

JAN MIELNICZUK

WKŁAD POLAKÓW W ROZWÓJ STATYSTYKI MATEMATYCZNEJ I STOSOWANEJ PO DRUGIEJ WOJNIE ŚWIATOWEJ

W artykule¹ omówię wkład Polaków pracujących na uniwersytetach, uczelniach technicznych i instytutach badawczych w rozwój statystyki matematycznej i stosowanej po roku 1945. Obejmuje on działania tych statystyków, którzy studiowali w Polsce i byli związani z nią w początkach swojej kariery naukowej. Czas powojenny został podzielony przeze mnie na trzy okresy: lata 1945–1970, 1970–1990 i okres po roku 1990 do dziś. W tym miejscu nie dyskutujemy wkładu najwybitniejszego polskiego statystyka, Jerzego Neymana, będącego tematem innych opracowań (por. Klonecki i Zonn, 1973; Klonecki, 1995 i Ledwina, 2012). Osiągnięcia polskich statystyków w biometrii przedstawiono w Caliński (2012).

LATA 1945–1970

Przy omawianiu wkładu Polaków do statystyki światowej warto zdać sobie sprawę z kondycji polskiej statystyki bezpośrednio po wojnie. Została ona trafnie opisana przez Jana Oderfelda: „*W statystyce matematycznej sytuacja jest wyjątkowo krytyczna. Tę gałąź nauki uprawiało przed 1939 rokiem bardzo niewielu specjalistów, z których po 1945 r. pozostała w kraju jedynie garstka.*”. Ze znanych stochastyków, którzy przeżyli wojnę, poza Polską pozostawali m.in. Jerzy Neyman, Stanisław Ulam, Marek Kac, Zygmunt Birnbaum (w USA) oraz Zbigniew Łomnicki i Stanisław Krystyn Zaremba (w Wielkiej Brytanii). Szczęśliwie byli matematycy biorący udział w organizacyjnym odtwarzaniu działalności polskiej matematyki po wojnie, tacy jak Edward Marczewski i Hugo Steinhaus, którzy rozumieli konieczność rozwijania zastosowań i statystyki. Działania organizacyjne i naukowe w tym kierunku zostały podjęte przede wszystkim we Wrocławiu i Warszawie, w nowo powstałych instytucjach: Uniwersytecie Wrocławskim i Politechnice Wrocławskiej oraz w Państwowym

¹ Artykuł jest zmodyfikowaną wersją wykładu wygłoszonego na Kongresie Statystyki Polskiej z okazji jubileuszu 100-lecia PTS, Poznań, kwiecień 2012.

Adres autora: Instytut Podstaw Informatyki PAN, Jana Kazimierza 5, 01-248 Warszawa, e-mail: miel@ipipan.waw.pl

Instytucie Matematycznym przekształconym wkrótce w Instytut Matematyczny Polskiej Akademii Nauk z oddziałami w Warszawie i Wrocławiu. W IM PAN zostaje utworzony Wydział Zastosowań Matematyki, którego Sekcja Aktuarialna przekształca się w Sekcję Statystyki Matematycznej (jej pierwszym kierownikiem był O. Lange). Powstają pisma: *Zastosowania Matematyki* (w 1953 roku) i *Listy Biometryczne* (w roku 1964). Głównymi animatorami rozwoju statystyki w tych instytucjach byli, we Wrocławiu: Hugo Steinhaus, Julian Perkal, Stefan Zubrzycki, Stanisław Trybuła, Józef Łukasiewicz i Bolesław Kopociński, a w Warszawie: Jan Oderfeld, Marek Fisz, Wiesław Sadowski, Oskar Lange i Mieczysław Warmus (który przeniósł się z Wrocławia do Warszawy w 1958 roku).

Centralną postacią i głównym animatorem rozwoju zastosowań matematyki i statystyki w pierwszym okresie powojennym był Hugo Steinhaus (1887–1972), pracujący na Uniwersytecie Wrocławskim i Instytucie Matematycznym PAN. Ten światowej renomy matematyk, współtwórca lwowskiej szkoły analizy funkcjonalnej, po wojnie, w wieku lat ponad pięćdziesięciu (w roku 1945 miał 57 lat) skupił się głównie na zastosowaniach matematyki i statystyce i robił to z wielkim talentem i zaangażowaniem. Jest on jednym z twórców (wspólnie z K. Florkiem, J. Łukasiewiczem, J. Perkalem i S. Zubrzyckim) tzw. taksonomii wrocławskiej, będącej zbiorem metod analizy danych opartych na konstrukcji najkrótszego dendrytu łączącego wszystkie punkty zbioru skończonego. Dendryt ten okazał się doskonałym narzędziem wizualizacji zbioru obiektów, odległości między którymi nie muszą być konieczne odległościami euklidesowymi i zadane są przez macierz ogólnie zdefiniowanych podobieństw. Metoda najkrótszego dendrytu jest dotąd stosowana przez praktyków wszędzie tam, gdzie istotne są metody grupowania obiektów i cech, m.in. w biologii, antropologii i astronomii. Steinhaus był również autorem nowatorskich prac dotyczących dochodzenia ojcostwa na podstawie pomiarów serologicznych i twórcą (wraz z J. Oderfeldem i K. Wiśniewskim) pierwszych polskich norm statystycznych dotyczących kontroli jakości. Prowadził też badania teoretyczne dotyczące bayesowskich estymatorów prawdopodobieństwa sukcesu i opublikował w 1957 roku na ten temat pracę *The problem of estimation* (Steinhaus, 1957). Był to pierwszy po wojnie artykuł w *Annals of Mathematical Statistics*, którego autorem był statystyk pracujący w Polsce. Ogromne znaczenie miały zainicjowane przez Steinhausza działania organizacyjne: założenie, przy współpracy Jana Oderfelda, pisma *Zastosowania Matematyki*, którego był pierwszym redaktorem naczelnym oraz zainicjowanie działalności i prowadzenie wtorkowego seminarium zastosowań matematyki w IM PAN we Wrocławiu. Może warto w tym miejscu przytoczyć zdania z artykułu wstępnego otwierającego pierwszy tom *Zastosowań Matematyki*, aby zdać sobie sprawę, jak bardzo poglądy i zamierzenia redaktorów pisma nie straciły dzisiaj na aktualności: *„Będziemy unikali artykułów, w których subtelność matematyczna jest celem samym w sobie, a moźół autora i czytelnika przewyższa korzyść poznawczą i praktyczną. Zastrzegamy sobie jednak ogłaszanie rezultatów, które nie służą bezpośrednio żadnej nauce poza samą matematyką, ale pośrednio służą im wszystkim naraz, przez to, że wyświełtlają kwestie*

zasadnicze i przyczyniają się do właściwego rozumienia matematyki i jej roli w kulturze i cywilizacji współczesnej.”.

Przedstawicielem następnego pokolenia polskich statystyków był Julian Perkal (1913–1965) (związany w różnych okresach życia z Uniwersytetem i Politechniką Wrocławską, Instytutem Matematycznym PAN i Wyższą Szkołą Rolniczą we Wrocławiu), wykształcony przed wojną matematyk (stopień magistra filozofii uzyskał w 1937 roku), który podobnie jak Steinhaus miał niezwykle szerokie zainteresowania i talent do zastosowań (por. Łukaszewicz (1967)). Najwięcej prac poświęcił analizie taksonomicznej i dyskryminacyjnej. Do jego najbardziej trwałych osiągnięć, obok udziału w stworzeniu taksonomii wrocławskiej należała koncepcja wskaźników Perkala, pozwalających stwierdzić, która z cech opisujących obiekt najbardziej wyróżnia go na tle innych obiektów. Bardzo ciekawą koncepcją zaproponowaną przez Perkala była konstrukcja nomogramów noszących dziś jego imię, stosowanych w ocenie rozwoju dzieci i będących ważną alternatywą dla powszechnie stosowanych nomogramów wzrostu i wagi dzieci. Julian Perkal był założycielem Polskiego Towarzystwa Biometrycznego (w 1961 roku) i pisma Listy Biometryczne (obecnie Biometrical Letters). Jego przedwczesna śmierć w wieku 53 lat znacznie osłabiła środowisko polskich statystyków, podobnie jak nagły zgon Stefana Zubrzyckiego, najzdolniejszego wrocławskiego ucznia Steinhaus (Stefan Zubrzycki zmarł w 1968 roku w wieku 41 lat).

Stefan Zubrzycki, związany z Uniwersytetem Wrocławskim i IM PAN był prekursorem badań dotyczących statystyki przestrzennej i opublikował na ten temat cykl znakomitych prac. Badania te motywowane potrzebami kopalnictwa w Polsce związane były z modelowaniem zasobności złóż $y(p)$ jako procesu losowego obserwowanego w pewnych punktach badanego obszaru D . Prace Zubrzyckiego dotyczyły między innymi problemu estymacji całkowitych zasobów złóż $V(D) = \iint_D y(p) dp$, w tym efektywności estymatorów liniowych tego parametru dla ustalonych punktów próbkowania p_1, \dots, p_n , problemu estymacji funkcji kowariancji dla obserwacji obciążonych błędami losowymi oraz wpływu kształtu regularnych sieci próbkowania na efektywność estymatora średniej \bar{Y} zasobów $V(D)$. Reprezentatywnymi pracami Zubrzyckiego dotyczącymi tych zagadnień są Zubrzycki (1957) i wspólna praca z Daleniusem i Hájkiem, Dalenius *et al.* (1961) (por. również bibliografię prac w Kopociński i Łukaszewicz, 1971). Podobnie jak Steinhaus i Perkal, Zubrzycki był niezwykle wszechstronnym badaczem i zajmował się również m.in. problemami kontroli jakości, oceny liczebności populacji i wykorzystaniem metod statystycznych w genetyce i astronomii. Napisał popularny podręcznik z rachunku prawdopodobieństwa i statystyki matematycznej (Zubrzycki, 1966).

Prace Stanisława Trybuły (1932–2008) (IM PAN, PWR) są istotnym wkładem w teorię estymacji minimaxowej i estymacji sekwencyjnej dla procesów stochastycznych oraz w teorię sterowania (por. Trybuła, 1958). Badania w zakresie statystyki kontynuowane są przez jego uczniów (R. Różański, R. Magiera, M. Wilczyński).

Z grupy statystyków warszawskich najtrwalszymi osiągnięciami z tego okresu są osiągnięcia Marka Fisz (1910–1963), pracującego w Polsce na Uniwersytecie

Warszawskim i IM PAN do 1960 roku. Do prekursorskich prac należą jego prace o rozkładach asymptotycznych pewnych nieparametrycznych statystyk testowych przy wykorzystaniu aparatu słabej zbieżności miar określonych na przestrzeniach metrycznych. Ta tematyka była rozwijana przez R. Bartoszyńskiego, promotorem pracy doktorskiej którego był właśnie Fisz. Był on również autorem doskonałego podręcznika z rachunku prawdopodobieństwa i statystyki, trzykrotnie wydanego w języku angielskim (Fisz, 1954).

Działalność Jana Oderfelda była ściśle związana z zastosowaniami matematyki i statystyki, przede wszystkim ze statystyczną kontrolą jakości. Warto wspomnieć tu o głównym nurcie działalności naukowej Oderfelda: był on znakomitym konstruktorem, przede wszystkim silników lotniczych, ale również pamięci zewnętrznych do maszyn matematycznych (jak kiedyś określano komputery). Z kolei, prace W. Sadowskiego (por. np. Łukaszewicz i Sadowski, 1960) i O. Langego dotyczyły przede wszystkim problemów statystycznych motywowanych lub związanych z problemami ekonomicznymi i ekonometrycznymi.

Mieczysław Warmus (1918–2007), pierwszy dyrektor powstałego w roku 1961 Centrum Obliczeniowego PAN, był jednym z prekursorów rozwoju metod statystyki obliczeniowej w Polsce. Prowadził również intensywną współpracę z lekarzami, dotyczącą m.in. modelowania i analizy statystycznej danych cukrzycowych i kardiologicznych. Jak można zorientować się nawet z tego krótkiego opisu, rozwój statystyki w pierwszym okresie powojennym dotyczył przede wszystkim statystyki stosowanej, a zagadnienia, które badano, były związane z potrzebami odbudowującej się gospodarki. Wyniki teoretyczne były nieliczne, ale za to znakomite (Fisz, Steinhaus, Trybuła, Zubrzycki).

LATA 1970–1990

Początek lat siedemdziesiątych ubiegłego stulecia przyniósł nowe działania organizacyjne mające ogromne znaczenie dla rozwoju statystyki polskiej. W 1972 roku powołano Komisję do spraw rozwoju Statystyki Matematycznej przy Komitecie Nauk Matematycznych PAN (jej pierwszym przewodniczącym był J. Łukaszewicz), rozpoczęto kształcenie statystyków na otwartych studiach doktoranckich przy IM PAN (Wrocław) i Uniwersytecie Wrocławskim (inicjatorami tych działań byli W. Klonecki i J. Łukaszewicz) oraz zaczęto organizować coroczne konferencje ze statystyki matematycznej (większość z nich odbyła się w Wiśle, pierwsza miała miejsce w 1973 roku, w roku 2012 odbyła się trzydziesta ósma jej edycja). W 1980 roku zaczyna wychodzić we Wrocławiu pismo *Probability and Mathematical Statistics* założone przez K. Urbanika, Cz. Rylla-Nardzewskiego i W. Kloneckiego. Większość statystyków, którzy byli uczestnikami wspomnianego studium doktoranckiego, wniosła znaczący wkład w rozwój statystyki światowej (T. Bednarski, S. Gnot, A. Kozek, T. Ledwina, M. Musiela, Cz. Stepniak, R. Zmyślony i S. Zontek). Do intensywnie rozwijanych

kierunków badań teoretycznych w statystyce matematycznej należały m.in.: analiza wielowymiarowa, przede wszystkim teoria modeli liniowych, teoria testowania hipotez statystycznych, metody statystycznej teorii decyzji, nieparametryczna estymacja krzywych, metody sekwencyjne i pomiar siły zależności stochastycznej. W wyborze kierunków badań podążano za najbardziej rozwijanymi w tym czasie kierunkami światowymi.

W. Klonecki zainicjował we Wrocławiu badania teoretyczne dotyczące tematyki modeli liniowych, która była równolegle intensywnie uprawiana w innych ośrodkach naukowych, przede wszystkim w Poznaniu i Lublinie. Rozwijano głównie teorię estymacji i testowania w modelach liniowych o różnorodnych strukturach macierzy kowariancji, w szczególności wnioskowanie o komponentach wariancyjnych, teorie planowania eksperymentu, w tym teorie układów blokowych, badano porządki stochastyczne związane z estymacją w modelach liniowych oraz stosowano wypracowane metody w konkretnych zagadnieniach biometrycznych. Rozwijano również metody statystycznej analizy wielowymiarowej, w tym analizy dyskryminacyjnej i analizy skupień. Główne postaci, dookoła których koncentrowały się powyższe badania, to: Jerzy Baksalary (1944–2005), Tadeusz Caliński, Bronisław Ceranka, Stanisław Gnot (1946–2002), Radosław Kala, Krystyna Katulska, Witold Klonecki (1930–2012), Mirosław Krzyśko, Wiktor Oktaba (1920–2009), Czesław Stępiak, Roman Zmysłony i Stefan Zontek. Warto wymienić kilka reprezentatywnych prac dla tego kierunku badań (modele liniowe): Zmysłony (1980), Gnot (1983), Gnot *et al.* (1985) (praca z J. Kleffe i R. Zmysłonym), Baksalary i Kala (1981), Stępiak *et al.* (1984) (praca z S. Wongiem i J. Wu), Farrell *et al.* (1984) (praca z W. Kloneckim i S. Zontkiem), Krzyśko (1983) (analiza dyskryminacyjna) i Caliński i Harabasz (1974) (analiza skupień). „Liniowcy” prowadzili intensywną współpracę międzynarodową, przede wszystkim ze statystykami niemieckimi.

Drugim intensywnie rozwijanym kierunkiem była statystyczna teoria decyzji, a w tym problemy porównania eksperymentów statystycznych i ich zastosowania w statystyce odpornej (Bednarski, 1982), badania nad teoriodecyzyjnym porównaniem estymatorów i szacowaniem ich ryzyka (Kozek, 1982a, 1982b; Brown i Gajek, 1990), nowe podejścia do estymacji odpornej (Zieliński, 1983) oraz redukowalności struktur statystycznych (Bartoszyński i Pleszczyńska, 1980).

Zostają zapoczątkowane w Polsce prace dotyczące tematyki testowania hipotez, bardzo intensywnie rozwijanej do dziś. Badania były prowadzone nad różnymi aspektami tej teorii: m.in. rozwijano pewną powojenną ideę Jerzego Neymana dotyczącą koncepcji tzw. Testów $C(\alpha)$ (Klonecki, 1977), prowadzono badania nad efektywnością testów, testami minimaxowymi i niezmienniczymi (Szkutnik, 1988) oraz nad testami używanymi do testowania hipotezy Hardy-Weinberga (Gnot i Ledwina, 1980).

Zajmowano się również nieparametryczną estymacją krzywych, przede wszystkim krzywej regresji i gęstości prawdopodobieństwa; tematyka ta w latach osiemdziesiątych ubiegłego stulecia należała do jednej z najżywiej rozwijanych dziedzin statystyki na świecie. Należy tu wymienić przede wszystkim osiągnięcia grupy W. Greblickiego

z Politechniki Wrocławskiej zajmującej się problemami zgodności estymatorów regresji (Greblicki *et al.*, 1984), jak również badania dotyczące estymacji gęstości metodą najmniejszej odległości i poprawiania estymatorów (Gajek, 1986a, 1986b), estymacji sekwencyjnej (Koronacki i Wertz, 1988) i estymacji gęstości w sytuacji ograniczonej obserwowalności próby (Mielniczuk, 1986).

We Wrocławiu i Warszawie prowadzono badania dotyczące metod sekwencyjnych (w 1981 r. R. Zieliński zorganizował w Warszawie międzynarodowy semestr w Centrum Banacha poświęcony tym zagadnieniom), we Wrocławiu zajmowano się wykorzystaniem tych metod w statystyce procesów stochastycznych (Musiała, 1981 i Magiera, 1987). W grupie E. Pleszczyńskiej w PAN prowadzono nowatorskie badania dotyczące pomiaru siły zależności stochastycznej przy użyciu pewnego zaproponowanego parametru funkcyjnego (Kowalczyk i Pleszczyńska, 1977 i Kowalczyk *et al.*, 1979). Owocem działalności naukowej tej grupy była również monografia *Statistical Inference. Theory and Practice* wydana w 1991 roku (wyd. polskie 1988), w której w bardzo nowoczesny sposób przedstawiano dwa nierozdzielne aspekty statystyki: badania teoretyczne i jej zastosowania.

Osobnego komentarza wymaga działalność Roberta Bartoszyńskiego (1933–1998) związanego w Polsce z warszawskim oddziałem IM PAN, wybitnego probabilisty, którego zainteresowania naukowe obejmowały również modelowanie matematyczne i statystykę. Z jego trwałych osiągnięć statystycznych należy tu wymienić klasyczną już pracę z Brownem,

McBride i Thompsonem (Bartoszyński *et al.*, 1981) na temat modelowania rozwoju nowotworu złośliwego i jego wykrycia, zakładającą poissonowski model wykrycia guza z intensywnością zależną od jego wielkości, dyskutującą również występujące tu problemy estymacji oraz prace dotyczące taksonomii opartej na subiektywnych klasyfikacjach (Bartoszyński, 1974) i modelu obiegu banknotów (Bartoszyński, 1972).

O ile pierwszy okres był związany przede wszystkim ze statystyką stosowaną i rozwiązywaniem problemów podsuwanych przez praktykę, o tyle okres następny przyniósł bardzo silny rozwój metod teoretycznych i osiągnięcie poziomu światowego w statystyce matematycznej. Świadczą o tym również, poza pracami w czołowych czasopismach statystycznych, takich jak *Annals of Statistics*, *Multivariate Analysis*, *Journal of Statistical Planning and Inference*, częste wizyty najlepszych statystyków zagranicznych w Polsce, którzy oprócz wizyt roboczych uczestniczyli m.in. w Sympozjum Neymanowskim w 1974 r., w semestrze statystycznym w Centrum Banacha w 1984 roku poświęconym metodom nieparametrycznym i odpornym (zorganizowanym przez T. Bednarskiego). Przyjeżdżali oni również licznie na konferencje w Wiśle, których stałym organizatorem w początkowym okresie był R. Zmyślony. Ważnym wydarzeniem było również Europejskie Spotkanie Statystyków (EMS) (Wrocław, 1980 r.).

LATA 1990 DO 2012

Cezurą ostatniego omawianego okresu jest rok 1990. Ponieważ mówimy o czasach nieodległych, ich omówienie staje się z konieczności coraz bardziej subiektywne i przede wszystkim odzwierciedla poglądy piszącego te słowa. Okres ten można określić jako okres konsolidacji, w którym pewne kierunki rozwijane poprzednio były kontynuowane, prace w innych praktycznie zamarły, natomiast pojawiły się zamiast nich nowe, rozwijane w Polsce obszary badań.

Intensywnie kontynuowano prace nad metodami testowania hipotez i modelami liniowymi. Rozwijano tematykę estymacji odpornej. Do nowych kierunków należą przede wszystkim: wnioskowanie dla łańcuchów Markowa i szeregów czasowych, optymalne oszacowania dla funkcjonalów statystycznych i problemy charakterystyczne związane z danymi uporządkowanymi oraz metody analizy trudnych danych, w szczególności danych o skomplikowanej strukturze lub dużej wymiarowości.

Tematyką bardzo konsekwentnie rozwijaną i najbardziej chyba rozpoznawalną na świecie jako tematyka badana w Polsce jest problematyka uprawiana przez grupę skupioną wokół Teresy Ledwiny dotyczącą testów adaptacyjnych. Została ona zapoczątkowana pracą z 1994 roku poświęconą adaptacyjnej wersji gładkiego testu Neymana dla problemu zgodności (Ledwina, 1994). Podejście to bazuje na opartym na danych wyborze jednego z potencjalnych modeli, a następnie konstrukcji w nim statystyki testowej było stosowane do otrzymania nowych rozwiązań w wielu problemach testowania (m.in. problemy testowania jednorodności, hipotez złożonych, niezależności, symetrii, hipotez w modelach regresyjnych i hipotezy o normalności). Równolegle badano koncepcje optymalności procedur testowych, w tym efektywności pośredniej wprowadzonej przez Kallenberg. Reprezentatywne prace grupy to Ingłot i Ledwina (1996), Kallenberg i Ledwina (1997), Ducharme i Ledwina (2003) i Ledwina i Wyłupek (2012).

Druga kontynuowana tematyka, konsekwentnie rozwijana przez T. Bednarskiego (IM PAN, UW) i jego współpracowników dotyczy konstrukcji estymatorów odpornych ze względu na zakłócenia modelowe, przede wszystkim dla modeli istotnych z punktu widzenia zastosowań, takich jak mieszane modele liniowe (Bednarski i Zontek, 1996), model Coxa (Bednarski, 1993) czy uogólniony model Poissona (Bednarski, 2004).

Wywodzące się z zastosowań prace Z. Szkutnika (AGH) dotyczą statystycznych problemów odwrotnych (Szkutnik, 2000, 2003). Ważne z punktu widzenia aplikacji są również badania grupy W. Niemiro (UW, UMK) nad teoretycznymi własnościami procedur Monte Carlo dla łańcuchów Markowa (MCMC), które wykorzystywane są m.in. w problemach estymacji stałoprecyzyjnej i w statystyce małych obszarów (Niemiro i Pokarowski, 2009 i Łatuszyński i Niemiro, 2011). T. Rychlik i współpracownicy (IM PAN i UMK) zajmują się własnościami statystyk bazujących na próbach uporządkowanych o ustalonej licznosci i przy dopuszczeniu zależności i niejednakowego rozkładu obserwacji oraz zagadnieniami pokrewnymi dotyczącymi

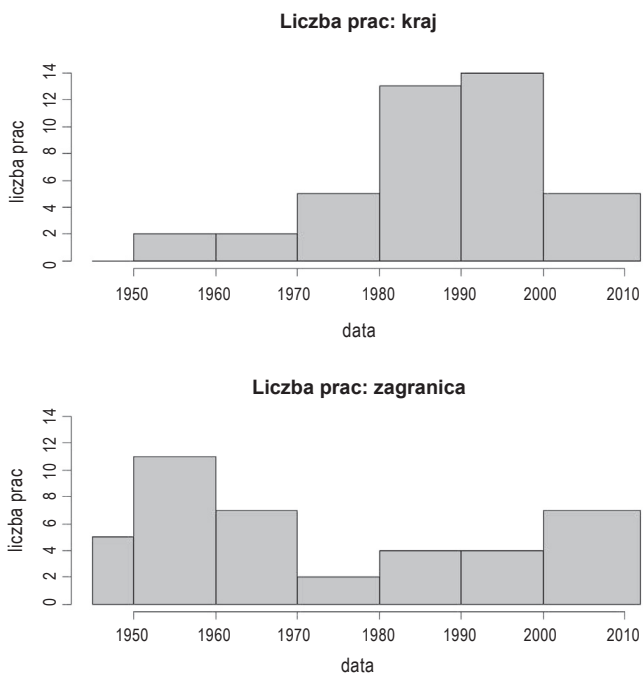
m.in. optymalnych oszacowań funkcjonałów statystycznych i zastosowań w teorii niezawodności (Rychlik, 1994; Gajek i Rychlik, 1998; Rychlik, 2001). Bliskie tym zagadnieniom są badania nad porządkami stochastycznymi prowadzone przez J. Bartoszewicza (UWr) i jego uczniów (por. np. Bartoszewicz, 1998). Zajęto się również tematyką wnioskowania dla szeregów czasowych, czyli danych obserwowanych w czasie, a prace dotyczyły szczególnie trzech obszarów: wnioskowania dla procesów niestacjonarnych o strukturze prawie periodycznej (Lenart *et al.*, 2008; praca z J. Leskowem i R. Synowieckim), dla procesów silnie zależnych (Csörgő i Mielniczuk, 1995a; Csörgő i Mielniczuk, 1995b i Mielniczuk *et al.*, 2010) i procesów z czasem ciągłym (por. np. Różański i Zagdański, 2006).

Liczne prace dotyczą rozwoju nowych metod wnioskowania związanych ze specyfiką badanych zjawisk empirycznych i związanych z nimi analizami „trudnych” danych. Specyfika ta może dotyczyć częściej np. w problemach genetycznych wysokiej wymiarowości danych przy jednoczesnej małej liczbie cech istotnie je opisujących (Bogdan *et al.*, 2004; Dрамиński *et al.*, 2008 i Bogdan *et al.*, 2012), specyficznej obserwowalności (Koronacki i Niemiro, 2005), prognozy dla obiektów bardziej skomplikowanych niż rozpatrywane w statystyce klasycznej, jak np. struktur białkowych (Pokarowski *et al.*, 2005) lub klasyfikacji wieloklasowej (Krzysko i Wołyński, 2009). Te coraz bardziej intensywne działania, które można zaliczyć do szeroko pojmowanego statystycznego wydobywania wiedzy z danych, są swoistą kontynuacją badań aplikacyjnych statystyków polskich zaraz po wojnie.

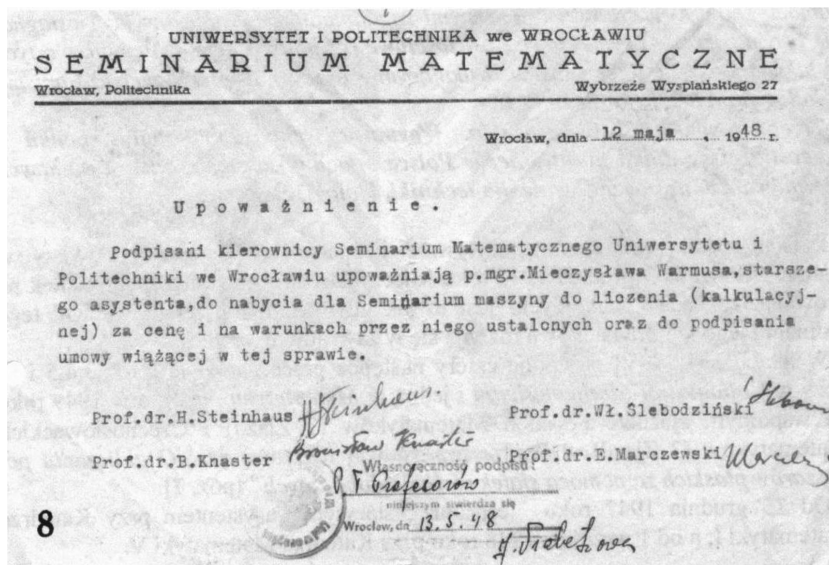
Warto wspomnieć o osiągnięciach Polaków w gałęziach probabilistyki stosowanej mających istotne znaczenie dla statystyki: w optymalizacji stochastycznej (R. Zieliński, J. Koronacki), teorii procesów punktowych (T. Rolski, W. Szczotka, R. Szekli) i teorii optymalnego stopowania (E. Ferenstein, K. Szajowski).

Dla zilustrowania wkładu do statystyki światowej posłużę się histogramami liczby prac opublikowanych przez Polaków w piśmie *Annals of Statistics* (i jego poprzedniku *Annals of Mathematical* oraz *Annals of Applied Statistics*, które się z *Annals of Statistics* wyodrębniło) w okresie powojennym do końca roku 2012 w rozbiciu na statystyków pracujących w kraju i pracujących poza jego granicami.

Łącznie opublikowano 81 prac. Moda w rozkładzie prac Polaków pracujących poza Polską jest głównie skutkiem niezwyklej płodności Zygmunta Birnbauma, który w latach [1945, 1970) opublikował w *Annals of Mathematical Statistics* 16 prac. Okres największej publikowalności statystyków krajowych w *Annals of Statistics* przypada na dwa ostatnie dekady ubiegłego wieku, gdy ukazało się łącznie 27 prac (13 w latach [1980, 1990) i 14 w latach [1990, 2000)). Ten wysyp prac ma niewątpliwie między innymi związek z działaniami organizacyjnymi podjętymi wcześniej i rosnącą świadomością wśród matematyków, że statystyka i matematyka stosowana winny być rozwijane. Poniżej, na zakończenie, przedstawiam wymowny przykład, że taka świadomość istniała już dużo wcześniej.



Rysunek 1. Histogramy liczby prac opublikowanych przez statystyków polskich pracujących w kraju i poza jego granicami w Annals of Statistics



Rysunek 2. Upoważnienie udzielone M. Warmusowi przez B. Knastera, E. Marczewskiego, H. Steinhaus i W. Ślebodzińskiego do zakupu „maszyny do liczenia”

LITERATURA

- [1] Baksalary J., Kala R., (1981), *Linear Transformations Preserving Best Linear Unbiased Estimators in a General Gauss-Markoff Model*, Ann. Statist., 9, 913–916.
- [2] Bartoszewicz J., (1998), *Characterizations of the Dispersive Order of Distributions by the Laplace Transform*, Stat. Probab. Letters, 40, 23–29.
- [3] Bartoszyński R., (1972), *Model of Circulation and Exchange of Banknotes*, Appl. Math., 13, 139–151.
- [4] Bartoszyński R., (1974), *On a Metric Structure Derived from Subjective Judgments: Scaling under Perfect and Imperfect Discrimination*, Econometrica, 42, 55–71.
- [5] Bartoszyński R., McBride C., Brown B., Thompson J., (1981), *Some Nonparametric Techniques for Estimating the Intensity Function of a Cancer Related Nonstationary Poisson Process*, Ann. Statist., 9, 150–160.
- [6] Bartoszyński R., Pleszczyńska E., (1980), *Reducability of Statistical Structures and Decision Problems*, Banach Center Publications, 6, 29–36.
- [7] Bednarski T., (1982), *Binary Experiments, Minimax Tests and 2-Alternating Capacities*, Ann. Statist., 10, 226–232.
- [8] Bednarski T., (1993), *Robust Estimation in the Cox Regression Model*, Scandinavian Journal of Statistics, 20, 213–225.
- [9] Bednarski T., (2004), *Robust Estimation in the Generalized Poisson Model*, Statistics, 38, 149–159.
- [10] Bednarski T., Zontek S., (1996), *Robust Estimation of Parameters in Mixed Unbalanced Models*, Ann. Statist., 24, 1493–1510.
- [11] Bogdan M., Chakraborti A., Frommlet F., Ghosh J., (2012), *Asymptotic Bayes Optimality under Sparsity of Some Multiple Testing Procedures*, Ann. Statist., 39, 1551–1579.
- [12] Bogdan M., Ghosh J., Doerge R., (2004), *Modifying the Schwarz Bayesian Information Criterion to Locate Multiple Interacting Quantitative Trait Loci*, Genetics, 167, 989–999.
- [13] Brown L., Gajek L., (1990), *Information Inequalities for the Bayes Risk*, Ann. Statist., 18, 1578–1594.
- [14] Caliński T., (2012), *Rozwój i osiągnięcia w biometrii polskiej*, Przegląd Statystyczny, Numer specjalny 1, 47–52.
- [15] Caliński T., Harabasz J., (1974), *Dendrite Method for Cluster Analysis*, Comm. Statist. Sim. Comp., 3, 1–27.
- [16] Csörgő S., Mielniczuk J., (1995a), *Density Estimation under Long-Range Dependence*, Ann. Statist., 23, 990–999.
- [17] Csörgő S., Mielniczuk J., (1995b), *Nonparametric Regression under Long-Range Dependence*, Ann. Statist., 23, 1000–1014.
- [18] Dalenius T., Hájek J., Zubrzycki S., (1961), *On Plane Sampling and Related Geometrical Problems*, Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, I, 125–150.
- [19] Damiński M., Rada-Iglesias A., Enroth S., Wadelius C., Komorowski J., Koronacki J., (2008), *Monte-Carlo Feature Selection for Supervised Classification.*, Bioinformatics, 24, 110–117.
- [20] Ducharme G., Ledwina T., (2003), *Efficient and Adaptive Nonparametric Tests for Two-Sample Problem*, Ann. Statist., 31, 2036–2058.
- [21] Farrell R., Klonecki W., Zontek S., (1984), *All Admissible Linear Estimators of the Vector of Gamma Scale Parameters with Application to Random Effects Models*, Ann. Statist., 12, 261–281.
- [22] Fisz M., (1954), *Rachunek Prawdopodobieństwa i Statystyka Matematyczna*, PWN, Warszawa.
- [23] Gajek L., (1986a), *Estimating a Density and its Derivatives via the Minimum Distance Method*, Probability Theory and Related Fields, 80, 601–617.
- [24] Gajek L., (1986b), *On Improving Density Estimators which Are Not Bona Fide Densities*, Ann. Statist., 14, 1612–1618.

- [26] Gajek L., Rychlik T., (1998), *Projection Method for Moment Bounds on Order Statistics from Restricted Families. II. Independent Case*, J. Multivar. Anal., 31, 156–182.
- [27] Gnot S., (1983), *Bayes Estimation in Linear Models: Coordinate Free Approach*, J. Mult. Anal., 13, 40–51.
- [28] Gnot S., Kleffe J., Zmyślony R., (1985), *Nonnegativity of Admissible Invariant Estimates in Linear Models with Two Variance Components*, J. Statist. Plann. Inference, 12, 249–258.
- [29] Gnot S., Ledwina T., (1980), *Testing for Hardy-Weinberg Equilibrium*, Biometrics, 36, 161–175.
- [30] Greblicki W., Krzyżak M., Pawlak M., (1984), *Distribution-Free Pointwise Consistency of Kernel Regression Estimate*, Ann. Statist., 12, 1570–1575.
- [31] Inglot T., Ledwina T., (1996), *Asymptotic Optimality of Data Driven Neyman's Tests for Uniformity*, Ann. Statist., 24, 1982–2019.
- [32] Kallenberg W., Ledwina T., (1997), *Data Driven Smooth Tests when the Hypothesis is Composite*, J. Amer. Statist. Assoc., 92, 1094–1104.
- [33] Klonecki W., (1977), *Optimal $C(\alpha)$ Tests for Homogeneity*, Proceedings of the Symposium to Honour J. Neyman, 161–175.
- [34] Klonecki W., (1995), *Jerzy Neyman (1894–1981)*, Probability and Mathematical Statistics, 15, 7–14.
- [35] Klonecki W., Zonn W., (1973), *Jerzy Splawa-Neyman*, Wiadomości Matematyczne, XVI, 53–70.
- [36] Koronacki J., Wertz W., (1988), *A Global Stopping Rule for Recursive Density Estimators*, J. Statist. Plann. Inference, 20, 23–39.
- [37] Koronacki S., Lasota K., Niemiro W., (2005), *Positron Emission Tomography by Markov Chain Monte Carlo with Auxiliary Variables*, Pattern Recognition, 38, 241–250.
- [38] Kowalczyk T., Kowalski A., Matuszewski A., Pleszczyńska E., (1979), *Screening and Monotonic Dependence Functions in the Multivariate Case*, Ann. Statist., 7, 607–614.
- [39] Kowalczyk T., Pleszczyńska E., (1977), *Monotonic Dependence Functions of Bivariate Distributions*, Ann. Statist., 5, 1221–1227.
- [40] Kozek A., (1982a), *Efficiency and Cramér-Rao Bounds for Convex Loss Functions*, J. Mult. Analysis, 12, 89–106.
- [41] Kozek A., (1982b), *Towards Calculus of Admissibility*, Ann. Statist., 10, 825–837.
- [42] Krzyśko M., (1983), *Asymptotic Distribution of the Discriminant Function*, Stat. Prob. Letters, 1, 243–250.
- [43] Krzyśko M., Wołyński W., (2009), *New Variants of Pairwise Classification*, European J. Operational Research, 199, 512–519.
- [44] Łatuszyński K., Niemiro W., (2011), *Rigorous Confidence Bounds for MCMC under a Geometric Drift Condition*, J. Complexity, 27, 23–38.
- [45] Ledwina T., (1994), *Data Driven Version of the Neyman Smooth Test of Fit*, J. Amer. Statist. Assoc., 89, 1000–1005.
- [46] Ledwina T., (2012), *Jerzy Neyman (1894–1981)*, Statistics in Transition – New Series, 13, 169–178.
- [47] Ledwina T., Wyłupek G., (2012), *Two-Sample Test Against One-Sided Alternatives*, Scand. J. Statist., 39, 758–781.
- [48] Lenart Ł., Leskow J., Synowiecki R., (2008), *Subsampling in Estimation of Autocovariance of PC Time Series*, J. Time Series Analysis, 29, 995–1018.
- [49] Łukaszewicz J., (1967), *Julian Perkal (1913–1965)*, Wiadomości Matematyczne, X, 29–36.
- [50] Łukaszewicz J., Sadowski S., (1960), *On Comparing Several Populations with a Control Population*, Zastosow. Mat., 5, 309–320.
- [51] Magiera R., (1987), *Admissible Sequential Polynomial Estimators for Stochastic Processes*, Sequential Analysis, X, 147–167.
- [52] Mielniczuk J. Zhou Z., Wu W., (2010), *On Nonparametric Prediction of Linear Processes*, J. Time Series Analysis, 30, 652–673.

-
- [53] Mielniczuk J., (1986), *Some Asymptotic Properties of Kernel Estimators of a Density Function in Case of Censored Data*, *Ann. Statist.*, 14, 1136–1139.
- [54] Musiela M., (1981), *On Sequential Estimation of Parameters of Continuous Gaussian Markov Processes*, *Probability and Mathematical Statistics*, 2, 37–53.
- [55] Niemiro W., Pokarowski P., (2009), *Fixed Precision MCM Estimation by Median of Products of Averages*, *J. Appl. Probab.*, 46, 309–329.
- [56] Pokarowski P., Kloczkowski A., Jernigan R., Kothari N., Pokarowska M., Koliński A., (2005), *Inferring Ideal Amino Acids Interaction Forms from Statistical Protein Contact Potentials*, *PROTEINS: Structure, Function and Bioinformatics*, 59, 49–57.
- [57] Rychlik T., (1994), *Distributions and Expectations of Order Statistics from Possibly Dependent Samples*, *J. Mult. Anal.*, 48, 31–42.
- [58] Rychlik T., (2001), *Projecting Statistical Functionals*, *Lecture Notes in Statistics 160*, Springer, New York.
- [59] Różański R., Zagdański A., (2006), *Estimation of the Drift Function for Ito Processes and Some Class of Semimartingales via Histogram Sieve*, *Appl. Math.*, 33, 21–40.
- [60] Steinhaus H., (1957), *The Problem of Estimation*, *Ann. Math. Statist.*, 28, 633–648.
- [61] Stępnik C., Wong S., Wu J., (1984), *Comparison of Linear Experiments with Known Covariances*, *Ann. Statist.*, 12, 358–365.
- [62] Szkutnik Z., (1988), *Most Powerful Invariant Tests for Binormality*, *Ann. Statist.*, 16, 292–301.
- [63] Szkutnik Z., (2000), *Unfolding Intensity Function of a Poisson Process in Models with Approximately Specified Folding Operator*, *Metrika*, 52, 11–26.
- [64] Szkutnik Z., (2003), *Doubly Smoothed EM Algorithm for Statistical Inverse Problems*, *J. Amer. Statist. Assoc.*, 98, 178–192.
- [65] Trybuła S., (1958), *Some Problems of Simultaneous Minimax Estimation*, *Ann. of Math. Statist.*, 29, 245–253.
- [66] Zielinski R., (1983), *Robust Statistical Procedures: A General Approach*, *Lecture Notes in Statistics*, 28, 285–295.
- [67] Zmyślony R., (1980), *A Characterization of Best Linear Unbiased Estimators in the General Linear Model*, *Lecture Notes in Statistics*, 2, 365–378.
- [68] Zubrzycki S., (1957), *O szacowaniu parametrów złóż geologicznych*, *Zastosowania Matematyki*, 3, 105–153.
- [70] Zubrzycki S., (1966), *Wykłady z Rachunku Prawdopodobieństwa i Statystyki Matematycznej*, PWN, Warszawa.