

Ocena rozprawy doktorskiej

**mgr Małgorzaty Kciuk pt. „PODEJŚCIE MODELOWE W STATYSTYCE MAŁYCH
OBSZARÓW I JEGO ZASTOSOWANIA W BADANIACH EKONOMICZNYCH”**

napisanej pod kierunkiem naukowym

**dr hab. , profesora Uniwersytetu Ekonomicznego w Katowicach,
Tomasza Żądło.**

1. Ocena formalna pracy

Przedstawiona do recenzji praca doktorska stanowi zwartą dysertację poświęconą, zgodnie z tytułem, podejściu modelowemu w statystyce małych obszarów i jego zastosowaniu w badaniach ekonomicznych. Sam tytuł pracy został określony właściwie i adekwatnie do zawartości pracy. Dysertacja ta składa się ze *wstępu*, *5 rozdziałów właściwych* (z czego cztery pierwsze mają charakter teoretyczny, a ostatni jest rozdziałem realizującym badanie symulacyjne) oraz *zakończenia*. Autorka we *wstępie* właściwie rozgranicza cele teoretyczno-poznawcze pracy (moim zdaniem wystarczyłoby nazwać je po prostu teoretycznymi) oraz cele praktyczne. Praca nie zawiera dodatkowych *spisów rysunków i tablic* (co z resztą nie zawsze jest praktyką), zawiera natomiast wnioski szczegółowe zawarte w rozdziale realizującym badanie symulacyjne oraz wnioski ogólne omówione w *zakończeniu*, które jednocześnie słusznie uzasadniają osiągnięcie celów pracy. Całość, wraz z wykazem literatury, liczy sobie aż 216 stron. Praca napisana jest bardzo starannie pod względem językowym i edytorskim (mimo pewnych uchybień opisanych w *załączniku* do recenzji), tworząc przy tym logiczną i spójną całość. Od strony formalnej nie zgłaszam istotnych uwag krytycznych.

2. Wybór kierunku badawczego i jego aktualność

Podjęty przez doktorantkę kierunek badawczy jest jednocześnie ciekawy, aktualny i ma duże znaczenie praktyczne. Jako recenzent pozwolę sobie na krótkie uzasadnienie tej opinii. Po pierwsze, o aktualności i zainteresowaniu podjętą tematyką świadczy zarówno relatywnie duża liczba ośrodków polskich i zagranicznych, których naukowcy stale rozwijają metodologię statystyki małych obszarów (w Polsce jest to m.in. ośrodek Warszawski, Poznański, Łódzki czy Katowicki), liczne,

współczesne publikacje z tego zakresu w prestiżowych czasopismach naukowych (np. *Statistical Papers*, *Journal of Official Statistics*, *Biometrika*, *Journal of the Royal Statistical Society (A)*, *Quality and Quantity*, czy *Journal of the American Statistical Association*) oraz monografie (np. Molina i Rao (2015) czy Pratesi (2016), Żądło (2015), Szymkowiak (2020)), wreszcie liczne seminaria, workshopy oraz konferencje branżowe (np. cykliczna, międzynarodowa konferencja *Small Area Estimation*). Po drugie, rozwijany w pracy kierunek badawczy ma niebagatelne znaczenie dla praktyki, a jego aplikacje omawianej tematyki sięgają takich obszarów jak rynek pracy, jakość życia i ubóstwo, polityka regionalna, ekonomika rolnictwa, czy ekologia (o czym zresztą nadmieniam wnikliwie autorka w podrozdziale 1.4). A zatem wybór tematyki pracy, jakiego dokonała doktorantka, uważam za w pełni zasadny i dobrze osadzony w stale rozwijającej się odpowiadającej jej metodologii badań statystycznych.

3. Ocena merytoryczna pracy

Zanim dokonam oceny merytorycznej sfery pracy pozwolę sobie na krótką charakterystykę jej pięciu rozdziałów właściwych z uwypukleniem ich głównych walorów. Pewne drobne uwagi, które wymagałyby ewentualnego komentarza lub korekty, zawarłem w załączniku dołączonego do niniejszej recenzji.

Rozdział 1

Rozdział ten w przystępny i jednocześnie precyzyjny sposób omawia teoretyczne podstawy statystyki małych obszarów (podejścia i trendy, pojęcia i definicje, twierdzenia). Bardzo przyjemnie czyta się podrozdział 1.3, omawiający genezę i rozwój statystyki małych obszarów zarówno w Polsce jak i na świecie, oraz podrozdział 1.4, który szeroko traktuje o potencjalnych zastosowaniach statystyki małych obszarów. Jednocześnie rozdział ten proponuje czytelnikowi autorskie propozycje szczególnych przypadków znanych modeli, które to przypadki uwzględniają skorelowanie wektorów efektów losowych (jest to główna idea pracy, w ślad za którą podążają również pozostałe rozdziały).

Rozdział 2

Rozdział ten omawia problematykę badań powtarzalnych w czasie, przy choć wspomina o badaniach prowadzonych w jednym okresie, to środek ciężkości kładzie na badaniach wielookresowych, w tym panelowych. Rozdział ten ma charakter przeglądowny, jednak świadczy o dużej świadomości autorskiej w zakresie wad i zalet, oraz ograniczeń i korzyści płynących z tego typu badań.



Rozdział 3

W rozdziale tym autorka podejmuje problem predykcji z wykorzystaniem predyktorów typu BLU (najlepsze liniowe nieobciążone estymatory) i EBLU (empiryczne najlepsze nieobciążone estymatory). Podobnie jak w każdym rozdziale (poza przeglądowym rozdziałem 2), autorka przedstawia tu również swoją propozycję, którą w tym przypadku jest wykorzystanie predyktora typu EBLU przy założeniu liniowego modelu mieszanego ze skorelowanymi efektami losowymi.

Rozdział 4

W rozdziale tym rozważana jest klasa empirycznych, najlepszych predyktorów oraz predyktorów plug-in, które dodatkowo bazują na oszacowaniach dla jednostek spoza próbki z wykorzystaniem estymatorów zgodnych (co wpisuje te predyktory do klasy predyktorów wielostopniowych). Również i w tym rozdziale nie zabrakło ciekawych, autorskich pomysłów – autorka proponuje tu wykorzystanie predyktorów należących do tych klas w oparciu o modele ze skorelowanymi wektorami efektów losowych. Również konsekwentnie do wcześniejszych rozdziałów, rozdział 4 jest zakończony podsumowaniem, które nie tylko stanowi krótkie streszczenie jego zawartości, ale również (co ułatwia pracę recenzentowi i odbiór czytelnikowi), omawia wkład własny doktorantki.

Rozdział 5

Rozdział stanowi szerokie badanie symulacyjne, które za punkt wyjścia (badaną cechę) obiera rzeczywiste dochody gmin pozyskane z BDL za lata 2018-2020. Rozdział co prawda pomija kanon w postaci ustalenia celów badania, ale dalej już jest tylko bardzo dobrze: mamy tu opis wykorzystanego zbioru danych, założenia dla eksperymentów, opis procedury symulacyjnej, prezentację wyników i szczegółowe wnioski dla każdego z trzech rozważanych wariantów korelacyjnych. Rozdział podnosi problem predykcji wartości globalnych oraz median w domenach przy założeniu liniowego modelu mieszanego z uwzględnieniem korelacji pomiędzy efektami losowymi. Autorka bada własności trzech zaproponowanych predyktorów (z klas BP, EBP, oraz plug-in) i porównuje je z wybranymi predyktorami zakładającymi brak zależności między efektami losowymi. Wnioski z rozdziału, potwierdzające dobre własności autorskich propozycji predyktorów oraz uzyskany dzięki nim zysk na dokładności i precyzji predykcji, mogą być ważne dla środowiska naukowego zajmującego się podjętą w pracy tematyką. Jednocześnie, dysertacja jako całość, a w szczególności wnioski z badania symulacyjnego, stanowią w moim mniemaniu doskonały punkt wyjścia dla przyszłych badań naukowych doktorantki.

Moja ocena merytorycznej sfery pracy jest jednoznacznie bardzo wysoka. Na szczególną uwagę zasługują bowiem jej następujące walory: bardzo dobry, wnikliwy przegląd literatury przedmiotu i stanu wiedzy na podjęty w pracy temat, mnoga liczba autorskich propozycji (np. modeli ze skorelowanymi wektorami efektów losowych, zastosowań tych modeli w statystyce małych obszarów, modyfikacji metod oceny MSE predyktorów) oraz rzetelne badanie symulacyjne, które dostarcza ciekawych wniosków i stanowi jednocześnie zachętę do dalszych badań nad autorskimi propozycjami modeli i predyktorów.

Na koniec tej części chciałbym raz jeszcze podkreślić, iż wszystkie omówione rozdziały dysertacji doktorskiej mgr Małgorzaty Krzciuk potwierdzają jej głęboką wiedzę z zakresu statystyki małych obszarów, umiejętności analityczne (o czym świadczy prezentowana kolejność terminów, twierdzeń, własnych propozycji i wywodów myślowych) oraz cenną umiejętność wiarygodnego prowadzenia badań naukowych (o czym świadczy szeroko przeprowadzone badanie symulacyjne i szereg wniosków jakie doktorantka wyciągnęła na podstawie przeprowadzonych analiz). Jako recenzent zwróciłem również uwagę, na bardzo rozległą literaturę z jakiej korzystała doktorantka (364 pozycje, głównie anglojęzycznych, w tym 3 pozycje wyłącznie autorskie doktorantki i 4 współautorskie) oraz niewątpliwą wartość dodaną tej pracy w postaci autorskich skryptów napisanych w środowisku R. Należy oczekiwać, iż w przyszłości doktorantka opublikuje swoje wyniki w dobrych czasopismach branżowych (obecne, zawarte w pracy samodzielne publikacje doktorantki dotyczą mniej znaczących periodyków lub są to jedynie wystąpienia konferencyjne).

4. Wnioski

Podsumowując recenzowaną pracę bez wątplenia należy określić jako oryginalną i twórczą. Praca wykracza poza ramy typowych opracowań z tematyki statystyki małych obszarów i rzuca nowe światło na te wciąż nie do końca zbadane problemy z zakresu podejścia modelowego. Poziom wiedzy teoretycznej i umiejętności analityczne mgr Małgorzaty Krzciuk w zakresie podjętej tematyki należy ocenić wysoko i szkoda, że nie zostały one zaakcentowane mocniej w spisie literatury poprzez bardziej prestiżowe publikacje. Jestem całkowicie przekonany, iż doktorantka jest w pełni gotowa do prowadzenia dalszej, samodzielnej pracy naukowej, na której wyniki z niecierpliwością czekam.

W zgodzie z powyższym stawiam wniosek o dopuszczenie pracy mgr Małgorzaty Krzciuk pt. „*Podejście modelowe w statystyce małych obszarów i jego zastosowania w badaniach ekonomicznych*” do publicznej obrony zaznaczając, iż spełnia one wszystkie ustawowe wymogi stawiane rozprawom doktorskim. Uważam również, iż recenzowana przeze mnie dysertacja doktorska zasługuje na ewentualną publikację i – jeśli taka jest praktyka reprezentowanej przez doktorantkę uczelni – wyróżnienie stosowną nagrodą, np. Nagrodą Rektora.



Załącznik

(uwagi szczegółowe dotyczące pracy i poszczególnych rozdziałów)

Praca zawiera bardzo niewielką liczbę usterek edytorskich, stylistycznych lub gramatycznych. Dla porządku wspomnę o niektórych: 1) We wstępie zdanie „Przedmiotem badań będzie (...)” nie doprecyzowuje lokalizacji (gdzie? w dysertacji?); 2) Autorka często nie stawia przecinka tam, gdy jest on wymagany (np. nagminnie nie ma go przed „gdy”, np. str. 11 pod wzorem (1.8) albo str. 92 pod wzorem (3.6)); 3) Zwrot „w ujęciu podejścia randomizacyjnego” (str. 7) można by z powodzeniem zastąpić zwyczajnie „w podejściu randomizacyjnym”. Jednak całościowo, o czym już wspominałem, praca jest przygotowana niezwykle starannie. Poniżej umieszczam drobne uwagi i komentarze, dotyczące poszczególnych rozdziałów pracy, jakie autorka powinna wziąć pod uwagę przed jej ewentualną publikacją:

Rozdział 1

- Brak precyzji w definiowaniu 4 typów małych obszarów – np. czytelnik nie wie, jak zaklasyfikować przypadek zawierający 1% badanej zbiorowości (typ „drugorzędny” czy „mini”?). Domyślnie zachowane jest pewnie prawostronne domykanie przedziałów.
- Zmienna losowa raz jest oznaczana dużą literą a raz małą, na tej samej stronie 7 (i dalej). Na tej samej stronie raz jest też wytłuszczona (\mathbf{y}), a raz nie (y).
- Na str. 11 pojawia się termin „p-nieobciążoności” estymatora, który to nie jest sprecyzowany wcześniej (na str. 9 jest mowa jedynie o „estymatorze nieobciążonym”, a wcześniej o „p-obciążeniu”). Oczekiwać by jednak można podobnej precyzji ja np. w przypadku definiowania ξ – nieobciążoności na str.15.
- We wzorze (1.31) pojawiają się symbole $\hat{\theta}_1$ i $\hat{\theta}_2$ estymatorów składowych przy definiowaniu estymatora złożonego. Natomiast przy określaniu jego MSE (formuła (1.32)) figurują już $\hat{\theta}^A$ i $\hat{\theta}^B$. Co więcej, zgodnie z aproksymacją (1.33), gdy $MSE_p(\hat{\theta}^A) \approx 0$, wówczas $q_d^* \approx 1$ i wówczas, wobec (1.32), mamy $MSE_p(\hat{\theta}^{COMP}) \approx MSE_p(\hat{\theta}^A) \approx 0$ (a więc MSE estymatora złożonego właściwie nie zależy od MSE estymatora drugiego rodzaju ani od jego wartości oczekiwanej lub jakiegokolwiek innej charakterystyki. Czy to poprawny wniosek?
- Predyktor w podejściu modelowym jest oznaczany jako $\hat{\theta}(Y^*)$ na str. 15 ale jako $\hat{\theta}(M^*)$ na str. 7.
- Czemu błąd predykcji określony wzorem (1.36), konsekwentnie do oznaczeń $B_\xi()$ czy $D_\xi^2()$, nie ma oznaczenia U_ξ ?

- Czy wzór (1.41) jest wyrażony w %?
- Implikacja (1.43) jest pod względem matematycznym zapisana nieprecyzyjnie, choć dobrze opisana wcześniej słownie. Zamiast „=min” (min po czym? z czego?) już lepszy byłby zapis „->” jaki został zastosowany np. na str. 21 w zapisie (1.60).
- W podrozdziale 1.2.3 trudno jest zorientować się, co jest autorską propozycją. Np. model określony przez (1.133) został zaanonsowany odnośnikiem do pracy autorki (Krzciuk, 2020, s.20) ale czytelnik już nie wie, czy dalsze propozycje (1.136) oraz (1.139) są również autorskie (domyślamy się jedynie, że TAK).
- Wzór na dewiancję nie dostał swojego numeru na str. 45;
- Na str. 47 testy normalności Shapiro-Wilka i Lilieforsa otrzymały swoje cytowania (prace odpowiednio z roku 1965 i 1967), ale test F i test Walda już nie.

Rozdział 2 – brak uwag

Rozdział 3

- W podrozdziale 3.1.1. dotyczącym predyktora typu BLU Royalla pojawia się na str. 93 założenie o symetrii rozkładów wektorów efektów i składników losowych. Czytelnik zna i rozumie zapewne pojęcie symetrii/asymetrii rozkładu jednowymiarowej zmiennej losowej, natomiast może mieć problem z właściwym odbiorem założenia dotyczącego wielu wymiarów (całego wektora zmiennych losowych). Myślę, że ten zapis wymaga doprecyzowania (czy chodzi np. o zerowe współczynniki asymetrii wyznaczone dla rozkładów brzegowych wektora zmiennych losowych?, czy chodzi o symetrię/asymetrię w sensie statystycznym czy też matematycznym?)
- Symbol $o(D^{-1})$, nie wyjaśniony na str. 96, 101 czy 102, może być dla mniej zmatematyzowanego czytelnika niezrozumiały.
- Może warto byłby również wyjaśnić czytelnikowi, czemu estymator (3.34) błędu średniokwadratowego EBLUP nosi miano „naiwnego” (str. 100).
- W koncepcji estymatorów odpornych na wartości odstające (Sekcja 3.5.2 czy 3.5.3) zabrakło mi zdefiniowania, co rozumiemy tutaj pod pojęciem „wartości odstające”?
- Sformułowanie „zapożyczenie mocy między domenami” (str.117) wymaga doprecyzowania.
- Jak pisałem wcześniej, często w pracy mam problem ze zdiagnozowaniem, co jest rzeczywistym wkładem autorki. Rozdział 3, podobnie jak inne, kończy się wylistowaniem autorskich propozycji jakie są w nim zawarte. Czytamy tu m.in. o autorskich modyfikacjach znanych metod oceny błędu średniokwadratowego predykcji uwzględniających korelacje pomiędzy wektorami efektów losowych. Odnoszę jednak wrażenie, że modyfikacje te zostały

zbyt słabo uwypuklone w tekście i trudno zorientować się, w którym dokładnie miejscu zaczyna się propozycja autorki (np. czy chodzi o propozycję (3.33)? jeśli tak, to jest to jedna modyfikacja a nie „modyfikacje”). Poza tym opisana na str. 99 modyfikacja dotyczy samego predyktora, a nie jak podkreślono w podsumowaniu tego podrozdziału – oceny błędu średniokwadratowego (nie ma nigdzie dyskusji nad $MSE(\hat{\theta}_{EBLUP}^*)$).

Rozdział 4

- Czy algorytm Monte Carlo opisany na str. 128 to propozycja autorska?
- W podrozdziale 4.3 znajduje się ciekawa propozycja autorska dotycząca zastosowania predyktora plug-in przy założeniu liniowego modelu ze skorelowanymi wektorami efektów losowych (wzory 4.5 i 4.6). Podrozdział ten jednak omawia (i tak jest w pracy z reguły) już przeprowadzone badanie empiryczne dotyczące tej propozycji (Krzciuk, 2019), aniżeli prezentuje nowe wyniki i kolejne wnioski. Cytowana publikacja jest zresztą „jedynie” referatem wygłoszonym na konferencji we Florencji, czyli nie stanowi recenzowanej publikacji w regularnie wydawanym czasopiśmie naukowym.
- Jako autorskie propozycje zostały tu wskazane oceny MSE dla rozważanych klas predyktorów, ale ponownie jako recenzent mam problem z oddzieleniem wyników już znanych od tych proponowanych. Jest co prawda zdawkowa wzmianka na ten temat na str. 135 i powołanie się na pracę Krzciuk i inni (2017), ale przywołano tu jedynie zastosowanie metody bootstrap do oceny MSE zaproponowanego predyktora, co samo w sobie nie jest nową propozycją wyznaczania MSE.

Rozdział 5

- Brakuje mi na wstępie do tego rozdziału wskazania celów badania, tego co autorka chce w rozdziale tym wykazać lub zbadać;
- Choć rozdział ten zapowiada badanie stricte symulacyjne, to jednak punktem wyjścia (badaną zmienną) są rzeczywiste dochody gmin pozyskane z BDL za lata 2018-2020. To tylko uwaga, która wskazuje na powiązanie przeprowadzonych badań z danymi empirycznymi.
- W podrozdziale 5.2.1 pojawiają się oznaczenia predyktorów, jakie mają być wzięte pod uwagę w badaniu symulacyjnym. Ale np. zaproponowany najlepszy predyktor uwzględniający korelację efektów losowych oraz proponowany empiryczny najlepszy predyktor przy założeniu liniowego modelu mieszanego ze skorelowanymi efektami

losowanymi otrzymuję to samo oznaczenie: EBP_ρ (szczęśliwie na wykresach jest rozróżnienie: BP_rho i EBP_rho). W opisie do wyników pominięto (choć oczywiście jest to bardzo domyślne), że estymator Horvitz-Thompsona będzie oznaczamy „H-T”, syntetyczny estymator ilorazowy oparty na estymatorze H-T będzie oznaczany przez „SYNT”, a estymator kalibrowany przez „CALIB”.

- W badaniu przyjęto dość specyficzne wartości parametru ρ , tj. -0.83 (wartość empiryczna), -0.95 i -0.65 (wartości ustalone w eksperymencie symulacyjnym). Trochę zbyt zdawkowo uzasadniono w pracy wybór dwóch dodatkowych wartości tego parametru – mamy jedynie (str. 180) wzmiankę o tym, iż wartości te wynikały z badania wstępnego. Ale co ono pokazało? Na jakiej próbie je przeprowadzono? W jaki sposób ustalono te wartości? Wydaje mi się, że interesującym byłoby rozważanie siatki wartości tego parametru korelacyjnego z przedziału powiedzmy $[-1,0]$ z krokiem choćby rzędu 0.01, a następnie sporządzenie wykresu, który pokazywałby zysk na dokładności i precyzji porównywanych predyktorów jako funkcję parametru ρ .
- Uważam, że wnioski z badania symulacyjnego (tak naprawdę empiryczno-symulacyjnego) są bardzo obiecujące i zachęcam autorkę do dalszych badań nad swoimi propozycjami predyktorów. Zauważmy jednak, iż punkt wyjścia do wszystkich scenariuszy i repetycji symulacyjnych jest statystyczny, tzn. oparty o ten sam zestaw danych rzeczywistych. Zestaw ten służy co prawda do generowania 3000 danych populacyjnych (algorytm 3), ale tak naprawdę w naturalny sposób ogranicza ten generator. Wydaje się, iż aby zyskać na możliwości generalizowania wniosków, należałoby rozważyć pewne spectrum danych wejściowych (np. zestawy danych z różnych krajów, zestawy danych generowanych wg ustalonego ekspercko rozkładu, itd.) i całość powtórzyć w podwójnej wówczas pętli (mam świadomość czasochłonności takiego przedsięwzięcia). To oczywiście jedynie uwaga czy sugestia, nie krytykująca wcale doskonale przeprowadzonego w pracy badania.

Michał Miutek
Jurysek, dnia 19.08.2022