

prof. dr hab. Jan Mielniczuk
Instytut Podstaw Informatyki PAN
i Wydział MiNI PW

30.01.2021

Recenzja osiągnięć naukowych dr Agnieszki Goroncy w związku z postępowaniem o nadanie stopnia doktora habilitowanego w dziedzinie nauk ścisłych i przyrodniczych, dyscyplina matematyka

Omówienie podstawowego osiągnięcia naukowego

Podstawowe osiągnięcie naukowe dr A. Goroncy zatytułowane *Oszacowania dla uporządkowanych funkcjonalów statystycznych pochodzących z nieparametrycznych rodzin rozkładów* składa się z 8 prac dotyczących zastosowań rachunku prawdopodobieństwa opublikowanych w latach 2014-2020, z czego cztery prace ([A2],[A3], [A7] i [A8]) są jedno-autorskie, reszta jest dwu-autorska. Prace zostały opublikowane w solidnych czasopismach statystycznych z listy JCR: *Statistics* (3 prace), *Metrika* (2 prace), *Communications in Statistics, Theory and Methods* (2 prace) i *Statistical Papers* (1 praca).

Prace poświęcone są tematyce optymalnych oszacowań funkcjonalów statystycznych dla danych uporządkowanych o dystrybuancie generującej pierwotne dane pochodzącej z nieparametrycznie określonych rodzin rozkładów. Wykorzystuje się w tym celu dwa podejścia: pierwsze oparte na technice rzutowania i posiłkujące się nierównością Marigutiego (Gajek i Rychlik (1996)) oraz drugie, wykorzystujące metodę minimalizacji normy (Goroncy i Rychlik (2011)) stosowaną wtedy, gdy pierwsze podejście daje trywialne górne oszacowanie zerowe.

Prace [A6] i [A7] dotyczą oszacowań wartości oczekiwanych standaryzowanych statystyk pozycyjnych i spacji w przypadku, gdy oryginalna próba była próbą niezależnych obserwacji z tego samego rozkładu zadanego dystrybuantą F .

Niech W będzie dystrybuantą z gęstością w o nośniku $[0, d)$, $\mathcal{C}_{\leq cW}$ oznacza stożek nierosnących, wklęsłych funkcji z $L^2([0, d), w)$ a $P_{\leq cW}$ będzie operatorem rzutu na ten stożek. W pracy [A7], w celu wykorzystania oszacowania górnego Gajka i Rychlika (1996) dla wartości oczekiwanej standaryzowanej L-statystyki $E(\sum_{i=1}^n c_i(X_{i:n} - \mu)/\sigma)$, (μ i σ^2 wartość oczekiwana i wariancja F) przez normę rzutu $\|P_{\leq cW}(\sum_{i=1}^n c_i(f_{i:n}W - 1))\|$ w przypadku, gdy F ma rosnącą W -uogólnioną intensywność awarii, bada się postać rzutów dla ogólnej

klasy funkcji spełniających warunek A (sekcja 2 pracy [A7]). Postać rzutu wykorzystuje się następnie do znalezienia optymalnych oszacowań dla wartości oczekiwanych standaryzowanych statystyk porządkowych i spacji z rozkładu o rosnącej gęstości (praca [A7]) i o rosnącej intensywności awarii (praca [A6]) oraz określenia rozkładów, dla których ograniczenia górne są osiągalne.

Następną pracą wykorzystującą technikę rzutowania jest praca [A3], gdzie wyznacza się oszacowania dla wartości oczekiwanych standaryzowanych n -tych wartości k -tego rekordu $R_n^{(k)}$ obliczanego na podstawie nieskończonego ciągu X_1, X_2, \dots , w sytuacji, gdy dystrybuanta F jest majoryzowana w porządku wypukłym przez uogólniony rozkład Pareto o dystrybuancie $W(\alpha)$ ($F \in IGFR(\alpha)$). Centralnym problemem, podobnie jak poprzednio jest wyznaczenie rzutu $g_n^{(k)}(W(\alpha)(x)) - 1$ ($g_n^{(k)}$ - gęstość $R_n^{(k)}$ dla rozkładu jednostajnego) na stożek analogiczny do powyżej zdefiniowanego.

W pracy [A4] stosuje się technikę maksymalizacji normy do otrzymania osiągalnych oszacowań dla wartości oczekiwanych standaryzowanych statystyk porządkowych dla klasy dystrybuant mających malejącą średnio gęstość (DDA). Wyniki otrzymano dla statystyk niskiego rzędu wykorzystując własność zmniejszającej się zmienności VDP dla wielomianów Bernsteina.

Prace [A1], [A2], [A5] i [A8] dotyczą analogicznych zagadnień dla uogólnionych statystyk porządkowych r -tego rzędu $X_\gamma^{(r)}$ dla różnych wariantów funkcjonału. Jedno-autorska praca [A8] bada właśnie takie zagadnienie: rozpatruje się w niej oszacowania wartości oczekiwanej scentrowanej i standaryzowanej $X_\gamma^{(r)}$, przy czym standaryzacja wynosi $\sigma_p = (E|X - \mu|^p)^{1/p}$. Podaje się w niej niedodatnie oszacowania górne w przypadku pierwszej uogólnionej statystyki, gdy $\gamma > 1$ oraz oszacowania dolne dla r -tych statystyk w zależności od kształtu gęstości $f_{\gamma,r}$ r -tej jednostajnej uogólnionej statystyki pozycyjnej.

Praca [A1] dotyczy tego samego zagadnienia co drugi problem w [A8] dla $p = 2$ i w sytuacji, gdy F majoryzuje pewną ustaloną dystrybuantę W w porządku wypukłym tj. $F^{-1}W$ jest wypukła. Problem sprowadza się do znalezienia oszacowania *górnego* dla funkcjonału $\int [(F^{-1}(u) - \mu)/\sigma] h_{\gamma,r}(u) du$, gdzie $h_{\gamma,r} = 1 - f_{\gamma,r}$, a po zamianie zmiennych i wykorzystaniu tw. Marigutiego, do znalezienia rzutu $\hat{h}_{\gamma,r} = h_{\gamma,r}W$ na stożek funkcji niemalejących i wypukłych w przestrzeni L^2 z wagą $w = W'$. Problem ten rozwiązuje się dla funkcji h spełniających warunek A1 będący modyfikacją warunku A (tw. 5 w pracy [A1]), a następnie dla $\hat{h}_{\gamma,r}$ dla W będącego uogólnionym rozkładem Pareto z $\alpha > -1/2$ (Proposition 3). Wy-

niki stosuje się do standaryzowanych statystyk porządkowych dla danych cenzurowanych typu II, będących przypadkiem szczególnym uogólnionych statystyk porządkowych.

Prace [A2] i [A5] zawierają oszacowania górne dla uogólnionych statystyk pozycyjnych dla klas rozkładów z malejącą gęstością oraz z malejącą intensywnością awarii oraz dla generowanych przez porządek gwiazdzisty.

Ocena głównego osiągnięcia naukowego

Praca [A7] zawiera istotny ogólny wynik (Proposition 1), który pozwala analizować wiele problemów szukania optymalnych oszacowań w jednolity sposób, przez redukcję problemu do problemu poszukiwania rzutu na zdefiniowany powyżej stożek $\mathcal{C}_{\leq c, W}$, gdzie W , w zależności od rozpatrywanej sytuacji, jest dystrybuantą rozkładu jednostajnego, wykładniczego, lub ogólnie, uogólnionego rozkładu Pareto. Mianowicie, warunek A postulujący postać rzutowanej funkcji jest na tyle ogólny, że pozwala objąć zarówno sytuację statystyk pozycyjnych jak i spacji, jak również, po odpowiedniej adaptacji, (Warunek A1 w [A1]) przypadek uogólnionych statystyk pozycyjnych. W szczególności podejście to pozwala uzyskać uprzedni wynik Rychlika (2014). Podejście to, eleganckie i użyteczne, sprowadza problem determinacji postaci rzutu, do problemu analitycznego wyznaczenia miejsca zerowania się pewnej pomocniczej funkcji, którą następnie analizuje się dla rozpatrywanych przypadków szczególnych. Zarówno Proposition 1 jak i wyznaczanie rzutów dla konkretnych przypadków wymagało bardzo konsekwentnej, i często żmudnej pracy, wskazującej na doskonałe opanowanie rzemiosła przez autorów. Praca [A6] zbiera niejako owoce podejścia wypracowanego w [A7]. Z nowych elementów warto zauważyć wykorzystanie wyników pracy do wyznaczenia ograniczeń górnych na oczekiwany czas życia układów niezawodnościowych. Praca [A3] nie zawiera również zasadniczo nowych pomysłów, lecz wymaga konsekwentnego i pracowitego wykorzystania metod [A7] dla jakościowo innej sytuacji analizy gęstości $R_n^{(k)}$.

Praca [A8] wykorzystuje również technikę rzutowania na stożek funkcji niemalejących i minimalizacji normy. Twierdzenie 4 w tej pracy wzmacnia twierdzenie Cramera i Kamps (2003) dla uogólnionej pierwszej statystyki pozycyjnej $X_\gamma^{(1)}$ dla $\gamma > 1$. Praca [A1] zawiera uogólnienie wyników Bieńka (2008) dotyczący rzutów na stożek funkcji niemalejących i wypukłych. Podobnie jak w przypadku [A7] zostało to uzyskane przez rozwiązanie problemu dla funkcji h spełniających warunek A1, a następnie zastosowanie tego wyniku do $\hat{h}_{\gamma, r}$.

Patrząc z nieco szerszej perspektywy: główne osiągnięcie naukowe dr Goroncy jak i większość prac w jej dorobku poświęcona analizie danych uporządkowanych opiera się na dwóch podejściach: metodzie rzutu na stożki i metodzie minimalizacji normy. Następnie te dwa podejścia stosuje się, metodycznie i konsekwentnie, do badania funkcjonałów dla różnych statystyk danych uporządkowanych jak przy różnych warunkach (głównie zadanych przez porządki) nałożonych na pierwotny rozkład generujący dane. Istotne Proposition 1 w pracy [A7] (i analogiczne Proposition 2 w [A1]) pozwalają znacznie ograniczyć liczbę tych problemów, ale nadal pozostaje ona bardzo duża. Jeśli praca wykonana przez A. Goroncy nie budzi mojego entuzjazmu, to głównie z powodu, że bardzo trudno znaleźć w pracach przekonującą motywację, dlaczego akurat te konkretne problemy były badane poza jedynie tą, że pojawia się w nich otwarty problem obliczeniowy, który warto rozwiązać. O zastosowaniach w niezawodności mówi się, ale w moim przekonaniu pozostają one poza głównym nurtem praktycznych aplikacji. Praktyczne przełożenie wyprowadzanych oszacowań jest w mojej opinii niewielkie: dotyczą one wartości oczekiwanej, a więc wielkości uśrednionej po wielu eksperymentach, a nie wartości dla konkretnego eksperymentu, jaką jest na przykład prawdopodobieństwo przekroczenia pewnego progu przez standaryzowaną statystykę. Pewną wsobność tematyki rozprawy potwierdza również analiza cytowalności prac wchodzących w jej skład: cytowane są tylko [A8] (8 cytowań, odczyt bazy Web of Science z 25.01.2021) i [A7] (7 cytowań), przy czym wszystkie cytowania pochodzą z prac badaczy polskich zajmujących się bardzo zbliżoną tematyką. Podobnie ma się sprawa z całym dorobkiem 'porządkowym': w odniesieniu do prac indeksowanych w WS, znalazłem tylko 6 cytowań w pracach, z których żaden autorów nie należał do powyższej grupy (ogólna liczba cytowań bez autocytowań wynosi 44, indeks Hirscha 6) .

Mimo tych uwag krytycznych, dotyczących przede wszystkim potencjalnych zastosowań, należy stwierdzić, że wysiłek badawczy włożony w prace podstawowego osiągnięcia po uwzględnieniu wkładu współautorów był bardzo duży i samodzielny, a prace głównego osiągnięcia są jednolite tematycznie. Nie mam wątpliwości, że osiągnięcie dr Goroncy spełnia wymogi ustawy o stopniach i tytułach naukowych. Mam jednak nadzieję, że na kolejnych etapach swojej pracy naukowej swoje imponującą sprawność rachunkową wykorzystywała ona do rozwiązywania problemów bliższych realnym zastosowaniom.

Ocena pozostałych osiągnięć naukowo-badawczych, dorobku dydaktycznego, popularyzatorskiego oraz współpracy międzynarodowej

Pozostałe osiągnięcia naukowe po doktoracie stanowi zbiór 11 prac (przed doktoratem Habilitantka opublikowała 5 prac), w większości opublikowanych w solidnych czasopismach statystycznych z listy JCR. Prace poświęcone są głównie tematyce habilitacyjnej danych uporządkowanych, ale 3 prace dotyczą zastosowań statystycznych w psychologii poznawczej i ekologii ([C1]-[C3]). Istotnie nowe problemy badawcze w porównaniu z głównym osiągnięciem naukowym to prace [C1] i [C5] dotyczące metod wyliczania momentów statystyk pozycyjnych dla ciągów zmiennych losowych niekoniecznie niezależnych i mających potencjalnie różne rozkłady oraz [C8] i [C9] dotyczące tematyki tak zwanych m -uogólnionych statystyk porządkowych. Szczególnie dwie pierwsze wymienione prace świadczą o tym, że dr Goroncy stara się rozszerzyć obszar swoich zainteresowań badawczych.

Z zainteresowaniem również odnotowałem istnienie zastosowaniowej pracy [C2], opublikowanej w prestiżowym piśmie *Ecology Indicators*, która wykorzystuje w analizie statystycznej nowoczesne metody uczenia maszynowego, a w szczególności zaproponowane niedawno metody wyboru zmiennych w oparciu o metodę lasów losowych (HapfelMeier i Ulm (2013)). Świadczy to bardzo dobrze o potencjale dr Goroncy i umiejętności przyswajania nowoczesnych narzędzi statystycznych.

Pozostały dorobek jest stosunkowo obszerny i niewątpliwie solidny, ale nie spotkał się z dużym zainteresowaniem: najczęściej cytowana praca z całego dorobku to praca [B1] przed doktoratem posiadająca 12 cytowań (wszystkie cytowania tej pracy, poza jednym, pochodzą od autorów polskich).

Niemniej, lektura autoreferatu i informacji o osiągnięciach naukowych utwierdziła mnie w przekonaniu, że dr Goroncy prowadzi systematyczną i intensywną działalność naukową, stale współpracuje z grupą polskich probabilistów i statystyków zajmujących się teorią uporządkowanych zmiennych losowych (T. Rychlik, M. Bieniek, A. Dembińska), aktywnie promuje swoje wyniki (5 zaproszonych wykładów na konferencjach międzynarodowych, ponadto 16 komunikatów na takich konferencjach). Była również przewodniczącą sesji na 2 konferencjach międzynarodowych. Na podkreślenie zasługuje stała i owocna (9 wspólnych prac) współpraca z promotorem pracy doktorskiej prof. T. Rychlikiem. Istotną współpracę międzynarodową nawiązała Habilitantka z prof. Kampsem podczas pobytu na stypendium DAAD, czego efektem była wspólna praca. Była i jest obecnie recenzentem prac naukowych

w 5-ciu czasopismach międzynarodowych. W polskim środowisku statystycznym jest osobą rozpoznawalną, wielokrotnie przedstawiała swoje wyniki na konferencjach statystycznych w Wiśle i Będlewie, była kierownikiem projektu Badawczego Sonata NCN i wykonawcą w projekcie badawczym kierowanym przez prof. Rychlika.

Również jej działalność dydaktyczna nie ogranicza się do wymaganych na jej stanowisku zajęć dydaktycznych (jest adiunktem na wydziale Matematyki i Informatyki UMK), prowadziła również kurs ze statystyki dla pracowników UMK oraz zajęcia popularyzujące statystykę dla gimnazjalistów i licealistów oraz w ramach Toruńskiego Festiwalu Nauki i Sztuki.

Moja ocena pozostałych osiągnięć jest zgodna z oceną podstawowego osiągnięcia naukowego.

Konkluzja

W mojej opinii główne osiągnięcie naukowe dr A. Goroncy stanowi znaczny wkład w rozwój problematyki własności funkcjonałów statystyk porządkowych i spełnia wymagania stawiane osiągnięciu naukowemu w dziedzinie nauk ścisłych i przyrodniczych, dyscyplina matematyka; zwyczajowe wymagania spełnia również jej pozostały dorobek. Podsumowując stwierdzam, że popieram wystąpienie do Rady Dyscypliny Matematyka UMK o przyznanie dr Goroncy stopnia doktora habilitowanego nauk ścisłych i przyrodniczych, dyscyplina matematyka.

Jan Mielniczek