza stosowana w kolejnym rozdziale, której celem jest ilustracja skutków finansowych proponowanych przez Doktoranta.

Rozdział czwarty poświęcony jest modelom liniowym i mieszanym wykorzystywanym do analiz i porównaniu ryzyka wynikającego z zastosowania modeli uzupełniających dotyczących taryfikacji ubezpieczeń majątkowych, zaproponowanych przez Autora. Sprawdzono także, jak zmiana kosztów stałych wpływa na konkurencyjność ofert i wyniki finansowe przedsiębiorstwa ubezpieczeniowego. Analiza została przeprowadzona na 12 zestawach danych, a wyniki porównano przy pomocy współczynnika determinacji.

Rozdział piąty to podsumowanie. Weryfikacji poddano sformułowane we wstępie hipotezy badawcze. Sformułowano również wnioski, niestety nie zawsze w sposób trafiony oraz zaproponowano dalsze kierunki badań.

Uwagi zacznę od generalnego stwierdzenia, że niewątpliwą zaletą recenzowanej pracy jest to, że napisana jest w sposób zrozumiały (może z wyjątkiem rozważań matematycznych, o czym w dalszej części recenzji – delikatnie rzecz ujmując - są one nieprawidłowe). Czasami rozprawa wygląda jak zwykłe sprawozdanie branżowe, nie zaś rozprawa doktorska! Można także znaleźć sporo niedociągnięć – również w tekście.

W rozdziale pierwszym, będącym głównie rysem historycznym, jest akapit (str.17¹), w którym nie wiadomo czy rozważania dotyczą pracy Doktoranta, czy też jakiejś innej. Pisanie o rezultatach końcowych w rozdziale pierwszym jest raczej dyskusyjne. We wstępie, owszem, można się pokusić o sformułowanie spodziewanych wniosków wynikających z proponowanych hipotez, ale nie podczas omawiania „historii” i prezentacji stosowanych pojęć. W rozdziale 1.3 Autor pisze o metodach ekonometrycznych stosowanych w dalszej

części pracy. Szkoda, że wymienia je tylko z nazwy, gdy tymczasem rozkład Poissona (znany ze szkoły średniej) omawia dokładnie! Niektóre pojęcia wymagają rzetelnego zdefiniowania (albo przynajmniej jasno wyartykułowanego określenia). Przykładem może być wielokrotnie wymieniana selekcja negatywna niestety ukryta w „późniejszych” przypisach (dopiero przypis 10). Biorąc pod uwagę rodzaj rozważań związanych z ubezpieczeniami problem hazardu moralnego wart byłby szerszego omówienia w kontekście wprowadzenia do modelu zagadnień związanych z relacjami pryncypał - agent (pryncypał to u Doktoranta menedżer).

Modele ekonometryczne pojawiają się wprawdzie w rozdziale drugim. Jednakże sposób ich prezentacji pozostawia wiele do życzenia. Parametry nie są opisane, np. nie jest jasne jaką rolę i przede wszystkim co oznaczają niektóre parametry, np. parametr θ we wzorze 2.16? Na pierwszy rzut oka wygląda jak „zbiór pusty”. Dalej czytamy, że jest to raczej parametr z przedziału $[0,1]$, „reprezentujący udział zer w próbie”. O co dokładnie chodzi?

Na stronie 44 symbole we wzorach i tekście nie są adekwatne. Przykładowo: $fI(0)$ to nie jest to samo co $f_1(0)$, podobnie ki w odniesieniu do wzoru (2.7). Takich przykładów jest znacznie więcej. Wygląda na to, że prawidłowe posługiwanie się językiem matematyki nie jest mocną stroną Doktoranta. Na plus można jednak przypisać stwierdzenie, że do estymacji parametrów Doktorant posługuje się metodą największej wiarygodności. Czytając prace przyszłych naukowców, niejednokrotnie bowiem spotykałam się z sytuacją, gdy do estymacji parametrów w modelach liniowych (niezależnie od własności modelu) używano metody najmniejszych kwadratów. Najczęściej tego robić nie można! W tym przypadku Doktorant miał dobre wyczucie. Pochwalić również należy, co już było wspomniane, to że badania prowadzone są na dużej próbie z danymi oraz rzetelny opis pozyskiwania danych. Zaprezentowane zostały nawet histogramy dotyczące występowania niektórych zmiennych takich jak, np. wiek pojazdu, czy też pojemność silnika.

Inne uwagi.

Str 65. Czy określenie „model przeuczony” to jest jakaś nazwa „własna” Autora, czy może jakiś „aktualny slang”; jeśli tak to co owo określenie oznacza?

Str. 66. j.w. „efekt losowy czasu”? Jaki to ma związek z wymienioną, w kolejnym akapicie, „marką modelu pojazdu”?

Str. 71. Czy parametr β „ukryty” we wzorze na wariancję (3.10) to estymowany parametr? Czego dotyczą i jaką mają postać wariancje σ_y^2 , σ_e^2 ze wzoru (3.9)? Dlaczego w dalszej części pracy wariancja oznaczana jest przy pomocy symbolu s ?

Str. 72 – wzory (3.11) i (3.12), co oznacza symbol d – zob. indeksy?

Rozdział 4 –wzory- są jakby z innej pracy. W rozdziale 3 Autor używa jako zmiennych - liter; w tym rozdziale zmiennymi są skróty nazw. W ten sposób komunikatywność pracy znacząco spada.

Zmienne w pierwszej kolumnie tabeli 4.1 nie są opisane w tekście. Z tego też powodu tabela staje się nieczytelna. Nie jest jasny związek z wynikami zaprezentowanymi w tabeli a cytowanym w nagłówku wzorem (2.16). Podobnie rzecz się ma z innymi tabelami.

Statystyka Z nie jest opisana. Podobnie rzecz się ma ze statystyką STAT. Jaką rolę pełnią w nich parametry „beta”? Jak odczytywać ich wartości? Podobnie rzecz się ma z błędami np. MSE. itd. Warto również napisać z jakiego oprogramowania korzysta Doktorant.

W podsumowaniu Doktorant formułuje wniosek, że efekt poprawy modeli taryfikacyjnych można uzyskać poprzez zastosowanie nowych źródeł danych (co nie jest raczej odkrywcze, ale niewątpliwie jest prawdą). Podobnie można odczytywać sformułowanie „zwiększenie mocy predykcyjnej” (czy to nowomowa, czy slang?) modeli poprzez „poprawę predykcji oszacowania estymatorów”. Zwracam uwagę, że estymator to po polsku „oszacowanie”.

Praca ma wiele niedociągnięć. Jeżeli chodzi o „komunikatywność” rozdział czwarty, to kompletna porażka, chociaż widać, że Doktorant włożył weń wiele pracy. Jednakże uwagi, chociaż liczne, nie wpływają na końcową ocenę. Po pierwsze temat jest wart rozważenia . Po drugie Doktorant może „pochwalić” się własnym pomysłem związanym z modyfikacją modeli taryfikacyjnych. Po trzecie, z pominięciem rozważań „matematycznych” i wspomnianego rozdziału 4 praca napisana jest raczej w sposób jasny.

Reasumując, stwierdzam, że recenzowana przeze mnie rozprawa spełnia wymagania stawiane rozprawom doktorskim i wnoszę o dopuszczenie do publicznej obrony.

Wc Dmela