

Łódź, 18.09.2023 r.

Dr hab. Aleksandra Baszczyńska
Uniwersytet Łódzki
Wydział Ekonomiczno-Socjologiczny
Katedra Metod Statystycznych



**Recenzja rozprawy doktorskiej mgr. Pawła Zdzisława Lańducha
pt. „Wykorzystanie technik imputacyjnych w szacowaniu informacji
wynikowych oraz w analizie struktury danych w statystyce przedsiębiorstw”**

przygotowanej w Szkole Głównej Handlowej w Warszawie
pod kierunkiem naukowym dra hab. Andrzeja Młodaka

1. Podstawa recenzji:

- Dokument z dnia 29.06.2023 r. informujący, że Rada Naukowa dyscypliny ekonomia i finanse uchwałą nr 721 z dnia 14.06.2023 r. powołała mnie na recenzenta rozprawy doktorskiej mgr. Pawła Zdzisława Lańducha nt.: „Wykorzystanie technik imputacyjnych w szacowaniu informacji wynikowych oraz w analizie struktury danych w statystyce przedsiębiorstw” w dziedzinie nauk ekonomicznych w dyscyplinie ekonomia,
- Ustawa z dnia 14.03.2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. z 2017 r., poz. 1789 z późn. zm.).

2. Ocena wyboru problemu badawczego

Techniki imputacyjne, rozumiane jako procedury uzupełniania braków w próbie przy jednoczesnym spełnieniu dobrych własności estymatorów standardowych, stanowią istotne narzędzie statystyczne zapewniające zadowalającą efektywność procedur estymacyjnych. Ma to szczególne znaczenie w analizach danych związanych ze statystyką publiczną. Procedury te są szeroko prezentowane w literaturze przedmiotu (zarówno z teoretycznego punktu widzenia, jak i oprogramowania), jednak zauważalna jest stosunkowo niewielka

liczba publikacji dotyczących technik imputacyjnych w zastosowaniach związanych ze statystyką przedsiębiorstw. Tak specyficzne zagadnienie, jakie zaprezentowane jest w rozprawie doktorskiej mgr. Pawła Łańducha, czyli analiza możliwości zastosowania imputacji masowej w badaniach przedsiębiorstw jest ściśle związane z potrzebami, jakie stawiane są obecnie przed statystyką przedsiębiorstw. Braki danych w bazach tworzonych przez statystykę przedsiębiorstw zdarzają się często, a techniki imputacji masowej zapewniają nie tylko dobre własności estymacyjne stosowanych procedur, ale również mogą skutkować zmniejszeniem kosztów badania statystycznego. Problematyka badawcza jest zatem niezwykle cenna z aplikacyjnego punktu widzenia. Dlatego też bardzo wysoko oceniam wybór problemu badawczego, zaś wyniki badań przedstawione w rozprawie oceniam jako wnoszące nowe elementy poznawcze w zakresie zastosowania imputacji masowej w statystyce przedsiębiorstw.

3. Ocena celów badawczych i hipotez rozprawy doktorskiej

Głównym celem rozprawy jest: „ocena efektywności zastosowania metod imputacji masowej, bazującej na integracji próby losowej i nielosowej w badaniu statystycznym przedsiębiorstw”. Cel pracy został sformułowany poprawnie. Moje wątpliwości budzi jednak sformułowanie celów szczegółowych analizy. Zostały one określone jedynie jako wyznaczenie wartości globalnych analizowanych cech w oparciu o estymator Horvitz-Thompsona przy różnych założeniach. W pracy przedstawione są wyniki analiz wykorzystujących te wartości globalne, co sugeruje, że celami szczegółowymi były analizy dokonane w oparciu o te procedury wyznaczenia wartości globalnych (a nie tylko samo wyznaczenie wartości określonych miar).

Stosownie do postawionych celów sformułowano główną hipotezę badawczą: „Imputacja masowa jako metoda integracji danych, w ujęciu łączenia próby losowej i nielosowej, może być użytecznym narzędziem – wspomagającym lub podstawowym – służącym uzyskaniu podstawowych, szybkich szacunków porównywalnej jakości w porównaniu z badaniami przedsiębiorstw prowadzonymi w oparciu o losową i warstwową próbę, kiedy zakładamy, że wartości analizowanych cech są obserwowane tylko w próbie nielosowej i stamtąd mogą być zaczerpnięte.” Hipoteza główna jest zgodna z celem głównym pracy. Moich zastrzeżeń nie budzi również hipoteza dodatkowa, prezentująca konkretne rozwiązanie w przypadku stosowania imputacji masowej w statystyce przedsiębiorstw zapewniające możliwość stosowania mniejszych prób w badaniach statystyki publicznej.

4. Ocena metod badawczych

Wybór metod badawczych dokonany jest w odpowiedni sposób, w pełnej zgodności z celami pracy (głównym celem pracy i dodatkowymi celami analizy). W pierwszej części pracy przedstawiono wyniki wnikliwych studiów literaturowych dotyczących ogólnej charakterystyki badań statystycznych przedsiębiorstw, integracji danych na poziomie mikro oraz imputacji masowej. Przegląd literatury został dokonany w oparciu o źródła polskojęzyczne i anglojęzyczne. Specyfika tematyki podjętej w pracy wymagała korzystania zarówno z dokumentów Głównego Urzędu Statystycznego i Eurostatu, jak również ze specjalistycznych prac naukowych publikowanych w uznanych wydawnictwach polskich i zagranicznych.

Część empiryczna pracy wykorzystuje metodę badawczą opartą na symulacji komputerowej. Na podkreślenie zasługuje fakt prawidłowego zaprojektowania badania statystycznego opartego na symulacji komputerowej, co nie zawsze jest zadanie łatwym, szczególnie w naukach ekonomicznych. W pracy zaprezentowano wyniki badania symulacyjnego przeprowadzonego dla dwóch zmiennych: przychodów netto ze sprzedaży produktów oraz wynagrodzenia brutto z wykorzystaniem danych pochodzących z zasobów statystyki publicznej. Rozważano w badaniu integrację danych dokonaną czterema metodami, w tym dwie metody są metodami parametrycznymi a dwie semiparametrycznymi. Szeroki zakres wariantów badania symulacyjnego pozwala Autorowi na dokonanie szczegółowych analiz i umożliwia wybór metody najlepszej z określonego punktu widzenia. Wybór procedur oceniam jako właściwy. Zastrzeżenie budzi jedynie brak podania oprogramowania, które zostało wykorzystane w poszczególnych etapach realizacji celów pracy.

5. Ocena układu pracy i zawartości poszczególnych części

Praca liczy 124 strony. Zawiera stronę tytułową, spis treści, streszczenie w języku polskim, streszczenie w języku angielskim, wprowadzenie, pięć rozdziałów, zakończenie, spis literatury, wykaz tablic, wykaz rysunków i wykresów oraz aneks.

Układ pracy jest czytelny i nie budzi zastrzeżeń. Zawartości poszczególnych części pracy są zgodne z tytułami tych części i cechują się wysokim poziomem merytorycznym. Drobne uwagi związane z układem pracy i zawartością poszczególnych części są podane poniżej.

- Nie każdy rozdział ma pierwszy podrozdział określony jako wstęp.
- W spisie treści niestety nie został uwzględniony podrozdział 2.3.2., występujący w pracy na stronie 39, chociaż nie jest on tam wyodrębniony graficznie.

- W spisie pracy, jak również w samej pracy, występują dwa podrozdziały o tym samym numerze (4.3.2), brak jest natomiast podrozdziału o numerze 4.3.1.
- We wprowadzeniu zastrzeżenie moje budzi określenie oceny jakościowej estymacji za pomocą współczynnika zmienności estymatora. Ocena dokonana na podstawie miary statystycznej powinna być raczej określona jako ilościowa.
- W rozdziale pierwszym błędnie podano liczbę faz procesu statystycznego na poziomie 1.
- Błędny zapis źródeł dla rysunku 1 i rysunku 2 (brak jest określenia „opracowanie własne na podstawie....”).
- W graficznej prezentacji modelu odpowiedzi w badaniach przedsiębiorstw (rysunek 3) brak jest określeń w języku polskim.
- Błędnie podany rozkład przedsiębiorstw w Polsce w tablicy 2 (podane wartości nie sumują się do 100%).
- Brak podania współautorów pracy Cibelli z 2008 roku.
- Brak ujednolicenia oznaczenia braków danych w tablicach 3 i 4.
- Brak ujednolicenia w pracy określenia statystyka publiczna i statystyka oficjalna.
- Niezrozumiałe zdanie na stronie 55.
- Dyskusyjne jest sformułowanie dotyczące ciągłości zmiennej „Liczba pracujących”.
- Brak ujednolicenia określenia semiparametrycznych metod („pół-parametryczna metoda” nie wydaje się być określeniem stosowanym w polskiej literaturze przedmiotu).
- Brak jest komentarza dotyczącego informacji (przedziałów ufności) zawartych w aneksie.

6. Ocena rozprawy pod względem formalnym i redakcyjnym

Praca niestety zawiera dość dużo błędów redakcyjnych i formalnych, które są wymienione poniżej. Nie wpływają one zasadniczo na ogólną pozytywną ocenę pracy.

Numer strony	Uwagi
Dotyczy całej pracy	• Występują wiszące spójniki. • Brak ujednolicenia zapisu powoływania się na tablicę (wielka czy mała litera). • Brak ujednolicenia zapisu (brak spacji i spacja) w zapisie numeru przypisu dolnego. • Brak ujednolicenia zapisu nazwy zmiennych. W wielu miejscach pracy, gdy Autor powołuje się dwie cechy zawarte w formularzy badania DG-1 jedna zmienna jest zapisana z wielkiej litery (Przychody) a druga z małej (wynagrodzenia). • Tytuły rozdziałów i podrozdziałów kończą się kropką, co przez wielu językoznawców jest traktowane jako drugorzędny błąd interpunkcyjny. • Brak ujednolicenia przy powoływaniu się na literaturę (na przykład na stronie 21 lub 24 jest raz użyty spójnik „i” a raz „and”). • Brak jest numeracji wzorów. • Błędnie użyte znaki interpunkcyjne lub spacje (strony: 20; 38; 39; 41; 45; 46; 49; 52; 53; 56; 58; 63; 64; 72; 75). • Występowanie w pracy tablic znajdujących się na dwóch stronach bez powtórzenia na drugiej stronie numeru i tytułu tablicy. • Nieliczne przypadki braku kontroli wód i bękartów.

2	Brak ujednolicenia zapisu numeru podrozdziału (kropka)
4	Określenie „częściowo parametryczne” powinno być zapisane bez użycia łącznika
8	Brak jest informacji dotyczącej skrótownicy ISAD
9	Niepotrzebny podział wiersza w trzecim punkcie (lewa kolumna) w tablicy 1
10	„między” zamiast „miedzy”
10	Praca Kim i in. pochodzi z 2021 roku
11	Brak początku zdania w tekście w przypisie dolnym
19	GSBPM zamiast GSBM (dwa miejsca)
22	Brak jest w spisie literatury powołania ze strony 22 (EWG, 1993)
24	„wcześniej” zamiast „Wcześniej”
25	Ohlsson zamiast Olson
25	„nieopartą” zamiast „nie opartą”
27	Pietsch zamiast Pietch
31	Praca Lynna pochodzi z 2009 roku
32	(Eurostat, 2017b) zamiast (Eurostat, 2017)
34	Błędne powołanie na tablicę 1 zamiast na tablicę 3
35	Brak „X” w opisie oznaczenia $f_{z x}$
35	Brak w spisie literatury ESS ADMIN, 2020
36, 37, 38	Heerschap zamiast Heerchap
37	Cibella (2014) zamiast Cibella (2013)
40	Tuoto (2014) zamiast Tuoto (2013)
44	Brak ujednolicenia zapisu współautorstwa (i in. zamiast et al.)
49	$P(I_i I_j = 1 i, j \in U)$ zamiast $P(I_i I_j = 1 i, j \in U)$
50	Wystąpienie innej czcionki
50	Brak informacji czy praca Yanga i Kima to 2020a czy 2020b
50	Kim i in. (2021) zamiast Kim i in. (2020)
50, 52	Brak jest w spisie literatury pracy Kim (2022)
62	Błędne powołanie na podrozdział
63	pkt. 3.4 zamiast pkt. 4.3
67	0,1 do 0,9 zamiast 0.1 do 0.9
68, 69	$y_{h,i}^*$ zamiast y_i^*
68	(ang. splines) zamiast (splines)
74, 76, 77	Błędne powołanie na tablicę 2
80-84	Brak źródła wykresów; brak informacji co oznacza ciemnoniebieski kolor na wykresie
Literatura	Niezachowana kolejność alfabetyczna; brak ujednolicenia zapisu stron; brak daty dostępu przy źródłach internetowych; brak miasta wydania przy pozycjach książkowych; brak ujednolicenia zapisu tytułów artykułów; brak ujednolicenia kroju czcionki.
96	(2020). Combining non-probability.... zamiast (2021). Combining non-probability...

7. Ocena rozprawy doktorskiej

Rozprawa doktorska mgr. Pawła Lańducha charakteryzuje się dobrym poziomem merytorycznym. Na podstawie części teoretycznej pracy należy stwierdzić, że Autor prezentuje ogólną wysoką wiedzę teoretyczną w dyscyplinie ekonomia i finanse. Zagadnienia dotyczące statystyki publicznej, w tym statystyki przedsiębiorstw, jak również matematyczne i statystyczne zagadnienia związane z tematem rozprawy doktorskiej prezentowane są w pracy w sposób świadczący o bardzo dobrym przygotowaniu w prezentacji niezwykle skomplikowanych zagadnień. Część empiryczna pracy dowodzi wyjątkowo sprawnego poruszania się w zakresie projektowania, przeprowadzania i analizy wyników badania symulacyjnego. Na uwagę zasługuje również swoboda Autora

i jednocześnie klarowność w prezentacji szczegółowych rozwiązań wynikających z przeprowadzonych badań. Świadczy to jednoznacznie o tym, że rozprawa doktorska wykazuje umiejętność samodzielnego prowadzenia pracy naukowej przez Autora.

Tak szczegółowe podejście do tematu imputacji danych wraz z dogłębną analizą narzędzi statystycznych mogących mieć zastosowanie w statystyce przedsiębiorstw uzasadnia twierdzenie, że rozprawa doktorska wnosi oryginalne rozwiązanie problemu naukowego w zakresie zastosowania wyników własnych badań naukowych w sferze gospodarczej lub społecznej.

Stwierdzam, że przedłożona do recenzji rozprawa doktorska mgr. Pawła Zdzisława Lańducha pt. „Wykorzystanie technik imputacyjnych w szacowaniu informacji wynikowych oraz w analizie struktury danych w statystyce przedsiębiorstw” spełnia wymogi stawiane pracom doktorskim określone w Ustawie z dnia 14.03.2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. z 2017 r., poz. 1789 z późn. zm.). W związku z powyższym wnoszę do Rady Naukowej dyscypliny ekonomia i finanse Szkoły Głównej Handlowej w Warszawie o jej przyjęcie i dopuszczenie Doktoranta do publicznej obrony.

Aleksandra Baszczyńska.