L Konferencja

STATYSTYKA MATEMATYCZNA

Chęciny, 8-11 grudnia 2025

WYDZIAŁ MATEMATYKI STOSOWANEJ AGH

WYDZIAŁ MATEMATYKI, INFORMATYKI I MECHANIKI

rok zał. 1919

Organizatorzy konferencji

Wydział Matematyki Stosowanej, Akademia Górniczo-Hutnicza w Krakowie
Wydział Matematyki, Informatyki i Mechaniki, Uniwersytet Warszawski
Polskie Towarzystwo Matematyczne

Instytucja wspierająca

Uniwersytet im. Adama Mickiewicza w Poznaniu

Komitet Organizacyjny

Anna Dudek (przewodnicząca)
Bogdan Ćmiel
Bartosz Majewski
Jakub Wojdyła
Błażej Miasojedow
Jan Mielniczuk
Jacek Miękisz

Miejsce konferencji

Europejskie Centrum Edukacji Geologicznej w Chęcinach
Korzecko 1C, 26-060 Chęciny

Sponsor zewnętrzny

AstraZeneca

Redakcja
Jakub Wojdyła

Pomoc przy redakcji
Zuzanna Maciszewska, Dominika Mosur, Bartosz Majewski

Przedmowa

W dniach 6–12 grudnia 1973 roku odbyła się w Wiśle konferencja „Statystyka matematyczna", inicjująca kontynuowaną w następnych latach serię spotkań środowiska polskich statystyków. I chociaż kolejne konferencje odbywały się w różnych miejscowościach, to właśnie Wisła pozostała tym miejscem szczególnym, do którego często powracano i które w środowisku naukowym stało się symbolem rozpoznawczym tego cyklu konferencji. Pełną listę kolejnych edycji tego cyklu, dokumentującą jego wieloletnią historię, można znaleźć na stronie `https://www.sm2025.agh.edu.pl/o-konferencji.html`.

Pierwsza konferencja została zorganizowana przez Komisję ds. Rozwoju Statystyki Matematycznej w Polsce, powołaną rok wcześniej przez Komitet Nauk Matematycznych Polskiej Akademii Nauk. Inicjatorem powstania Komisji był prof. Tadeusz Caliński, a pierwszym przewodniczącym — prof. Józef Łukaszewicz. Więcej informacji o historii i bieżącej działalności Komisji, która w późniejszym czasie została przemianowana na Komisję Statystyki, można znaleźć na stronie internetowej `https://www.ibspan.waw.pl/komisja-statystyki/`.

Na przestrzeni lat konferencje „Statystyka matematyczna" zyskiwały coraz większe znaczenie, przyciągając uczestników nie tylko z Polski, ale również z zagranicy. Wysoki poziom merytoryczny referatów, wykłady zaproszonych gości, w tym wybitnych statystyków o międzynarodowej renomie, konkurs na najlepszą prezentację dla młodych badaczy, dyskusje i kuluarowe rozmowy będące często zalążkiem przyszłej współpracy naukowej oraz towarzyszące obradom spotkania towarzyskie, stworzyły klimat, dzięki któremu konferencje stały się wydarzeniem unikatowym, niezwykle ważnym dla polskich statystyków — zarówno uczonych zajmujących się teorią, jak i praktyków oraz młodych badaczy stawiających pierwsze kroki w tej dziedzinie.

Obecnie konferencje z tego cyklu postrzegane są jako wydarzenie wyjątkowe i niezwykle istotne dla polskiej społeczności statystycznej. Integrują badaczy zajmujących się teorią statystyki, praktyków wykorzystujących metody statystyczne w różnych dziedzinach oraz doktorantów i młodych naukowców rozpoczynających swoją drogę zawodową. Niezależnie od zmian organizacyjnych i miejsca obrad, „wiślańskie" spotkania stały się trwałym elementem tradycji, podkreślającym ciągłość i współpracę środowiska statystyków w Polsce.

Tegoroczna pięćdziesiąta konferencja „Statystyka Matematyczna — Chęciny 2025" ma wyjątkowy charakter. Pragniemy bowiem nie tylko wymienić się najnowszymi wynikami z zakresu statystyki matematycznej i jej zastosowań, lecz także uczcić historię i dorobek całego cyklu konferencji. W programie przewidziano m.in. specjalną sesję jubileuszową, podczas której wybitni polscy statystycy, prof. Jacek Koronacki oraz prof. Roman Zmyślony, opowiedzą o historii konferencji, jej rozwoju oraz znaczeniu dla środowiska naukowego w Polsce.

Konferencja odbędzie się w dniach 8–11 grudnia 2025 r. w Europejskim Centrum Edukacji Geologicznej w Chęcinach (`https://www.eceg.uw.edu.pl`). Jej organizatorami są: Wydział Matematyki Stosowanej Akademii Górniczo-Hutniczej w Krakowie, Wydział Matematyki, Informatyki i Mechaniki Uniwersytetu Warszawskiego, Polskie Towarzystwo Matematyczne oraz Komisja Statystyki Komitetu Nauk Matematycznych PAN.

Życzymy Państwu inspirujących wystąpień, owocnych dyskusji, a także wielu okazji do spotkań, wspomnień i refleksji nad półwieczem „Statystyki matematycznej" w Polsce.

W imieniu Komitetu Organizacyjnego

dr hab. Anna Dudek

Preface

In December 6–12, 1973, the conference "Mathematical Statistics" was held in Wisła, initiating a series of meetings of the Polish statistical community that continued in the following years. And although subsequent conferences took place in various locations, it was Wisła that remained the special place to which people often returned and which, within the academic community, became the hallmark of this conference meetings. A full list of subsequent editions of the series, documenting its long-standing history, can be found at `https://www.sm2025.agh.edu.pl/o-konferencji.html`.

The first conference was organized by the Commission for the Development of Mathematical Statistics in Poland, established a year earlier by the Committee on Mathematics of the Polish Academy of Sciences. The initiator of the Commission's creation was Prof. Tadeusz Caliński, and its first chair was Prof. Józef Łukaszewicz. More information on the history and current activities of the Commission can be found at the website `https://www.ibspan.waw.pl/komisja-statystyki/`.

Over the years, the "Mathematical Statistics" conferences grew in importance, attracting participants not only from Poland but also from abroad. The high substantive level of the presentations, the lectures by invited guests, including internationally renowned statisticians, the competition for the best presentation by young researchers, discussions and informal conversations that often became the seed of future scientific collaboration, as well as the social events accompanying the sessions, created an atmosphere that made the conferences a unique event. They became extremely important for Polish statisticians: for theoreticians, practitioners, and young researchers taking their first steps in the field.

Today, the conferences in this series are regarded as exceptional events of great importance to the Polish statistical community. They bring together researchers working on statistical theory, practitioners applying statistical methods across diverse fields, as well as doctoral students and early-career scientists at the beginning of their professional paths. Regardless of organizational changes or the location of the meetings, the "Wisła" conferences have become a lasting part of the tradition, emphasizing the continuity and collaboration within the community of statisticians in Poland.

This year's fiftieth conference, "Mathematical Statistics — Chęciny 2025", has a special character. We wish not only to discuss the latest results in mathematical statistics and its applications, but also to celebrate the history and achievements of the entire conference series. The program includes, among other things, a special jubilee session during which distinguished Polish statisticians, Prof. Jacek Koronacki and Prof. Roman Zmyślony, will talk about the history of the conference, its development, and its importance for the academic community in Poland.

The conference will take place on 8–11 December 2025 at the European Centre for Geological Education in Chęciny (`https://www.eceg.uw.edu.pl`). It is organized by the Faculty of Applied Mathematics of the AGH University of Krakow, the Faculty of Mathematics, Informatics and Mechanics of the University of Warsaw, the Polish Mathematical Society, and the Commission of Statistics of the Committee on Mathematics of the Polish Academy of Sciences.

We wish you inspiring presentations, fruitful discussions, and many opportunities to meet, reminisce, and reflect on half a century of "Mathematical Statistics" in Poland.

On behalf of the Organizing Committee

dr hab. Anna Dudek

Mirosław Krzyśko

Historia Komisji Statystyki Matematycznej

Zachęcony przez prof. Przemysława Grzegorzewskiego postanowiłem krótko opisać historię Komisji Statystyki Matematycznej.

Jej historia rozpoczęła się w Poznaniu od spotkania w roku 1972 prof. Władysława Orlicza, ówczesnego przewodniczącego Komitetu Nauk Matematycznych PAN, z doc. Tadeuszem Calińskim, który zwrócił się do prof. Władysława Orlicza z prośbą o podjęcie intensywnych kroków dotyczących rozwoju statystyki matematycznej w Polsce. Wynikiem tego spotkania było zaproszenie doc. Tadeusza Calińskiego na posiedzenie Komitetu Nauk Matematycznych PAN w dniu 16 czerwca 1972 roku. Na tym posiedzeniu, doc. Tadeusz Caliński przedstawił stan statystyki matematycznej w Polsce. Następnie w dniu 7 września 1972 roku przesłał na ręce prof. Bogdana Bojarskiego, sekretarza naukowego Komitetu Nauk Matematycznych PAN, pismo precyzujące zakres działania oraz skład Komisji ds. Rozwoju Statystyki Matematycznej w Polsce.

W dniu 19 września 1972 roku Prezydium Komitetu Nauk Matematycznych PAN na posiedzeniu wyjazdowym w Lublinie powołało Komisję. Stwierdzono wtedy, że Komitet w miarę swych organizacyjnych i finansowych możliwości będzie udzielał Komisji wszelkiej niezbędnej pomocy organizacyjnej i finansowej.

Pierwsze posiedzenie plenarne Komisji odbyło się w dniu 13 lutego 1973 roku we Wrocławiu. Pozwolę sobie przytoczyć skład osobowy Komisji I Kadencji: Prof. Józef Łukaszewicz (przewodniczący), doc. Tadeusz Caliński (wiceprzewodniczący), doc. Bolesław Kopociński (sekretarz), doc. Robert Bartoszyński, doc. Witold Klonecki, dr Mirosław Krzyśko, dr Edward Niedokos, prof. Wiktor Oktaba, dr Elżbieta Pleszczyńska, doc. Stanisław Trybuła, prof. Mieczysław Warmus, dr Ryszard Zieliński. Zauważmy, że w Komisji reprezentowane były cztery ośrodki naukowe: Wrocław, Warszawa, Poznań i Lublin.

Dostąpiłem ogromnego zaszczytu bycia członkiem tej Komisji, jako świeżo upieczony doktor. Od roku 1975 Komisją kierował prof. Tadeusz Caliński, a ja zostałem jej sekretarzem.

Pierwsza konferencja pod patronatem Komisji odbyła się w Wiśle w dniach 6–12 grudnia 1973 roku. W latach 1973–1990 zorganizowano 16 konferencji, w tym 8 pierwszych w Wiśle. Sześć spośród nich to konferencje międzynarodowe.

Pierwsze konferencje wiślane nieodłącznie kojarzą mi się z Romanem Zmyślonym, wspaniałym organizatorem tych konferencji. Niezapomniane wrażenie pozostało także po Skalnicy, miejscu obrad. W przeszklonej sali wiało chłodem, a uczestnicy byli okryci kocami. Ale początki są zawsze trudne.

W latach 1991–2012, z woli Komitetu Matematyki PAN, kierowanego przez prof. Andrzeja Białynickiego-Birulę, przewodniczyłem tej Komisji. W tym czasie zorganizowane zostały 22 konferencje ze statystyki matematycznej, w tym 16 w Wiśle, 2 w Będlewie i po jednej w Poznaniu, Jachrance, Łagowie oraz Szklarskiej Porębie. Sześć z nich były konferencjami międzynarodowymi.

Słowo „Wisła”, w środowisku statystyków i nie tylko, jest rozpoznawalną marką do dzisiaj.

Z biegiem lat warunki zakwaterowania i obrad poprawiały się. Mile wspominam Gawrę, znakomity hotel z bardzo dobrym zakwaterowaniem i wyżywieniem, wieczorne spotkania z opowieściami Antoniego Dawidowicza, wspaniałe bankiety i wycieczki po okolicach Wisły. Ale to już wspomnienia.

Moim głównym zadaniem, jako przewodniczącego Komitetu, było wskazywanie organizatorów poszczególnych konferencji, zadanie trudne, bo chętnych było niewielu oraz zabieganie o środki finansowe na organizację kolejnych konferencji. Korzystny był grudniowy termin odbywania konferencji. Nikt w tym terminie już konferencji nie organizował. Ko-

mitetowi Matematyki pod koniec roku zostawały zawsze pewne kwoty i zwracano się do mnie z propozycją ich wykorzystania. Nigdy nie odmawiałem.

Oprócz organizacji konferencji ze statystyki matematycznej warto wspomnieć o innych działaniach przyczyniających się do rozwoju statystyki matematycznej w Polsce.

Dzięki życzliwości prof. Władysława Orlicza, w roku 1966 w Poznaniu, powołano środowiskowe seminarium ze statystyki matematycznej i jej zastosowań, którym kierował doc. Tadeusz Caliński. Na seminarium tym gościło wielu znakomitych statystyków z całego świata, między innymi prof. C.R. Rao. Byłem pierwszym wykształconym przez profesora Tadeusza Calińskiego doktorem. Doktoraty były przeprowadzane przez moją macierzystą Uczelnię, Uniwersytet im. Adama Mickiewicza. Przetarłem szlak. Szlakiem tym podążyły jeszcze 24 inne osoby wypromowane przez Profesora, między innymi profesorowie: Jerzy Baksalary, Bronisław Ceranka, Radosław Kala, Michał Karoński oraz Stanisław Mejza.

Na początku lat siedemdziesiątych we Wrocławskim Oddziale Instytutu Matematycznego PAN uruchomione zostały studia doktoranckie ze statystyki matematycznej kierowane przez doc. Witolda Kloneckiego. Z tych studiów wywodzą się tacy znakomici profesorowie jak: Tadeusz Bednarski, Stanisław Gnot, Teresa Ledwina, Czesław Stępniak oraz Roman Zmyślony.

W dniach 21–23 września 1970 roku w Lublinie, z inicjatywy prof. Wiktora Oktaby, zorganizowana została pierwsza konferencja pod nazwą Kolokwium Biometryczne. W tym roku w Olsztynie w dniach 7–10 września odbyło się pięćdziesiąte czwarte Międzynarodowe Kolokwium Biometryczne. Konferencje te przyczyniły się w znakomitym stopniu do rozwoju biometrii i statystyki matematycznej w Polsce.

W tym roku w Chęcinach organizowana jest po raz pięćdziesiąty konferencja „Statystyka Matematyczna", a więc Jubileusz. Niestety ze względów zdrowotnych nie wezmę udziału w tej konferencji. Tą drogą Uczestnikom przesyłam serdeczne pozdrowienia i życzenia miłego spędzenia czasu konferencyjnego.

Zaproszeni wykładowcy

Rafał Kulik

University of Ottawa, Kanada

Rafał Kulik jest profesorem zwyczajnym w Instytucie Matematyki i Statystyki Uniwersytetu w Ottawie. Uzyskał tytuł magistra na Uniwersytecie w Ulm (Niemcy) oraz stopień doktora na Uniwersytecie Wrocławskim (Polska). Po doktoracie spędził trzy lata na stażu podoktorskim na Uniwersytecie w Ottawie. Następnie został zatrudniony jako profesor w Szkole Matematyki i Statystyki Uniwersytetu w Sydney (Australia). Do Uniwersytetu w Ottawie dołączył w 2007 roku.

Zainteresowania badawcze Rafała Kulika obejmują: twierdzenia graniczne dla procesów stacjonarnych i niestacjonarnych, szeregi czasowe z ciężkimi ogonami, teorię wartości ekstremalnych, prywatność danych. Opublikował dwie monografie naukowe i ponad 60 artykułów w wiodących czasopismach naukowych. Jego badania nad teorią wartości ekstremalnych i uczeniem maszynowym są obecnie finansowane przez NSERC i CANSSI (Kanadyjski Instytut Nauk Statystycznych), podczas gdy badania nad prywatnością danych są finansowane przez MITACS i Biuro Komisarza ds. Prywatności.

Rafał Kulik jest członkiem komitetów redakcyjnych w *Stochastic Processes and their Applications*, *Extremes*, *Electronic Journal of Statistics*, *Canadian Journal of Statistics*. Od 1 stycznia 2026 roku obejmie stanowisko redaktora naczelnego czasopisma *Extremes*.

Rafał Kulik pełnił funkcję dyrektora Instytutu Matematyki i Statystyki w latach 2016–2019 i 2022–2023. Jest członkiem Rady Gubernatorów Uniwersytetu w Ottawie oraz Grupy Ewaluacyjnej ds. Matematyki i Statystyki NSERC (National Science and Engineering Research Council).

Non-stationary time series with heavy tails

Data in finance, insurance, climate or machine learning algorithms exhibit non-standard features such as non-stationarities, slowly decaying correlations or heavy tails. Such data cannot be modelled by classical stationary Gaussian-like time series. Rather, one needs to consider non-stationary time series with heavy tails. Dealing with such time series may lead to non-standard behaviour in central limit theorem (CLT) or extreme value theory (EVT). As such, the goal of this talk is to review current results on heavy tails and extremes of non-stationary heavy-tailed time series.

FIRST LECTURE:

1) We will start with briefly reviewing briefly limit theorems (CLT and EVT) for stationary, light-tailed (Gaussian-like) time series.

2) In the second step we will extend the aforementioned results to stationary, heavy-tailed time series.

3) In the next step, we will present different non-stationary models that have appeared in the literature, underlying practical motivations (data modelling, ML algorithms).

4) We will then discuss different limit theorems for very simple non-stationary models. In articular, we will give limiting results for sample maxima. These results will reveal different behaviour of the maxima for the entire sample and for blocks. It has an influence on statistical inference. For example, the classical estimators of the tail index (that characterizes how heavy is the tail) are inconsistent just due to non-stationarity.

SECOND LECTURE:

In this lecture I will discuss probabilistic properties of a special class of non-stationary time series that stem from Stochastic Gradient Descent, the optimization algorithm that is widely used in Machine Learning. It was empirically observed that some characteristic of SGD exhibit heavy tails.

In this talk we will (partially) explain the mechanisms that lead to emergence of heavy tails in both online and offline SGDs. Furthermore, we will study SGD through the lenses of stochastic processes obtaining stable convergence.

We will finish with a list of (many) open problems.

Hernando Ombao

King Abdullah University of Science and Technology, Arabia Saudyjska

Hernando Ombao jest profesorem statystyki na King Abdullah University of Science and Technology (KAUST) w Arabi Saudyjskiej. Pełni funkcję kierownika grupy badawczej Biostatistics Group. Jego badania dotyczą modeli statystycznych dla szeregów czasowych o dynamicznych i złożonych strukturach, motywowane problemami z zakresu neurobiologii, medycyny oraz zdrowia publicznego. Przed podjęciem pracy w KAUST pracował na uniwersytetach w Stanach Zjednoczonych: University of California, Irvine; Brown University; University of Illinois oraz University of Pittsburgh.

Profesor Ombao jest współautorem monografii „Handbook of Statistical Methods for Neuroimaging". W 2019 roku pełnił funkcję przewodniczącego sekcji Statistics in Imaging American Statistical Association (ASA), a w 2020 roku przewodniczącego komitetu nagród Council of Presidents of the Statistical Societies. Był członkiem panelu Biostatystyka w sekcji oceny projektów Narodowych Instytutów Zdrowia (NIH). W 2017 roku otrzymał nagrodę Mid-Career za wybitne osiągnięcia badawcze na UC Irvine. Był także głównym wykonawcą kilku projektów finansowanych przez US National Science Foundation.

Pełnił funkcję zastępcy redaktora w czasopiśmie *Journal of the Royal Statistical Society, Series B*, a obecnie pełni tę funkcję w czasopismach *Journal of the American Statistical Association*, *Annals of Applied Statistics* oraz *Data Science in Science*. Jest także współzałożycielem i współredaktorem nowego czasopisma *Statistics and Data Science in Imaging*. W 2016 roku został wybrany członkiem (Fellow) American Statistical Association, a w 2024 roku — Institute of Mathematical Statistics.

Overview of Methods for Characterizing Dependence in Multivariate Time Series

Multivariate time series data is being collected in all facets of life. In the world of finance, stock shares from different companies and sectors. In medical science, blood glucose, heart rate, blood pressure are continuously measured by wearable devices. In neuroimaging, brain signals are observed from many different locations or nodes in a brain network. There is a canonical goal in analyzing these data. The stock trader or personal investor is keen to understand the dynamics across stocks, in particular, how a stock index in one company may influence that of another. In neuroscience, modeling dependence between nodes in a brain network is central to understanding underlying neural mechanisms such as perception, action, and memory. In this talk, structured into two parts, we present a broad range of statistical methods for characterizing dependence in a brain network. We first review a general framework which decomposes each signal into various frequency components and then characterize the dependence properties through these oscillatory activities. The unifying theme across the talk is to explore the strength of dependence and possible lead-lag dynamics through filtering. The proposed framework has the capacity

to represent both linear and non-linear dependencies that could occur instantaneously or after some delay (lagged dependence).

In the first part of the talk, some of the most prominent frequency domain measures such as coherence, partial coherence, and dual-frequency coherence can be derived as special cases under this general framework. Coherence can be viewed in the proposed framework as the maximal squared correlation between oscillations from a pair of nodes. It gives a more specific information that correlation because it identifies the frequency bands that drive the dependence between nodes. This is followed by a discussion on methods for characterizing dependence between a pair of communities (of nodes) under the high dimensional setting. Here, we will introduce our proposed Kendall's Canonical Coherence (KenCoh). The classical method of Canonical Correlation Analysis (CCA) is limited to capturing linear associations, cross-sectional studies, and is sensitive to heavy-tailed observations. The proposed KenCoh method mitigates the limitations by using a rank-based approach under elliptic symmetry. It captures non-linear associations and is robust against outliers. The work presented here is limited to multiple independent trials of multivariate time-series, which can be extended to accout for across-trials correlation which is the usual set-up of experimental studies.

In the second part of the talk, we will briefly describe current work on information-theoretic measures of connectivity. This class of measures are based on joint (and conditional) distributions rather than just the first two moments. The proposed spectral transfer entropy (STE) quantifies the magnitude and direction of information flow from a certain frequency-band oscillation of a channel to an oscillation of another channel. The main advantage of our proposed approach is that it allows adjustments for multiple comparisons to control family-wise error rate. Another novel contribution is a simple yet efficient estimation method based on vine copula theory that enables estimates to capture zero (boundary point) without the need for bias adjustments. With the vine copul representation, a null copula model, which exhibits zero STE, is defined, making significance testing for STE straightforward through a standard resampling approach. The advantages of our proposed measure through some numerical experiments and provide interesting and novel findings on the analysis of EEG recordings linked to a visual task. Following the discussion on STE, we will present an approach to testing for (Granger) causality using deep neural networks. Here, the conditional dependence of expressed via the conditional mean which is a non-specified functional of all past observations. This approach is very general though highly computationally intensive and thus we will adopt the principle of dropouts to execute the computation. The last part of this talk will cover some new techniques for characterizing lead-lag dependence which may differ across different scales. For example, the lead-lag relationship between a pair of stocks may differ depending on the time scale (e.g., hourly vs weekly).

This is joint work with the Mara Talento, Sarbojit Roy, Sipan Aslan, Malik Sultan and Paolo Redondo of the Biostatistics Group at KAUST. Moreover, Adam Sykulski (Imperial College) is also collaborating with the KAUST Group on the last topic on frequency-specific lead-lag dependence.

Sesja historyczna

Jacek Koronacki

Od roku 2021 jest profesorem emerytowanym Instytutu Podstaw Informatyki PAN, którego wcześniej był długoletnim dyrektorem. Tytuł profesora nauk technicznych uzyskał w roku 1999. Stopień dra habilitowanego nauk matematycznych w zakresie matematyki (statystyka matematyczna) nadała mu Rada Naukowa Instytutu Matematycznego PAN w roku 1988.

Prace dra J. Koronackiego z lat 1970-tych, dotyczące algorytmów optymalizacji metodami szukania losowego, wskazały nowy kierunek rozwoju tego obszaru badań. Prace dra hab. Koronackiego z przełomu lat 1980-tych i 1990-tych, dotyczące trudnych problemów praktycznej estymacji funkcji statystycznych, zostały zauważone przez uznanych specjalistów na świecie i miały pewien udział w wyznaczeniu nowych kierunków badań również w tym obszarze. W ostatnich latach swojej pracy zawodowej prof. Koronacki zaproponował oryginalne podejście do ważnej klasy trudnych problemów klasyfikacyjnych w dziedzinie maszynowego uczenia się, otwierając drogę do uzyskiwania na świecie nowych wyników m.in. na polu badań genomicznych i proteomicznych.

Jest autorem albo współautorem kilkudziesięciu publikacji, zwykle o zasięgu międzynarodowym, w tym 6 monografii (dwóch o zasięgu międzynarodowym). Zredagował wiele tomów prac zbiorowych o zasięgu międzynarodowym. Był członkiem komitetów programowych dziesiątków konferencji międzynarodowych oraz członkiem komitetów redakcyjnych kilku czasopism o zasięgu międzynarodowym.

Wykładał przez wiele lat w Polsce, głównie na Politechnice Warszawskiej, ale również w USA, Australii i Argentynie. Prowadził prace B+R dla wielu podmiotów, np. dla Ford Motor Co., dla Bank of America oraz Koncernu Orlen S.A.

Był członkiem Rad Naukowych kilku instytutów badawczych (najdłużej Rady Naukowej Instytutu Badań Systemowych PAN, w tym przez dwie kadencje jej przewodniczącym) oraz wielu zespołów NCN i NCBiR. Przez wiele lat był członkiem Komisji Statystyki Komitetu Matematyki PAN. Przez 8 lat był członkiem Europejskiego Komitetu Regionalnego Towarzystwa Statystyki Matematycznej i Prawdopodobieństwa im. Bernoulliego (w kadencji 1998–2000 jego przewodniczącym). Od roku 2002 jest członkiem (Fellow) brytyjskiego Instytutu Matematyki i Jej Zastosowań (Institute of Mathematics and Its Applications).

Roman Zmyślony

Był profesorem (prof. dr hab.) Uniwersytetu Zielonogórskiego. Przez jedną kadencję pełnił funkcję dyrektora Instytutu Matematyki Uniwersytetu Zielonogórskiego. Do 2024 roku pracował w Zakładzie Statystyki Matematycznej i Ekonometrii, którym kierował.

Urodził się w 1946 roku. Studia matematyczne ukończył na Uniwersytecie Wrocławskim w 1969 r. Doktorat obronił w Instytucie Matematycznym PAN w 1974 r. Temat pracy: „Kwadratowe nieobciążone estymatory komponentów wariancyjnych w modelach losowych". Habilitację uzyskał w 1987 r. na podstawie pracy „Nieobciążona estymacja w modelach liniowych". W 2001 r. otrzymał tytuł profesora nauk matematycznych.

Specjalizuje się w statystyce matematycznej. Wypromował pięciu doktorów i około 80 magistrów. Jest autorem ponad 80 publikacji naukowych, w tym wielu prac międzynarodowych. Współpracuje z matematykami z różnych krajów (m.in. Meksyk, Niemcy, USA, Austria, Słowacja). Był współorganizatorem 13 konferencji z zakresu statystyki matematycznej, w tym 5 międzynarodowych. Redaktor naczelny czasopisma Discussiones Mathematicae — Probability and Statistics.

Pracował na Uniwersytecie Zielonogórskim od 1993 r. (wcześniej Instytut Matematyki Politechniki Zielonogórskiej). Brał udział w szeregu grantów naukowych. Za swoją działalność naukową i dydaktyczną otrzymał nagrody, m.in. dwie od Ministra Edukacji Narodowej oraz jedną od Sekretarza Naukowego PAN.

W 2016 r. obchodzono jego jubileusz — 70. urodziny i 44 lata pracy naukowej i dydaktycznej. Jest ceniony za wkład w rozwój statystyki matematycznej w Polsce.

Program konferencji

Poniedziałek, 8 grudnia 2025 r.

Referaty konkursowe

07:30–08:30	**Śniadanie**	
08:45–08:55	**Otwarcie konferencji**	
08:55–09:30	**Roman Zmyślony**, **Jacek Koronacki**	Sesja historyczna
09:30–10:30	**Rafał Kulik**	Non-stationary time series with heavy tails (I)
10:30–10:50	**Przerwa**	
10:50–11:10	**Małgorzata Bogdan**	Statistics and Machine Learning in Astrophysics: Current Work and Future Directions
11:10–11:30	**Filip Wichrowski**, Katarzyna Kaczmarek-Majer	Półnadzorowane niejednorodne ukryte modele Markowa
11:30–11:50	**Katarzyna Szczerba**, Laurent Malisoux, Christophe Ley	AI-informed Non-linear Cox Regression for Time-to-event Analysis
11:50–12:10	Alicja Jokiel-Rokita, **Sylwester Piątek**	Estimation of inequality measures for grouped data
12:10–12:40	**Przerwa**	
12:40–13:00	**Krzysztof Rudaś**, Magdalena Bartczak	Selekcja zmiennych w modelowaniu przyczynowym przy użyciu informacji wzajemnej
13:00–13:20	**Wojciech Niemiro**	Directed information in graphical models of causality
13:20–13:40	**Szymon Jaroszewicz**, Krzysztof Rudaś	Czy modele przyczynowe są niestabilne? Znaczenie randomizacji
14:00–15:00	**Obiad**	
16:00–16:20	Narayanaswamy Balakrishnan, He Yi, **Agnieszka Goroncy**	Sygnatura dyskretnych systemów niezawodnościowych
16:20–16:40	**Krzysztof Jasiński**, Anna Dembińska	Estymatory największej wiarygodności oparte na dyskretnych czasach życia komponentów w systemie k-spośród-n
16:40–17:00	Narayanaswamy Balakrishnan, Maria Kateri, **Magdalena Szymkowiak**	ϕ–Divergence of System Lifetimes

17:00–17:20	**Tomasz Rychlik**, Magdalena Szymkowiak	Warunki zachowania porządków transformacji czasów życia komponentów o jednakowym rozkładzie przez czas życia systemu
17:20–17:50	**Przerwa**	
17:50–18:10	Anna Dudek, **Jakub Pięta**	Bootstrap with random block length for periodically correlated time series
18:10–18:30	Patrice Bertail, Anna E. Dudek, **Karolina Łukasiewicz**	Bootstrap for Single-Index Markov Chains
18:30–18:50	Patrice Bertail, Anna Dudek, **Zuzanna Maciszewska**	Validity of parametric bootstrap in time series models
19:00–20:00	**Kolacja**	
	Zebranie Komisji Statystyki Komitetu Matematyki PAN	

Wtorek, 9 grudnia 2025 r.

Referaty konkursowe

08:00–09:00	**Śniadanie**	
09:30–10:30	**Rafał Kulik**	Non-stationary time series with heavy tails (II)
10:30–10:50	**Przerwa**	
10:50–11:10	**Łukasz Lenart**	Stacjonarność szeregów czasowych o własnościach cyklicznych
11:10–11:30	Anna E. Dudek, Łukasz Lenart, **Bartosz Majewski**	Statistical Properties of Oscillatory Processes with Stochastic Modulation in Amplitude and Time
11:30–11:50	**Cristian Felipe Jiménez-Varón**, Marina I. Knight	A GNAR-Based Framework for Spectral Estimation of Network Time Series: Application to Global Bank Network Connectedness
11:50–12:10	Dominique Dehay, Anna E. Dudek, **Jean-Marc Freyermuth**	Harmonizable VARMA processes: spectrum, simulation, and parameter estimation
12:10–12:40	**Przerwa**	
12:40–13:00	**Filip Pazio**	Własności asymptotyczne estymatora największej wiarogodności w dyskretnych próbach cenzurowanych typu II i układach k-z-n
13:00–13:20	Agnieszka Goroncy, Krzysztof Jasiński, **Jakub Sadowy**	Lifetimes and joint distributions of numbers of components in each state for multi-state discrete k-out-of-n systems
13:20–13:40	**Katarzyna Mamla**, Krzysztof Rudaś	Selekcja zmiennych w modelowaniu różnicowym przy założeniu ograniczonego budżetu
14:00–15:00	**Obiad**	
16:00–16:20	Małgorzata Bogdan, **Adam Przemysław Chojecki**, Ivan Hejný, Bartosz Kołodziejek, Jonas Wallin	PCGLASSO: wykrywanie hubów w modelach grafowych
16:20–16:40	**Iza Danielewska**, Bartosz Kołodziejek	Dyskretne parametryczne modele grafowe

16:40–17:00	Anna Dudek, Antonio Napolitano, **Jakub Rutkowski**, Jakub Wojdyła	Asymptotic normality of Fourier coefficients estimators with unknown lag-dependent cycle frequencies for generalized almost-cyclostationary processes
17:00–17:20	Bogdan Ćmiel, **Bartłomiej Gibas**	The post-hoc detection of dependence
17:20–17:50	**Przerwa**	
17:50–18:30	**Lesław Gajek**, **Marcin Rudź**	Zastosowania twierdzenia Banacha o punkcie stałym do szacowania rozkładu deficytu w momencie ruiny
18:30–18:50	**Agata Boratyńska**	Odporność składki ubezpieczeniowej w bayesowskim modelu ryzyka kolektywnego na zaburzenia niezależności częstości liczby szkód i wartości oczekiwanej ich wielkości
19:30	**Uroczysta kolacja**	

Środa, 10 grudnia 2025 r.

08:00–09:00	**Śniadanie**	
09:30–10:30	**Hernando Ombao**	Overview of Methods for Characterizing Dependence in Multivariate Time Series (I)
10:30–10:50	**Przerwa**	
10:50–11:10	**Konrad Furmańczyk**, Kacper Paczutkowski	Klasyfikacja PU przy naruszeniu założenia SCAR z zastosowaniem klasteryzacji i regresji logistycznej
11:10–11:30	**Paweł Teisseyre**, Jan Mielniczuk	Positive-Unlabeled Regression: learning from partially labeled quantitative outcomes
11:30–11:50	Irène Gijbels, Małgorzata Łazęcka, **Jan Mielniczuk**, Johan Segers	Assessing variability of kernel-based estimators of label shift
11:50–12:10	Jan Mielniczuk, **Wojciech Rejchel**, Paweł Teisseyre	Prior shift estimation for positive unlabeled data
12:10–12:40	**Przerwa**	
12:40–13:00	Xabier Iriarte, Julen Bacaicoa, Jokin Aginaga, Aitor Plaza, **Anna Szczepańska-Alvarez**	D-optymalne rozmieszczenie tensometrów
13:00–13:20	Mateusz John, **Adam Mieldzioc**	Testowanie hipotez dotyczących macierzy kowariancji o strukturze wstęgowej Toeplitza
13:20–13:40	**Wasyl Kowalczuk**	Deformed sphere as a new reference model for the geoid
14:00–15:00	**Obiad**	
15:15–17:00	**Wycieczka**	
17:50–18:10	**Piotr Nowak**	Nieobciążona estymacja wykładniczej niezawodności na podstawie danych lewostronnie cenzurowanych
18:10–18:30	**Augustyn Markiewicz**	Własności BLUE w błędnie wyspecyfikowanym modelu liniowym
18:30–18:50	**Monika Mokrzycka**, Paweł Krajewski	Improved entropy loss estimation of the separable covariance structure under high-dimensionality
19:00–20:00	**Kolacja**	

Czwartek, 11 grudnia 2025 r.

08:00–09:00	**Śniadanie**	
09:30–10:30	**Hernando Ombao**	Overview of Methods for Characterizing Dependence in Multivariate Time Series (II)
10:30–10:50	**Przerwa**	
10:50–11:10	**Grzegorz Wyłupek**	A bootstrap test for the one-sided two-sample right-censored model
11:10–11:30	**Katarzyna Filipiak**	Estimation and sufficiency under the extended growth curve model with random nuisance parameters
11:30–11:50	**Aldo William Medina Garay**, Yessenia A. Gil, Victor H. Lachos	The interval censored regression models under asymmetric distributions
11:50–12:10	**Francyelle de Lima Medina**, Patrice Bertail, Aldo M. Garay	Parameter Estimation and Forecasting for Zero-Inflated Discrete Time Series
12:10–12:40	**Przerwa**	
12:40–13:00	**Przemysław Grzegorzewski**, Yan Sun, Maha Moussainst	O testowaniu jednorodności rozproszenia na podstawie danych przedziałowych
13:00–13:20	**Piotr Szulc**	Czy znamy wynik tegorocznych wyborów prezydenckich?
13:20–13:30	**Zakończenie konferencji**	
13:30	**Obiad**	

Streszczenia referatów

Statistics and Machine Learning in Astrophysics: Current Work and Future Directions

Małgorzata Bogdan

University of Wroclaw

Language of the presentation: English

Gamma-Ray Bursts (GRBs) are among the most powerful electromagnetic explosions in the universe, reaching us from distances corresponding to nearly 13 billion years ago. They provide a rare window into the earliest phases of cosmic evolution. Yet, their potential for cosmological studies is limited, as only a small fraction of GRBs have spectroscopically confirmed redshifts obtained from ground- and space-based observations. To address this, we are developing statistical and machine learning techniques for estimating redshifts from other observable properties and for the early identification of high-redshift GRBs that can be rapidly followed up by satellite telescopes.

In this talk, I will present ongoing work carried out in collaboration with Prof. Maria Dainotti's team and discuss future directions of research in this area.

References

[1] Narendra A., Dainotti M. G., Sarkar M., Lenart A. Ł., Bogdan M., Pollo A., Zhang B., Rabeda A., Petrosian V., Iwasaki K., *Gamma-ray burst redshift estimation using machine learning and the associated web app*, Astrophysics&Astronomy, 698 (2025), A92.

[2] Dainotti M. G., Narendra A., Pollo A., Petrosian V., Bogdan M., Iwasaki K., Prochaska J. X., Rinaldi E., Zhou D., *Gamma-ray bursts as distance indicators by a statistical learning approach*, The Astrophysical Journal Letters, 967 (2024), L30.

[3] Dainotti M. G., Bhardwaj S., Cook C., Ange J., Lamichhane N., Bogdan M., McGee M., Nadolsky P., Sarkar M., Pollo A., Nagataki S., *GRB Redshift Classifier to Follow up High-redshift GRBs Using Supervised Machine Learning*, The Astrophysical Journal Supplement Series, 277 (2025), 31.

Odporność składki ubezpieczeniowej w bayesowskim modelu ryzyka kolektywnego na zaburzenia niezależności częstości liczby szkód i wartości oczekiwanej ich wielkości

Agata Boratyńska

Kolegium Analiz Ekonomicznych, Szkoła Główna Handlowa SGH

Język prezentacji: polski

Jednym z głównych zadań w ubezpieczeniach jest kalkulacja składki, która zależy od zmiennych losowych opisujących liczbę i wartości szkód. Rozważamy problem odporności składki kolektywnej i bayesowskiej przy niepełnej wiedzy a priori dotyczącej zaburzenia niezależności pomiędzy zmiennymi opisującymi częstość i wartość oczekiwaną wielkości szkody. Tradycyjnie w modelach te zmienne są niezależne, ale zastosowania pokazują, że nie musi tak być. Celem jest zbadanie wpływu zależności i wyznaczanie składek optymalnych.

Rozpatrzono dwie klasy rozkładów a priori. W pierwszej wykorzystano kopule FGM, w drugiej opisano pewne zależności pomiędzy dwoma rozkładami typu epsilon-zaburzenie. W obu klasach rozkłady a priori mają formę pewnych liniowych kombinacji dwuwymiarowych rozkładów a priori. Estymowana jest składka indywidualna i do otrzymania składki kolektywnej oraz bayesowskiej wykorzystano kwadratową funkcję straty. Przy obu rodzinach rozkładów a priori wyznaczono oscylację składki kolektywnej i bayesowskiej oraz, jako składki optymalne, wyliczono składki o gamma minimaksowej utracie a priori i a posteriori. Zaprezentowano sytuację, gdzie zależność pomiędzy zmiennymi opisującymi częstość i wartość oczekiwaną wielkości szkody nie ma wpływu na składkę kolektywną.

Wyniki wykorzystano w przykładzie numerycznym, w którym mimo niewielkiej zależności (wyrażonej współczynnikami korelacji) wpływ na składkę bayesowską jak i mnożnik składki w systemie bonus-malus jest istotny.

PCGLASSO: wykrywanie hubów w modelach grafowych

Małgorzata Bogdan[1], Adam Chojecki[2], Ivan Hejný[3], Bartosz Kołodziejek[2], Jonas Wallin[3]

[1]Instytut Matematyczny, Uniwersytet Wrocławski
[2]Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska
[3]Katedra Statystyki, Uniwersytet w Lund

Język prezentacji: polski

W pracy przedstawiono metodę *Partial Correlation Graphical LASSO* (PCGLASSO) – niezależny od skalowania odpowiednik klasycznego GLASSO, w którym karana jest macierz częściowych korelacji zamiast macierzy precyzji. Takie ujęcie niestety prowadzi do niewypukłego zadania optymalizacyjnego, jednak umożliwia dokładniejsze odtwarzanie struktury grafów o zróżnicowanej wariancji wierzchołków, zwłaszcza w sieciach typu *hub* (grafów z nielicznymi krawędziami, lecz wierzchołkami o dużym stopniu).

Opracowano nowy, efektywny algorytm oparty na naprzemiennej optymalizacji względem macierzy skalującej, oraz względem macierzy korelacji, wykorzystujący zmodyfikowany szybki solver GLASSO.

Przeprowadzono analizę teoretyczną, obejmującą warunek reprezentowalności słabszy niż w GLASSO (co sprzyja wykrywaniu struktur typu *hub*) i gwarantujący spójność wyboru modelu. Wyniki symulacji oraz analiza danych genowych potwierdzają, że PCGLASSO skutecznie identyfikuje *huby*, czyli węzły centralne w sieciach.

Literatura

[1] Bogdan M., Chojecki A., Hejný I., Kołodziejek B., Wallin J., *Identifying Network Hubs with the Partial Correlation Graphical LASSO*, arXiv preprint (2025), arXiv:2508.12258.

[2] Friedman J., Hastie T., Tibshirani R., *Sparse inverse covariance estimation with the graphical lasso*, Biostatistics, 9 (2008), 432–441.

[3] Carter J., Lin S., Yuan M., *The Partial Correlation Graphical LASSO*, arXiv preprint (2024), arXiv:2405.01639.

Dyskretne parametryczne modele grafowe

Iza Danielewska, Bartosz Kołodziejek

Politechnika Warszawska

Język prezentacji: polski

Przy modelowaniu złożonych zależności między zmiennymi losowymi pomocne są narzędzia skutecznie łączące strukturę grafową z ujęciem probabilistycznym. W przypadku danych zliczeniowych rolę taką pełnią dyskretne parametryczne modele grafowe, które pozwalają opisywać zależności między licznościami w powiązanych grupach zmiennych.

W dotychczasowej pracy [1] wprowadzono grafowe odpowiedniki rozkładów wielomianowego i ujemnego wielomianowego, opisujące schematy losowania ze zwracaniem. Przedstawiane podejście rozszerza tę ideę, definiując grafowy rozkład hipergeometryczny oraz grafowy ujemny rozkład hipergeometryczny, które modelują procesy losowania bez zwracania. Modele te posiadają globalną własność Markowa dla grafów dekomponowalnych oraz wykazują strukturalną spójność z wcześniejszymi konstrukcjami, co pozwala interpretować je jako część szerszej rodziny rozkładów grafowych.

Proponowane modele mogą znaleźć zastosowanie w analizie złożonych zależności wielowymiarowych w danych dyskretnych, gdzie występują ograniczenia zasobów lub specyficzne zależności między zmiennymi. Jednocześnie ich konstrukcja otwiera możliwość dalszych uogólnień – zarówno w kierunku interpretacji probabilistycznej, jak i zastosowań.

Literatura

[1] Danielewska I., Kołodziejek B., Wesołowski J., Zeng X., *Graphical Negative Multinomial and Multinomial Models with Dirichlet-type priors*, arXiv preprint (2025), arXiv:2301.06058.

Estimation and sufficiency under the extended growth curve model with random nuisance parameters

Katarzyna Filipiak

Poznań University of Technology

Language of the presentation: Polish

An extended growth curve model with fixed and random effects is considered. Under the assumption of multivariate normality, the maximum likelihood estimators of the fixed effects and the dispersion matrix are determined in a model with random nuisance parameters without any assumption on the covariance structure and under the assumption of compound symmetry. Furthermore, when the experiments are designed in balanced complete blocks, particular symmetric matrices appear in the likelihood equations, allowing closed-form expressions for the estimators. Moreover, it is shown that the vector of sufficient statistics for the fixed effects extended growth curve model for each covariance structure remains sufficient for the model with random nuisance parameters.

The presented results are illustrated with a real data example.

References

[1] Filipiak K., Markiewicz A., Krajewski P., Ćwiek-Kupczyńska H., *Estimation and sufficiency under the mixed effects extended growth curve model with compound symmetry covariance structure*, Symmetry, 17 (2025), 1901.

Harmonizable VARMA processes: spectrum, simulation, and parameter estimation

Dominique Dehay[1], **Anna E. Dudek**[2,3], **Jean-Marc Freyermuth**[3]

[1]IRMAR-UMR CNRS 6625, University of Rennes, Rennes, France
[2]Faculty of Applied Mathematics, AGH University of Krakow, Krakow, Poland
[3]Aix Marseille University, CNRS, I2M,Marseille, France

Language of the presentation: English

Harmonizable time series are natural extensions of stationary time series with a spectral decomposition whose components are correlated. Thus, the covariance function of a harmonizable time series is bivariate and admits a two-dimensional Fourier decomposition (Loeve spectrum). They form a broad class of nonstationary processes that has been a subject of investigation for a long time. In this talk, we introduce a parametric form for these harmonizable processes, namely Harmonizable Vector AutoRegressive and Moving Average models (HVARMA). A method is then provided to generate finite time sample realizations of HVARMA with known Loeve spectrum, and finally, a first approach to the parameter estimation problem is discussed.

References

[1] Dehay D., Dudek A. E., Freyermuth J.M., *Spectral characteristics of Harmonizable VARMA processes*, HAL preprint (2025), HAL-04324913v3.

Klasyfikacja PU przy naruszeniu założenia SCAR z zastosowaniem klasteryzacji i regresji logistycznej

Konrad Furmańczyk, Kacper Paczutkowski

Szkoła Główna Gospodarstwa Wiejskiego, Instytut Informatyki Technicznej

Język prezentacji: polski

W referacie zostanie przedstawiony algorytmu korekty etykiet w klasyfikacji typu Positive–Unlabeled (PU) oparty na klastrowaniu (por. [3]). Rozważane są przypadki, w których warunek Selected Completely At Random (SCAR) jest spełniony, jak również sytuacje, w których założenie to zostaje naruszone. W pierwszym kroku proponowanego algorytmu korekta etykiet uzyskana jest za pomocą klasteryzacji metodą 2-means. Następnie wykonywana jest regresja logistyczna na oczyszczonym zbiorze danych, w której etykiety pozytywne przypisywane są obserwacjom oznaczonym jako pozytywne przez algorytm w pierwszym kroku oraz rzeczywistym obserwacjom pozytywnym, natomiast pozostałe obserwacje otrzymują etykietę negatywną. Skuteczność metody oceniono na podstawie 11 rzeczywistych zbiorów danych pochodzących z repozytoriów uczenia maszynowego oraz jednego zbioru syntetycznego. Uzyskane wyniki potwierdzają efektywność zaproponowanego algorytmu w przypadkach, gdy warunek SCAR jest naruszony, oraz wskazują na umiarkowaną odporność metody LassoJoint (poz. [1]-[2]) na takie naruszenia.

Literatura

[1] Furmańczyk K., Dudziński M., Dziewa-Dawidczyk D., *Some proposal of the high dimensional PU learning classification procedure*, Computational Sciences-ICCS 2021, Lecture Notes In Computer Science 12744, Springer, 2021, ss. 18-25.

[2] Furmańczyk K., Paczutkowski K., Dudziński M., Dziewa-Dawidczyk D., *Classification methods based on fitting logistic regression to positie and unlabeled data*, Computational Sciences-ICCS 2022, Lecture Notes In Computer Science 13350, Springer, 2022, ss. 31-45.

[3] Furmańczyk K., Paczutkowski K., *A proposal for PU classification under Non-SCAR using clustering and logistic model*, USB Proceedings of the 22nd International Conference on Modeling Decisions for Artificial Intelligence: MDAI 2025, Valencia, Spain 15 - 18 September, 2025 / Torra Vicenç, Narukawa Yasuo, Domingo-Ferrer Josep (red.), 2025, ss.117-128.

Zastosowania twierdzenia Banacha o punkcie stałym do szacowania rozkładu deficytu w momencie ruiny

Lesław Gajek, Marcin Rudź

Instytut Matematyki Politechniki Łódzkiej

Język prezentacji: polski

Zaprezentowane zostaną nowe zastosowania twierdzenia Banacha o punkcie stałym do rozwiązywania pewnych problemów matematyki ubezpieczeniowej. W zadanej przestrzeni probabilistycznej skonstruowane zostaną monotoniczne ciągi dolnych i górnych oszacowań rozkładu deficytu w momencie ruiny $\Psi(u,\, y)$ rozważanego jako funkcja początkowej nadwyżki u ubezpieczyciela i wysokości deficytu y w chwili ruiny. Bezpośrednie zastosowanie twierdzenia Banacha nie jest możliwe, gdyż operatory ryzyka rozważane dotychczas w literaturze mają zwykle nieskończenie wiele punktów stałych. Nasuwa się zatem pytanie, czy istnieje przestrzeń, w której są one kontrakcją. Podamy rozwiązanie tego problemu, rozważając monotoniczny operator ryzyka L w odpowiednio zdefiniowanej metrycznej przestrzeni zupełnej $\langle \mathcal{R}, d_r \rangle$, w której $\Psi(u,\, y)$ jest jedynym punktem stałym. Pokażemy ponadto, że L jest kontrakcją na $\langle \mathcal{R}, d_r \rangle$. W konsekwencji możliwe będzie efektywne szacowanie $\Psi(u,\, y)$ (kontrolując przy tym błąd aproksymacji), iterując L na dowolnej funkcji z $\langle \mathcal{R}, d_r \rangle$. Zaprezentowana metodologia może być stosowana w modelach ryzyka niewypłacalności z czasem dyskretnym jak i ciągłym.

The interval censored regression models under asymmetric distributions

Aldo M. Garay[1,2], **Yessenia A. Gil**[1], **Victor H. Lachos**[3]

[1]Federal University of Pernambuco, UFPE - Brazil
[2] MODAL'X UMR 9023 - UPL, University Paris Nanterre, France
[3]University of Connecticut - UConn - USA

Language of the presentation: English

This paper proposes an interval-censored linear regression model based on the class of asymmetric heavy-tailed distributions, denoted by scale mixtures of skew-normal distributions (SMSN), providing a robust alternative to the usual Gaussian assumption in censored regression models. A novel Expectation/Conditional Maximization algorithm is proposed for maximum likelihood estimation, with analytical expressions at the E-step, as opposed to Monte Carlo simulations. Our methodology is illustrated through intensive simulations and analysis of a real data set from the Household Survey OHS99 conducted by Statistics South Africa.

References

[1] Dempster A., Laird N., Rubin D., *Maximum likelihood from incomplete data via the EM algorithm*, Journal of the Royal Statistical Society, Series B (Statistical Methodology), 39 (1977), 1-38.

[2] Garay A.M., Lachos V.H., Bolfarine H., Cabral C.R.B., *Linear censored regression models with scale mixtures of normal distributions*, Statistical Papers, 58 (2017), 247-278.

[3] Gil Y.A., Garay A.M., Lachos V.H., *Likelihood-based inference for interval censored regression models under heavy-tailed distributions*, Statistical Methods & Applications, 34 (2025), 519-544.

[4] Lachos V.H., Garay A.M., Cabral C.R.B., *Moments of truncated skew-normal/independent distributions*, Brazilian Journal of Probability and Statistics, 34 (2020), 478-494.

The post-hoc detection of dependence

Bogdan Ćmiel[1], **Bartłomiej Gibas**[2]

[1]AGH University of Krakow, Faculty of Applied Mathematics
[2]AGH University of Krakow, Faculty of Computer Science, Electronics and Telecommunications

Language of the presentation: Polish

The concept of independence plays a crucial role in probability theory and has been the subject of extensive research in recent years. Numerous approaches have been proposed to validate this dependency, but most of them address the problem only at a global level. From a practical perspective, it is important not only to determine whether the data are dependent, but also to identify where this dependence occurs and how strong it is. The graphical presentation of results is another essential aspect that should not be neglected, as it considerably enhances interpretability.

The main objective of this work is to propose a solution that considers both of these aspects. Relaying on copula-based results presented in [1], we introduce a novel method for testing statistical independence using the quantile dependence function. Rather than assessing whether the value of the test statistic exceeds a single critical threshold and subsequently deciding whether to reject the independence hypothesis, we use a so-called critical surfaces that guarantee locally equal probability of exceeding it under independence. This approach enables a detailed examination of local discrepancies and an assessment of their statistical significance while preserving the overall significance level of the test.

In this paper, we derive the theoretical foundations of the method and provide a proof of the test's consistency using one of the Berry–Esseen type bound established in [3]. Furthermore, we define these critical surfaces, propose an effective test's implementation procedure, and demonstrate how the method performs on several illustrative examples. Finally, we compare the empirical power of the proposed test with that of several existing approaches.

References

[1] Ćmiel B., Ledwina T., *Detecting dependence structure: visualization and inference*, arXiv preprint (2025), arXiv:2410.05858.

[2] Hájek J., Šidák Z., Sen P.K., *Theory of Rank Tests*, Academic Press, 1999.

[3] Lahiri S.N., Chatterjee A., Maiti T., *Normal approximation to the hypergeometric distribution in nonstandard cases and a sub-Gaussian Berry–Esseen theorem*, Journal of Statistical Planning and Inference, 137 (2007), 3570-3590.

[4] Ledwina T. *Visualizing association structure in bivariate copulas using new dependence function*, Springer Proceedings in Mathematics & Statistics, 122 (2015), 19-27.

[5] van der Vaart A.W., *Asymptotic Statistics*, Cambridge University Press, 1998.

Sygnatura dyskretnych systemów niezawodnościowych

Narayanaswamy Balakrishnan[1], **He Yi**[2], **Agnieszka Goroncy**[3]

[1]McMaster University, Kanada
[2]Beijing University of Chemical Technology, Chiny
[3]Uniwersytet Mikołaja Kopernika w Toruniu

Język prezentacji: polski

W badaniach systemów koherentnych w teorii niezawodności pojęcie sygnatury, wprowadzone pierwotnie przez Samaniego [1], stało się kluczowym narzędziem w analizie właściwości systemów niezawodnościowych, ocenie ich strukturalnej efektywności oraz w opracowywaniu metod wnioskowania opartych na danych dotyczących czasu życia systemu.

W oparciu o koncepcję sygnatury zaproponowano w literaturze kilka jej rozszerzeń, takich jak sygnatura przetrwania, sygnatura rozszerzona, sygnatura łączona czy sygnatura uporządkowana. Dotychczasowe badania koncentrowały się głównie na systemach koherentnych złożonych z komponentów o ciągłym czasie życia. W praktyce jednak systemy tego typu często składają się z elementów o czasie życia dyskretnym, na przykład gdy monitorowanie odbywa się jedynie w dyskretnych momentach czasu lub gdy systemy działają cyklicznie, a obserwacje dotyczą liczby cykli zakończonych sukcesem przed wystąpieniem awarii.

W pracy [2] rozważamy systemy koherentne złożone z niezależnych komponentów o dyskretnym czasie życia i wprowadzamy pojęcie sygnatury w czasie dyskretnym. Omawiamy jej podstawowe własności, przedstawiamy wyniki dotyczące uporządkowania stochastycznego oraz formuły przekształceniowe umożliwiające porównywanie sygnatur systemów o różnej liczbie komponentów. Dodatkowo prezentujemy przykłady ilustrujące uzyskane wyniki teoretyczne.

Literatura

[1] Samaniego F.J., *On closure of the IFR class under formation of coherent systems*, IEEE Transactions on Reliability, 34 (1985), 69–72.

[2] Balakrishnan N., Yi H., Goroncy A., *On the notion of discrete-time signature and some associated properties and results*, submitted (2025).

O testowaniu jednorodności rozproszenia na podstawie danych przedziałowych

Przemysław Grzegorzewski[1], **Yan Sun**[2], **Maha Moussainst**[2]

[1] Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska
[2]Department of Mathematics & Statistics, Utah State University, USA

Język prezentacji: polski

Niektóre procedury statystyczne, jak ANOVA, zakładają jednorodność rozproszenia rozkładów, z których pochodzą próbki. Jeśli analizujemy dane z rozkładów normalnych, postulowaną równość wariancji można weryfikować np. za pomocą testu Levene'a. W tym celu można też posłużyć się testem Browna-Forsythe'a, który jest bardziej odporny od testu Leven'a na odstępstwa od normalności rozkładów.

Celem niniejszego referatu jest zaprezentowanie konstrukcji testu do weryfikacji hipotezy o jednorodności rozproszenia na podstawie danych przedziałowych. Tego typu dane pojawiają się przy modelowaniu nieprecyzyjnych pomiarów, podczas analizy fluktuacji mierzonych wielkości, czy też w sytuacjach, w których dokładne wartości liczbowe są zastępowane przedziałami w celu ochrony prywatności.

Główne wyzwanie przy opracowywaniu testów statystycznych dla tego typu danych wynika z faktu, że tradycyjne pojęcie rozkładu prawdopodobieństwa (np. normalność) nie daje się w naturalny sposób rozszerzyć na przedziały losowe. Stąd też większość istniejących testów dla danych przedziałowych opiera się na metodach bootstrapowych lub technikach permutacyjnych (por. np. [1], [2]).

W referacie pokażemy, w jaki sposób można wyprowadzić rozkład graniczny statystyki testowej dla przedziałów losowych przy dość naturalnych założeniach i niezbyt ograniczających warunkach regularności (por. [3]). Badania symulacyjne wskazują na wysoką skuteczność zaproponowanego testu nawet w przypadku próbek o niezbyt dużej liczności.

Literatura

[1] Grzegorzewski P., *Permutation k-sample goodness-of-fit test for fuzzy data*, Proceedings of the 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2020, pp. 1-8.

[2] Ramos-Guajardo A.B., Lubiano M.A., *K-sample tests for equality of variances of random fuzzy sets*, Computational Statistics and Data Analysis, 56 (2012), 956-966.

[3] Sun Y., Rios Z., Bean B., *Multi-sample means comparisons for imprecise interval data*, International Journal of Approximate Reasoning, 176 (2025), 109322.

Czy modele przyczynowe są niestabilne? Znaczenie randomizacji

Szymon Jaroszewicz, Krzysztof Rudaś

Instytut Podstaw Informatyki PAN
Politechnika Warszawska

Język prezentacji: polski

W pracy zajmujemy się oceną modeli predykcji różnicowej (uplift modeling), służącej do szacowania przyczynowych efektów danego działania (terapii medycznej, akcji marketingowej) na podstawie indywidualnych cech danej osoby. Aby zapewnić przyczynową interpretację modeli, dane uczące pochodzą zwykle z prób randomizowanych.

Chociaż stosowanie modeli różnicowych jest wskazane w wielu dziedzinach a ich popularność rośnie, wielu autorów [1, 2, 3] zauważa ich 'niestabilne' zachowanie: bardzo duże zmiany modelu i miar jego jakości przy niewielkich zmianach parametrów czy metod oceny.

W literaturze proponowane są różne wyjaśnienia tego zjawiska. W niniejszej pracy wskazujemy czynnik, który chociaż był do tej pory ignorowany, jest naszym zdaniem kluczowy. Czynnikiem tym jest jakość randomizacji.

Pokazujemy, że w przypadku wielu publicznie dostępnych zbiorów danych, przypisanie terapii nie jest całkowicie losowe, wbrew deklaracjom ich autorów. Pokazujemy, że niestabilne zachowanie modeli różnicowych występuje właśnie na tych zbiorach. Co więcej, nawet niewielkie zaburzenia randomizacji mogą znacząco wpływać na modele i miary ich oceny.

W pracy przedstawiamy również szereg narzędzi diagnostycznych służących do oceny wpływu ewentualnych wad randomizacji na zbudowany model. Proponujemy również metody oparte o wskaźnik skłonności (*propensity score*) pozwalające, w niektórych sytuacjach, na poprawną ocenę modelu mimo braku pełnej randomizacji. Wprowadzamy również dwie nowe metody estymacji wskaźnika skłonności lepiej dostosowane do oceny modeli różnicowych.

Literatura

[1] Verbeken B., Guerry M., Verbeke W., Verboven S., *Uplift model evaluation with ordinal dominance graphs*, Journal of Machine Learning Research, 26(90), 2025, 1-56.

[2] Nyberg O., Klami A., *Quantifying uncertainty of uplift: Trees and t-learners*, Neurocomputing, vol. 590, 127741, 2024.

[3] Rößler J., Tilly R., Schoder D., *To treat, or not to treat: Reducing volatility in uplift modeling through weighted ensembles*, Proc. of the 54th Hawaii International Conference on System Sciences, 2021, p. 1601–1610.

Estymatory największej wiarygodności oparte na dyskretnych czasach życia komponentów w systemie k-spośród-n

Krzysztof Jasiński[1], **Anna Dembińska**[2]

[1]Wydział Matematyki i Informatyki, Uniwersytet Mikołaja Kopernika w Toruniu
[2]Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska

Język prezentacji: polski

Referat opiera się na pracach [1], [2]. Prezentowane wyniki dotyczą estymacji metodą największej wiarygodności nieznanego parametru dyskretnego rozkładu czasów życia komponentów w binarnym systemie k-spośród-n, zakładając, że te czasy życia są IID zmiennymi losowymi. W estymacji uwzględniamy liczbę popsutych komponentów w systemie. Rejestrujemy ją aż do momentu (włącznie) jego awarii. Kluczowe znaczenie ma tutaj łączny rozkład tej liczby i czasów awarii odpowiadających jej komponentów. Najpierw omówimy przypadek ogólny, czyli dowolnego rozkładu dyskretnego, który spełnia określone łagodne warunki regularności. Następnie przeanalizujemy typowe rozkłady dyskretne, takie jak rozkład Poissona, dwumianowy, dwumianowy ujemny oraz geometryczny.

Literatura

[1] Dembińska A., Jasiński K., *Maximum likelihood estimators based on discrete component lifetimes of a k-out-of-n system*, TEST 30 (2021), 407-428.

[2] Dembińska A., Jasiński K., *Likelihood inference for geometric lifetimes of components of k-out-of-n systems*, Journal of Computational and Applied Mathematics 435 (2024), nr art. 115267.

A GNAR-Based Framework for Spectral Estimation of Network Time Series: Application to Global Bank Network Connectedness

Cristian F. Jiménez-Varón, Marina I. Knight

University of York, UK

Language of the presentation: English

Understanding how dependencies evolve across time and network structures is key to analyzing complex financial systems such as global banking networks. In this talk, I will present a novel spectral analysis framework for Generalized Network Autoregressive (GNAR) processes, which extends traditional models by incorporating multi-stage neighbourhood effects. This approach allows us to study frequency-specific interactions between network nodes, providing deeper insights into dynamic connectivity patterns. I will introduce the definition of the GNAR spectral density, coherence, and partial coherence, together with parametric and network-penalized nonparametric estimators. Simulation studies and an empirical analysis of global bank connectedness demonstrate that the GNAR spectral approach captures both temporal and structural features of volatility transmission across financial networks.

References

[1] Jiménez-Varón C. F., Knight M. I., *A GNAR-Based Framework for Spectral Estimation of Network Time Series: Application to Global Bank Network Connectedness*, arXiv preprint (2025), arXiv:2510.06157.

Deformed sphere as a new reference model for the geoid

Wasyl Kowalczuk

Department for Theory of Continuous Media and Nanostructures
Institute of Fundamental Technological Research, Polish Academy of Sciences

Language of the presentation: English

In our research we concentrate on some applications of differential geometry to geodesy and cartography. In particular, we are interested in development of a new reference model for the geoid (originally proposed by Faridi and Schucking back in 1987 in their paper [1]) that turns out to be a valid alternative to the classical ellipsoidal reference model (used, e.g., in World Geodetic System 1984). The considered new reference model is based on the concept of deformed spheres with the parameter of deformation $\lambda < 1/3$ (where $\lambda = 0$ corresponds to the usual, non-deformed sphere).

In our paper [2] we have compared the main geometric characteristics of the deformed spheres and the standard spheroids (ellipsoids of revolution), such as semi-major (equatorial) axis a and semi-minor (polar) axis b, quarter-meridian length m_P, surface area S, volume V, sphericity index Ψ, and tipping (bifurcation) point $T^{1/2}$ for geodesics. Using the minimization process of the RMS error for individual pairs of the discussed geometric characteristics, we obtain a proposition of the optimized value of the deformation parameter $\lambda_{\mathrm{RMS}} \approx 0.003\ 349\ 672\ \ldots$ for the deformed sphere's reference model of the geoid.

Quite interestingly, it turns out that the value of the standard spheroid's (Earth's) flattening factor $f = 1 - b/a \approx 0.003\ 352\ 811\ \ldots$ is quite a good approximation for the obtained optimized value of the deformation parameter λ_{RMS} with the relative error $\delta f = (f - \lambda_{\mathrm{RMS}})/f \approx 9.362\ \ldots \times 10^{-4}$.

In papers [3, 4] we have also found the explicit solutions of the direct and inverse problems for loxodromes (rhumb lines) and orthodromes (geodesics) for navigation purposes on the deformed spheres and compared the obtained results for the selected important air routes on the Earth's surface with the corresponding results obtained from the standard ellipsoidal reference model WGS 84. It has turned out that there is a good agreement in the predictions of both reference models wit relative errors of the order 10^{-6}–10^{-7} in distances and reversed azimuths and of the order 10^{-5}–10^{-7} in azimuths.

References

[1] Faridi A., Schucking E., *Geodesics and deformed spheres*, Proc. Am. Math. Soc., 100 (1987), 522–525.

[2] Kowalczuk W., Mladenov I., *Comparison of main geometric characteristics of deformed sphere and standard spheroid*, Bulletin of the Polish Academy of Sciences: Technical Sciences, 71 (2023), e147058, 1–9.

[3] Kowalczuk W., Mladenov I., *Explicit solutions for geodetic problems on the deformed sphere as a reference model for the geoid*, Geom. Integr. Quant., 25 (2023), 73–94.

[4] Kowalczuk W., Mladenov I., *λ-spheres as a new reference model for geoid: explicit solutions of the direct and inverse problems for loxodromes (rhumb lines)*, Mathematics, 10 (2022), 3356-1–10.

Stacjonarność szeregów czasowych o własnościach cyklicznych

Łukasz Lenart

Uniwersytet Ekonomiczny w Krakowie

Język prezentacji: polski

W badaniach rozważono klasę szeregów czasowych $\{y_t : t \in \mathbb{Z}\}$ o własnościach cyklicznych definiowanych przez reprezentację (dla uproszczenia z jedną częstotliwością $\lambda \in (0, \pi)$)

$$y_t = \alpha_t \sin(\lambda t) + \beta_t \cos(\lambda t),$$

gdzie $\{(\alpha_t, \beta_t) : t \in \mathbb{Z}\}$ stanowi dwuwymiarowy szereg czasowy. Równoważnie, można zapisać

$$y_t = A_t \sin\left(\lambda(t + P_t)\right),$$

gdzie A_t jest procesem amplitudy, a P_t – procesem przesunięcia fazowego. Celem pracy jest sformułowanie warunków gwarantujących stacjonarność szeregu czasowego $\{y_t : t \in \mathbb{Z}\}$.

W literaturze analizowany jest głównie przypadek, gdy (α_t, β_t) jest stacjonarnym szeregiem czasowym gaussowskim o zerowej średniej (por. [1], [2], [3]), co zapewnia jednocześnie stacjonarność i gaussowskość procesu $\{y_t : t \in \mathbb{Z}\}$. Przypadek ten jednak wprowadza ścisły związek pomiędzy pierwszymi dwoma momentami procesu amplitudy $A_t = \sqrt{\alpha_t^2 + \beta_t^2}$, co istotnie ogranicza jego zastosowania (np. w analizie sygnałów EKG, plam na słońcu czy cykliczności koniunkturalnej).

W pracy zostaną sformułowane warunki konieczne i wystarczające stacjonarności (w sensie węższym i szerszym) szeregu czasowego $\{y_t\}$. W szczególności rozważone zostaną przypadki:

a) $(\alpha_t, \beta_t) = \sum_{k=0}^{\infty} \psi_k(\epsilon_{t-k}, \epsilon'_{t-k})$, gdzie $\sum_{k=0}^{\infty} |\psi_k| < \infty$, a $(\epsilon_t, \epsilon'_t)$ jest dwuwymiarowym białym szumem o rozkładzie sferycznym,

b) proces amplitudy A_t jest stacjonarny i niezależny od procesu przesunięcia fazowego P_t, który jest procesem błądzenia przypadkowego z szumem gaussowskim.

Dodatkowo przeanalizowana zostanie własność α-mieszania w przypadku (b).

Literatura

[1] Hannan E., *The estimation of a changing seasonal pattern*, Journal of the American Statistical Association, 1964, 59(308), 1063–1077.

[2] Proietti T., Maddanu F., *Modelling cycles in climate series: The fractional sinusoidal waveform process*, Journal of Econometrics, Elsevier, ,2024, 239(1).

[3] Trimbur T. M., *Properties of higher order stochastic cycles*, Journal of Time Series Analysis, 2006, 27(1), 1–17.

Bootstrap for Single-Index Markov Chains

Patrice Bertail[1], **Anna E. Dudek**[2,3], **Karolina Łukasiewicz**[1,3]

[1]MODAL'X, Université Paris-Nanterre, Paris, France
[2]Aix Marseille University, CNRS, I2M, Marseille, France
[3]AGH University of Krakow, Krakow, Poland

Language of the presentation: English

Single-index models are widely used in many models because they combine the flexibility of nonparametric methods with the efficiency of parametric ones, helping to overcome the curse of dimensionality. We introduce the Single-Index Approximative Regenerative Block Bootstrap (SARBB), a new bootstrap algorithm designed for single-index models of Markov processes of the form

$$X_{t+1} = g(\beta' X_t) + \epsilon_{t+1}.$$

We propose a consistent estimation method for both the parametric and nonparametric parts of the single-index Markov chain of the above model. We propose a reduction algorithm that transforms the original high-dimensional process into a one-dimensional Markov chain. After the reduction step, we apply the Nummelin splitting technique together with the bootstrap algorithm developed for one-dimensional setting (see [1]). We present theoretical results on the consistency of the SARBB method and the convergence rate of the estimator for the nonparametric component. Finally, we illustrate the application of our algorithm with a simulation example.

References

[1] Bertail P., Clémençon S., *Regeneration-based statistics for Harris recurrent Markov chains*, Dependence in Probability and Statistics. Lecture Notes in Statistics, 187 (2006).

[2] Bertail P., Dudek A. E., Łukasiewicz K., *Bootstrap for Single-Index Markov Chains.*, working paper.

Validity of parametric bootstrap in time series models

Patrice Bertail[2], **Anna Dudek**[1,3], **Zuzanna Maciszewska**[1,2]

[1]AGH University of Krakow, Poland
[2]Université Paris Nanterre, France
[3]Aix-Marseille University, CNRS, I2M, France

Language of the presentation: English

Beran in [1] established necessary and sufficient conditions for the validity of the bootstrap in certain regular models (in the Hájek–Le Cam sense) for locally regular statistics. Although these results were originally formulated for the i.i.d. case, they remain applicable to dependent data under suitable regularity assumptions. The theoretical framework adopted here originates from Le Cam (see [3]), who introduced the notion of contiguity and the class of Locally Asymptotically Normal (LAN) experiments to study the asymptotic properties of parametric models in a unified and elegant way. Within this framework, asymptotic inference can be developed in terms of local experiments that approximate the original statistical models.

It turns out that, for models possessing the LAN property, the local asymptotic equivariance of an estimator is equivalent to the convergence of the intuitive bootstrap distributions to the correct limit distribution. As noted in [2], many time series models naturally fall within this framework, which allows one to employ these asymptotic arguments for establishing the validity and consistency of bootstrap estimators in parametric time series settings.

In this work, we extend Beran's results to dependent data and propose a generalization of Theorem 2.1 from [1] to models satisfying the Locally Asymptotically Mixed Normal (LAMN) property. This broader framework is particularly relevant for nonstationary time series models, many of which exhibit the LAMN structure, and therefore provides a theoretical foundation for understanding bootstrap procedures in such contexts.

References

[1] Beran R., *Diagnosing Bootstrap Success*, Annals of the Institute of Statistical Mathematics 49, 1997, pp. 1–24.

[2] Jeganathan P., *Some Aspects of Asymptotic Theory with Applications to Time Series Models*, Econometric Theory 11.5, 1995, pp. 818–887.

[3] Le Cam L. M., *Locally asymptotically normal families of distributions*, University of California Press, 1960, pp. 37-98.

Statistical Properties of Oscillatory Processes with Stochastic Modulation in Amplitude and Time

Anna E. Dudek[1], **Łukasz Lenart**[2], **Bartosz Majewski**[1]

[1]AGH University of Krakow
[2]Krakow University of Economics

Language of the presentation: English

Analysis of cyclical data is a key component of statistical analysis of random processes. However, many real-world signals exhibit irregular cyclic patterns, which pose challenges for standard modeling approaches. To address this, we introduce a semiparametric continuous-time model for signals with irregular cyclicities. This model is based on stochastic modulation of the amplitude and phase of a deterministic signal (an almost periodic function). We investigate its theoretical properties, focusing on the behavior of the mean and autocovariance functions, and we demonstrate that the process gradually loses its cyclic structure over time due to random disturbances. Estimators of the asymptotic mean and autocovariance functions are introduced. The performance of the autocovariance function estimator is examined in a simulation study.

References

[1] Dudek A. E., Lenart Ł., Majewski B., *Statistical properties of oscillatory processes with stochastic modulation in amplitude and time*, Journal of Nonparametric Statistics, 37 (2025), 1-21, doi: 10.1080/10485252.2025.2556414.

Selekcja zmiennych w modelowaniu różnicowym przy założeniu ograniczonego budżetu

Katarzyna Mamla[1], **Krzysztof Rudaś**[1,2]

[1]Politechnika Warszawska
[2]Instytut Podstaw Informatyki, PAN

Język prezentacji: polski

Modelowanie różnicowe zajmuje się przewidywaniem efektu działania, takiego jak terapia medyczna czy kampania marketingowa, na poziomie pojedynczych obserwacji. Budowa modeli różnicowych wymaga wykorzystania dwóch zbiorów uczących: grupy eksperymentalnej, poddanej działaniu, oraz grupy kontrolnej. Wynikowy model umożliwia oszacowanie różnicy między prawdopodobieństwami klas w obu grupach. Kluczowym zagadnieniem w modelowaniu różnicowym jest określenie, które zmienne istotnie wpływają na poprawne oszacowanie efektu działania. Dotychczasowe metody selekcji cech zakładają jednak, że wszystkie zmienne mają jednakowy koszt pozyskania, co w praktycznych zastosowaniach może stanowić zbyt duże uproszczenie. W klasycznym modelowaniu predykcyjnym rozważano już podejścia, które uwzględniają różne koszty cech, prowadząc do zadania selekcji zmiennych z ograniczonym budżetem [1].

W prezentacji omówimy problem selekcji zmiennych z ograniczonym budżetem dla regresji logistycznej w kontekście modelowania różnicowego. Wykorzystamy trzy modele różnicowe: model podwójny (ang. *Double Model*), model CVT (ang. *Class Variable Transformation*) oraz model, bazujący na metodzie opisanej w pracy [2]. Przedstawimy modyfikacje istniejących metod selekcji kosztowej z klasycznego zadania predykcyjnego, bazujących na regularyzacji Lasso i adaptacyjnym Lasso, które dostosowaliśmy do specyfiki modeli różnicowych. Eksperymenty przeprowadzone na syntetycznych i rzeczywistych zbiorach danych wskazują, że uwzględnienie kosztów zmiennych w procesie selekcji pozwala uzyskać modele o lepszej jakości predykcyjnej przy ograniczonych zasobach.

Literatura

[1] Klonecki T., Teisseyre P., *Feature selection under budget constraint in medical applications: analysis of penalized empirical risk minimization methods*, Applied Intelligence, 53 (2023), 29943-29973.

[2] Tian L., Alizadeh A.A., Gentles A.J., Tibshirani R., *A simple method for estimating interactions between a treatment and a large number of covariates*, Journal of the American Statistical Association, 109 (2014), 1517-1532.

[3] Rudaś K., Jaroszewicz S., *Regularization for uplift regression*, w: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD'23), Springer, 2023, ss. 593-608.

Własności BLUE w błędnie wyspecyfikowanym modelu liniowym

Augustyn Markiewicz

Uniwersytet Przyrodniczy w Poznaniu

Język prezentacji: polski

Błędna specyfikacja kowariancji błędu w modelach liniowych zazwyczaj prowadzi do niepoprawnego wnioskowania statystycznego. Rozważamy dwa modele liniowe, $\mathcal{A}$ i $\mathcal{B}$, z tą samą macierzą układu, ale różnymi macierzami kowariancji błędu. Warunki, w których każda reprezentacja najlepszego liniowego nieobciążonego estymatora (BLUE) dowolnego estymowalnego wektora parametrycznego dla modelu $\mathcal{A}$, pozostaje BLUE dla modelu $\mathcal{B}$, są dobrze znane od czasu publikacji artykułu C.R. Rao z 1971 roku: Unified theory of linear estimation, Sankhya Ser. A, tom 33, s. 371–394. Jednakże, nie ma wcześniej opublikowanych wyników dotyczących zachowania ważonej sumy kwadratów błędów (SSE) gdy macierz układu lub macierz kowariancji błędu są niepełnego rzędu. W pracy podajemy warunki konieczne i dostateczne na to aby dla założonej kowariancji błędu zarówno BLUE, jak i jego macierz kowariancji pozostała niezmieniona.

Parameter Estimation and Forecasting for Zero-Inflated Discrete Time Series

Francyelle L. Medina[1,2], **Patrice Bertail**[1], **Aldo M. Garay**[1,2]

[1]MODAL'X UMR 9023 - UPL, University Paris Nanterre, France
[2]Federal University of Pernambuco – Recife – Brazil

Language of the presentation: English

In this work, we proposed a class of p-order non-negative integer-valued autoregressive (INAR(p)) process with innovations following zero-inflated (ZI) distributions, referred to as ZI-INAR(p) process. We propose some parameter estimation procedures for the model. We also developed a bootstrap method to construct confidence intervals for the parameters and to estimate the forecasting distributions of future values. The performance of the proposed methods is evaluated through simulation studies and analysis of original datasets.

Keywords: Discrete time series; ZI-INAR(p) process; Zero-inflated distribution.

References

[1] Bertail P., Clémençon S., *Regenerative block bootstrap for markov chains*, Bernoulli, 12 (2006), 689–712.

[2] Bertail P., Clémençon S., Tressou J., *Bootstrapping robust statistics for markovian data applications to regenerative r-statistics and l-statistics*, Journal of Time Series Analysis, 36 (2015), 462-480.

[3] Garay A.M., Medina F.L., Cabral C.R.B., Lin T.I., *Bayesian analysis of the p-order integer-valued ar process with zero-inflated poisson innovations*, Journal of Statistical Computation and Simulation, 90 (2020), 1943-1964.

[4] Garay A.M., Medina F.L., Jales I.C.S., Bertail P., *First-order integer valued ar processes with zero-inflated innovations*, In: Chaari F, Leskow J, Wylomanska A, Zimroz R, Napolitano A. (editors), Nonstationary systems: theory and applications, Springer, 2022, ss. 19-40.

[5] Bertail P., Garay A.M., Medina F.L., Jales I.C.S., *A maximum likelihood and regenerative bootstrap approach for estimation and forecasting of INAR(p) processes with zero-inflated innovations*, Statistics, 58 (2024), 336-363.

Testowanie hipotez dotyczących macierzy kowariancji o strukturze wstęgowej Toeplitza

Mateusz John, Adam Mieldzioc

Politechnika Poznańska

Język prezentacji: polski

Macierze kowariancji o strukturze liniowej mają szerokie zastosowania w wielu dziedzinach nauki i techniki. Przykładem takich struktur są wstęgowe macierze Toeplitza. Identyfikacja odpowiedniej struktury jest kluczowa do wyznaczenia odpowiedniej oceny macierzy kowariancji.

Klasyczne estymatory macierzy kowariancji często nie spełniają wymagań warunków dotyczących posiadania właściwej struktury, dodatniej określoności lub dobrego uwarunkowania. Dlatego w pracy zastosowano estymatory największej wiarogodności z nałożonymi ograniczeniami oraz estymatory typu *shrinkage* (por. [1]), przy założeniu, że dane pochodzą z wielowymiarowego rozkładu normalnego.

Dlatego porównano estymatory największej wiarogodności z nałożonymi ograniczeniami z estymatorami typu *shrinkage* w kontekście testowania hipotez dotyczących struktury macierzy kowariancji zarówno w przypadku testu ilorazu wiarogodności (LRT, ang. *Likelihood Ratio Test*) oraz testu wynikowego Rao (RST, ang. *Rao Score Test*) (por. [2]).

W przypadku struktur liniowych, które nie posiadają własności kwadratowości wyznaczenie estymatorów największej wiarogodności z nałożonymi ograniczeniami stanowi problem numerycznie złożony i czasochłonny. Dlatego celem pracy było pokazanie, że estymatory typu *shrinkage* mają równe dobre własności jak estymatory największej wiarogodności z nałożonymi ograniczeniami i są szybsze do wyznaczenia. Dzięki czemu procedura testowania staje się bardziej wydajna oraz stabilna numerycznie.

Literatura

[1] Filipiak K., Markiewicz A., Mieldzioc A., Mrowińska M., *On estimation of a partitioned covariance matrix with linearly structured blocks*, Statistical Papers, 66 (2025), 98.

[2] Rao C.R., *Score Test: Historical Review and Recent Developments*, w: Balakrishnan N., Nagaraja H.N., Kannan N. (red.), Advances in Ranking and Selection, Multiple Comparisons, and Reliability. Statistics for Industry and Technology, Birkhäuser Boston, 2005.

Assessing variability of kernel-based estimators of label shift

Irène Gijbels[1], **Małgorzata Łazęcka**[2], **Jan Mielniczuk**[2], **Johan Segers**[2]

[1]Katholieke Unversiteit Leuven, Belgium
[2]Institute of Computer Science, Polish Academy of Sciences, Poland

Language of the presentation: English

We consider non-standard binary classification problem in which prior probability π' of class indicator for (unlabeled) target data differs from prior probability π for source distribution, with the aim to estimate π' (see [1], [2]). We use Reproducing Kernel Hilbert Space (RHKS) approach with kernel K and introduce two competing estimators: projection estimator and kernelized ratio estiamator. We derive an asymptotic distribution of both estimators via finding their linear representations based on expansions of U-statistics. We show that even for moderate sample sizes estimated asymptotic variance approximates well the variances of both estimators and use it to construct confidence intervals for the unknown π'.

References

[1] Iyer A., Nath S., Sarawagi S., *Maximum mean discrepancy for class ratio estimation: convergence bounds and kernel selection*, Proceedings of the 31st International Conference on Machine Learning, PMLR 32(1), 2014, 530-538.

[2] Vaz A., Izbicki R., Stern R., *Quantification Under Prior Probability Shift: the Ratio Estimator and its Extensions*, Journal of Machine Learning Research, 2019, p. 1-33.

Improved entropy loss estimation of the separable covariance structure under high-dimensionality

Monika Mokrzycka, Paweł Krajewski

Institute of Plant Genetics, Polish Academy of Sciences, Poznań, Poland

Language of the presentation: English

The datasets from experiments where several characteristics are measured repeatedly, e.g., in time are called doubly multivariate. In standard model of such experiment the matrix of observations for each object has, after vectorization, the normal distribution while the matrix of all observations, has the matrix normal distribution. For doubly multivariate data the most intuitive assumption would be the separable covariance structure (Kronecker product of two matrices), which means that the time points are correlated independently of characteristics and characteristics are correlated independently of time points.

We present estimation methods of the covariance matrix with a separable structure, utilizing the improved entropy loss function. These approaches are particularly applicable in high-dimensional settings where the standard entropy loss function is not suitable. Special cases of the general Kronecker product structure, incorporating compound symmetry, autoregression of order one, and identity structure, are also discussed. Simulation studies are presented to illustrate the behavior of the estimates with respect to changes in dimensionality and parameter values, as well as to demonstrate how estimation methods can be used to identify underlying true covariance structure. Finally, the impact of developmental heterochrony on the covariance matrix estimates is presented for plant data obtained by image phenotyping.

This study was funded by National Science Centre, Poland, under the OPUS call in the Weave programme, proj. no. 2021/43/I/NZ9/02519 and the NAWA Bekker program under project no. BPN/BEK/2023/1/00142.

Directed information in graphical models of causality

Wojciech Niemiro

University of Warsaw
Nicolaus Copernicus University, Toruń

Language of the presentation: Polish

I will speak of using the tools of information theory to quantify the strength of causal relations. In my talk I focus on models based on possibly cyclic directed graphs, with nodes corresponding to time series, as in [3]. In this setup the so called 'transfer entropy', introduced by Thomas Schreiber in [5] serves as a natural measure of Granger causality. It quantifies to what extent the past of X helps predict the future of Y, given the past of Y.

Graphical models allow to rigorously define the notion of intervention, ie. an actual or imagined controlled experiment in which values of some variables are set by the experimenter. Measuring the effect of intervention is quite different from measuring the predictive power. Several approaches based on information theory have been proposed in the literature. I will compare different definitions of 'interventional directed information', in particular those due to Ay & Polani [1] and Simoes & al. [4]. I will point out some problems and rather controversial interpretations.

References

[1] Ay N., Polani D., *Information flows in causal networks.* Advances in Complex Systems 11 (2008), 17–41.

[2] Niemiro W., *Causal graphs, composable stochastic processes and conditional independence.* Applicationes Mathematicae 51 (2024), 109–130.

[3] Niemiro W., Rajkowski Ł., *Local Dependence Graphs for Discrete Time Processes.* Proceedings of the Second Conference on Causal Learning and Reasoning, PMLR 213 (2023), 772-790.

[4] Simoes F. N. Q., Dastani M., van Ommen T., *Properties of Causal Entropy and Information Gain.*, in: Proceedings of the Third Conference on Causal Learning and Reasoning, PMLR 236 (2024), 188–208.

[5] Schreiber T., *Measuring Information Transfer.* Physical Review Letters 85 (2000), 461–464.

Nieobciążona estymacja wykładniczej niezawodności na podstawie danych lewostronnie cenzurowanych

Piotr Bolesław Nowak

Uniwersytet Wrocławski, Instytut Nauk Ekonomicznych

Język prezentacji: polski

W referacie przedstawimy zastosowanie techniki Rao-Blackwellizacji do konstrukcji nieobciążonego estymatora funkcji niezawodności dla rozkładu wykładniczego w przypadku danych lewostronnie cenzurowanych. Uzyskany estymator jest funkcją statystyki dostatecznej, jednak nie możemy stwierdzić, że ma on minimalną wariancję, ponieważ statystyka ta nie jest zupełna. Zaprezentujemy również asymptotyczny przedział ufności dla funkcji niezawodności wyznaczony na podstawie proponowanego estymatora, oraz porównamy jego własności z estymatorami wyznaczanymi innymi metodami.

Własności asymptotyczne estymatora największej wiarogodności w dyskretnych próbach cenzurowanych typu II i układach k-z-n

Filip Pazio

Politechnika Warszawska

Język prezentacji: polski

Prezentacja będzie poświęcona zagadnieniu estymacji największej wiarogodności w dyskretnych próbach cenzorowanych typu II. Cenzurowanie to polega na tym, że podczas eksperymentu obserwujemy n obiektów z *i.i.d.* czasami przeżycia i przerywamy eksperyment w momencie r-tej awarii, gdzie $r < n$ jest ustalone. Na to zagadnienie można też spojrzeć ze strony teorii niezawodności, gdzie stanowi ono problem wnioskowania na podstawie czasów awarii elementów układu k-z-n, dla $k = n - r + 1$, zaobserwowanych do momentu awarii tego układu. Układy k-z-n składają się z n elementów i działają dopóki k z ich elementów pozostaje sprawnych.

W trakcie prezentacji będę rozważał przypadek estymacji wielowymiarowego, nieznanego parametru dyskretnego rozkładu czasów przeżycia. Skupię się na osiągniętych uogólnieniach twierdzenia [1, Theorem 1]. Będę rozważał sytuację, w której $k = k(n) = \lfloor (1-q)n \rfloor$, gdzie $q \in (0,1)$, a $\lfloor x \rfloor$ oznacza największą liczbę całkowitą, mniejszą lub równą x. Przedstawię dwa twierdzenia mówiące o asymptotycznych własnościach estymatora oraz warunki regularności wymagane do ich zachodzenia. Omawiane będą istnienie, mocna zgodność oraz, w przypadku gdy q-ty kwantyl rozkładu czasów przeżycia jest jednoznacznie wyznaczony, asymptotyczna normalność i efektywność estymatora. Następnie przedstawię przykład pokazujący zastosowanie twierdzenia na wybranym rozkładzie prawdopodobieństwa.

Literatura

[1] Dembińska A., Jasiński, K., *Maximum likelihood estimators based on discrete component lifetimes of a k-out-of-n system*, TEST, 30 (2021), 407-428.

Estimation of inequality measures for grouped data

Alicja Jokiel-Rokita[1], **Sylwester Piątek**[1,2]

[1]Wrocław University of Science and Technology
[2]Institute of Mathematics, Polish Academy of Science

Language of the presentation: English

The studies on the income inequality are often based on grouped data in predefined ranges. We explore a parametric approach to the problem of estimation of quantile inequality curves and measures, focussing on distributions commonly used for income data, such as the Dagum, log-Student t, and generalised beta distributions. We study the performance of minimum distance estimators using various types of distance (including Hellinger distance) to estimate distribution parameters from grouped data, which leads to plug-in estimators of the inequality measures. A simulation study assesses the accuracy of these methods and their robustness, particularly under model misspecification. Moreover, real-world data from US census records are analysed to demonstrate the example of a practical application of these techniques.

Bootstrap with random block length for periodically correlated time series

Anna Dudek[1,2], **Jakub Pięta**[2]

[1]Faculty of Applied Mathematics, AGH University of Krakow, Krakow, Poland
[2]Aix Marseille University, CNRS, I2M, Marseille, France

Language of the presentation: English

Periodically Correlated (PC) time series generalize the concept of weak stationarity by allowing the first two moments to vary periodically in time with the same period. Despite their importance in applications such as climatology and economics, few resampling methods have been developed for PC data. Two main block bootstrap approaches exist for this type of data: the Generalized Seasonal Block Bootstrap (GSBB), introduced in [2], and the Extension of the Moving Block Bootstrap (EMBB), proposed in [3]. In this presentation, we introduce the Geometric Block-Length Bootstrap (GBLB), a generalization of the Stationary Bootstrap (see [1]) that extends the method to accommodate periodic structures in the data. Unlike other block bootstrap methods for stationary data that rely on a fixed block length, the SB method employs blocks of random lengths drawn from a geometric distribution, which reduces sensitivity to block length selection. We discuss preliminary results on the GBLB consistency for the overall mean of PC time series. In a simulation study, we compare the performance of the GBLB, GSBB, and EMBB methods when applied to PC data. In particular, we examine the empirical coverage probabilities of the resulting confidence intervals for the overall mean.

References

[1] Politis D.N., Romano J.P., *The Stationary Bootstrap*, Journal of the American Statistical Association, 89 (1994), 1303-1313.
[2] Dudek A.E., Leśkow J., Paparoditis E., Politis D.N., *A Generalized Block Bootstrap for Seasonal Time Series*, Journal of Time Series Analysis, 35 (2014), 89-114.
[3] Dudek A.E., *Block Bootstrap for Periodic Characteristics of Periodically Correlated Time Series*, Journal of Nonparametric Statistics, 30 (2018), 87-124.

The author acknowledges funding from A*MIDEX (*Aix-Marseille Initiative d'Excellence*), grant no. AMX-22-CEX-057, supported by the French government under the France 2030 investment plan, for doctoral research at the Institut de Mathématiques de Marseille, Aix-Marseille University.

Prior shift estimation for positive unlabeled data

Jan Mielniczuk[1,2], **Wojciech Rejchel**[3], **Paweł Teisseyre**[1,2]

[1]Polish Academy of Sciences
[2]Warsaw University of Technology
[3]Nicolaus Copernicus University

Language of the presentation: English

We study estimation of a class prior for unlabeled target samples which possibly differs from that of the source population. Moreover, it is assumed that the source data is partially observable: only samples from the positive class and from the whole population are available (PU learning scenario). We introduce a novel direct estimator of the class prior which avoids estimation of posterior probability in both populations and has a simple geometric interpretation. It is based on a distribution matching technique together with kernel embedding in a Reproducing Kernel Hilbert Space and is obtained as an explicit solution to an optimisation task.

We establish its asymptotic consistency, as well as an explicit, non-asymptotic bound on its deviation from the unknown prior, which is computable in practice. We also study finite sample behaviour for synthetic and real data and show that the proposal works consistently on par or better than its competitors.

Selekcja zmiennych w modelowaniu przyczynowym przy użyciu informacji wzajemnej

Krzysztof Rudaś[1,2], **Magdalena Bartczak**[2]

[1]Instytut Podstaw Informatyki PAN
[2]Politechnika Warszawska

Język prezentacji: polski

Modelowanie przyczynowe zajmuje się oszacowaniem efektu działania (terapii medycznej, kampanii reklamowej) na daną obserwację. Aby móc oszacować ten efekt, populację dzieli się losowo na część eksperymentalną, poddaną działaniu i część kontrolną, nie poddaną działaniu. Bardzo istotną kwestią w kontekście modelowania przyczynowego jest problem selekcji zmiennych. Może on poprawić jakość predykcyjną modelu i zwiększyć jego interpretowalność. W referacie przedstawię różne sposoby użycia informacji wzajemnej w dwóch podstawowych metamodelach przyczynowych (T-learner i X-learner) opisanych w pracach [1] i [2]. Przedstawione metody selekcji przeanalizuję pod kątem jakości predykcyjnej, a także umiejętności identyfikacji prawdziwie istotnych cech na danych symulacyjnych. Dodatkowo przedstawię rezultaty zaproponowanych metod na danych rzeczywistych.

Literatura

[1] Rudaś K., Jaroszewicz S., *Linear regression for uplift modeling*, Data Mining and Knowledge Discovery, Springer, 32 (2018), 1275–1305.

[2] Künzel S., Sekhon J., Bickel P., Yu B., *Metalearners for estimating heterogeneous treatment effects using machine learning*, Proceedings of the National Academy of Sciences, 116 (2019), 4156-4165.

Asymptotic normality of Fourier coefficients estimators with unknown lag-dependent cycle frequencies for generalized almost-cyclostationary processes

Anna Dudek[1,2], **Antonio Napolitano**[3], **Jakub Rutkowski**[1], **Jakub Wojdyła**[1]

[1]Faculty of Applied Mathematics, AGH University of Krakow, al. Mickiewicza 30, 30-059 Krakow, Poland
[2]Aix-Marseille University, CNRS, I2M, Marseille, France
[3]Parthenope University of Naples, Via Ammiraglio Ferdinando Acton, 38, 80133 Napoli NA, Italy

Language of the presentation: Polish

Generalized almost-cyclostationary processes (GACS) are a class of continuous stochastic processes, which, for every fixed delay, have statistical moments that are almost-periodic functions of time. The autocovariance function of such processes can be expanded into generalized Fourier series, whose frequencies are functions of delay (those frequencies are called lag-dependent cycle frequencies). The problem of estimating Fourier coefficients with unknown lag-dependent cycle frequencies will be addressed. The asymptotic normality of the considered estimator with estimated lag-dependent cycle frequencies, as well as the method of estimation of said frequencies, will be presented. Theoretical results will be supported by the results of numerical simulations.

References

[1] Napolitano A., *Cyclostationary Processes and Time Series. Theory, Applications, and Generalizations*, Elsevier, 2019.

[2] Napolitano A., *Generalizations of Cyclostationary Signal Processing: Spectral Analysis and Applications*, John Wiley & Sons, 2012.

[3] Dudek A., Napolitano A., Rutkowski J., Wojdyła J., *Asymptotic Properties of Generalized Almost-Cyclostationary Processes Estimators With Unknown Cycle Frequencies*, working paper.

Warunki zachowania porządków transformacji czasów życia komponentów o jednakowym rozkładzie przez czas życia systemu

Tomasz Rychlik[1], **Magdalena Szymkowiak**[2]

[1]Instytut Matematyczny PAN
[2]Politechnika Poznańska

Język prezentacji: polski

Istotną dziedziną teorii prawdopodobieństwa i statystyki są problemy zachowania wybranych własności rozkładów zmiennych losowych w efekcie działania wybranych operacji. Przykładami takich operacji są konwolucje, mieszaniny, granice, statystyki pozycyjne czy systemy niezawodnościowe. W referacie opiszemy warunki zachowania relacji w porządkach transformacji z ustalonym rozkładem jednakowo rozłożonych czasów życia komponentów systemów semikoherentnych przez czasy życia systemów. Wspólną nazwą porządków transformacji są określane częściowe porządki wypukłej transformacji, gwiaździsty i superaddytywny. Relacje porządków transformacji z wybranymi rozkładami wyznaczają rozmaite nieparametryczne klasy rozkładów czasów życia: o monotonicznej intensywności awarii IFR i DFR i jej uogólnieniach α-IGFR i α-DGFR, malejącej odwróconej intensywności awarii DRFR i jej uogólnieniach α-DRFR oraz α-IRFR, o monotonicznej uśrednionej intensywności awarii IFRA i DFRA i jej uogólnieniach α-IGFRA i α-DGFRA, rozkładów typów "lepszy niż używany"i "gorszy niż używany"NBU i NWU i ich uogólnieniach α-GNBU i α-GNWU czy o malejącym zlogarytmowanym ilorazie szans DLOR. Weryfikacja warunków zachowania tych klas polega na sprawdzeniu elementarnych własności funkcji zależnych od struktury systemu, kopuli zależności czasów życia jego komponentów i ekstremalnej dystrybuanty klasy rozkładów w wybranym porządku. Podane warunki są konieczne i dostateczne w przypadku, gdy ta ekstremalna dystrybuanta na nośnik ograniczony z dołu.

Literatura

[1] Rychlik T., Szymkowiak M., *Conditions for preserving transform order relations of identically distributed component lifetimes by system lifetimes*, Journal of Computational and Applied Mathematics, 476 (2026), Paper No. 117109, 16 pp.

Lifetimes and joint distributions of numbers of components in each state for multi-state discrete k-out-of-n systems

Agnieszka Goroncy, Krzysztof Jasiński, Jakub Sadowy

Nicolaus Copernicus University in Toruń

Language of the presentation: English

We consider multi-state discrete **k**-out-of-n systems composed of components which lifetimes are modeled by independent and identically distributed discrete random variables, with a special emphasis on four-state **k**-out-of-n systems - that is systems with two intermediate states between perfect functioning and failure. Based on the definition introduced by Huang et al. (2000) we focus on the lifetimes of such systems, including the distribution of the random vector representing the numbers of components in each state during system breakdown. We illustrate the theoretical results for four-state **k**-out-of-n systems with numerical examples concerning systems with components with geometrically distributed lifetimes following the Markov degradation process.

References

[1] Goroncy A., Jasiński K., Sadowy J., *Four-state **k**-out-of-n discrete system failure and the numbers of components in each state*, in review (2025).

[2] Goroncy A., Jasiński K., *Discrete time three-state **k**-out-of-n system's failure and numbers of components in each state.*, Journal of Computational and Applied Mathematics, 457 (2025), 116255.

D-optymalne rozmieszczenie tensometrów

Xabier Iriarte[1], Julen Bacaicoa[1], Jokin Aginaga[1], Aitor Plaza[1], Anna Szczepańska-Álvarez[2]

[1]Public University of Navarre (UPNA), Spain
[2]Uniwersytet Przyrodniczy w Poznaniu

Język prezentacji: polski

W referacie omówione zostanie wykorzystanie teorii układów optymalnych do opracowania algorytmu doboru czujników naprężeń. Na podstawie macierzy informacji dla estymatora najwiekszej wiarogodności w modelu stałym z efektami zakłócającymi oraz w modelach przekształconych określone zostanie położenie i orientacja tensometrów na powierzchni wału. Zaprezentowane wyniki będą konfrontacją rozważań teoretycznych z praktyką stosowaną w inżynierii mechanicznej.

Literatura

[1] Baksalary J.K., *A study of equivalence between Gauss Markoff model and its argumentation by nuisance parameters*, Math. Operationsforch u. Statist., ser. ststist.,15 (1984), 3-35.

[2] Iriarte X., Bacaicoa J., Aginaga J., Plaza A., Szczepańska-Álvarez A., *D-optimal strain sensor placement for mechanical load estimation in the presence of nuisance loads and thermal strain*, Sensors and Actuators A: Physical, 382 (2025), 116110.

AI-informed Non-linear Cox Regression for Time-to-event Analysis

Katarzyna Szczerba[1,3], **Laurent Malisoux**[2], **Christophe Ley**[1]

[1]University of Luxembourg
[2]Luxembourg Institute of Health
[3]Information Technology for Translation Medicine (ITTM)

Language of the presentation: English

The Cox proportional hazards model is the most commonly used method for multivariate survival analysis. Despite its many advantages, such as simplicity and interpretability, it has a serious drawback: it fails to capture non-linear relationships. In this study, we propose AI-informed Non-linear Cox Model, a method that uses insights from a highly predictive machine learning model, extracted with an interpretable machine learning tool, to integrate non-linear relationships into the traditional Cox model via means of splines. On simulated data with a deliberately introduced non-monotonic relationship between the predictor and the outcome variable, the AI-informed Cox model outperformed the traditional proportional hazards (PH) Cox model. Its concordance index (C-index) was also comparable to that of the best-performing machine learning model - gradient boosted Cox model. Similar results were observed when the models were applied to a prospective dataset in running.

References

[1] Chen Y., Jia Z., Mercola D., and Xie X., *A gradient boosting algorithm for survival analysis via direct optimization of concordance index*, Computational and Mathematical Methods in Medicine, 2013, Article ID 873595.

[2] Moncada-Torres A., van Maaren M.C., Hendriks M.P., et al., *Explainable machine learning can outperform Cox regression predictions and provide insights in breast cancer survival*, Scientific Reports, 11 (2021), 6968.

[3] Felice F., Ley C., Bordas S.P., and Groll A., *Boosting any learning algorithm with Statistically Enhanced Learning*, Scientific Reports, 15 (2025), 1605.

Czy znamy wynik tegorocznych wyborów prezydenckich?

Piotr Szulc

danetyka.com

Język prezentacji: polski

Tegoroczne wybory prezydenckie wzbudziły powszechne zainteresowanie statystyką. Po ogłoszeniu wyników pojawiło się wiele analiz wykazujących nieprawidłowości w liczeniu głosów. Z części z nich wynikało, że z bardzo wysokim prawdopodobieństwem w niewielkiej części komisji podano niepoprawne wyniki, z innych z kolei, że ich skala była na tyle duża, że mogły mieć wpływ na ogólnopolski wynik, niwelując różnice w liczbie głosów między kandydatami. Wiele z tych analiz było błędnych, mimo to zdobyły sporą popularność, bardzo silnie wpływając na opinię publiczną [1].

W reakcji na te analizy Prokurator Generalny powołał biegłych w celu ustalenia skali nieprawidłowości i ich potencjalnego wpływu na wynik [2]. Ich opinie oraz raport Fundacji Batorego, którego jestem współautorem [3, 4], były dyskutowane w Sądzie Najwyższym, w trakcie rozprawy stwierdzającej ważność wyborów.

Podczas swojej prezentacji przedstawię, jak przy pomocy regresji liniowej wskazać komisje, w których z dużym prawdopodobieństwem podano nieprawidłowe wyniki w drugiej turze wyborów. Zastanowimy się, jakie są ograniczenia tej metody i czy jesteśmy w stanie odróżnić fałszerstwa od przypadkowych pomyłek. W końcu odpowiem na pytanie, czy znamy ostateczny wynik wyborów.

Literatura

[1] https://polskieradio24.pl/artykul/3541213,ponowne-liczenie-glosow-tak-uwaza-niemal-polowa-polakow

[2] https://www.gov.pl/web/prokuratura-krajowa/komunikat-dot-opinii-bieglych-ws-wyborow-prezydenckich3

[3] https://www.batory.org.pl/publikacja/falszerstwa-czy-falszywe-alarmy-statystyczna-kontrola-wynikow-ii-tury-wyborow-prezydenckich-2025/

[4] https://danetyka.com/model-detekcji-anomalii/

ϕ–Divergence of System Lifetimes

Narayanaswamy Balakrishnan[1], **Maria Kateri**[2], **Magdalena Szymkowiak**[3]

[1]McMaster University, Canada
[2]RWTH Aachen University, Germany
[3]Poznan University of Technology, Poland

Language of the presentation: Polish

We propose new Jensen-ϕ divergence measures JD_ϕ, for comparing coherent and mixed systems with identically distributed components, and present their properties. Among them is the Jensen-Cressi-Read power divergence JD_λ, which is a parametric family of divergences. They provide information about the system's predictability - the smaller the divergence value, the higher the system's predictability. In general, the divergence measure of the system depends on both: the lifetime distribution of system's homogeneous components and its design through its system signature (see, e.g., Rychlik and Szymkowiak [2]).

Moreover, we show that the use of different values of λ, may result in different ordering of JD_λ divergences. Therefore, determining the right value of λ to properly order the predictability of the systems is the interesting challenge. Further on, we proof that JD_λ divergence measure is proportional to the power of the scale parameter of components' lifetime distribution. To demonstrate the practical application of Jensen-ϕ divergence of a system, we perform the analysis of Jensen-Cressie-Read divergence for systems with 3 and 4 components and we compare it with Jensen-Shannon divergence JS proposed by Asadi et al. [1] and the preferable system criterion $PS(T_{(0)})$ proposed by Toomaj et al. [3].

References

[1] Asadi M., Ebrahimi N., Soofi E.S., Zohrevand Y., *Jensen-Shannon information of the coherent system lifetime*, Reliability Engineering and System Safety, 156 (2016), 244–255.

[2] Rychlik T., Szymkowiak M., *Signature conditions for distributional properties of system lifetimes if component lifetimes are iid exponential*, IEEE Transactions on Reliability, 71 (2022), 590–602.

[3] Toomaj A., Chahkandi M., Balakrishnan N., *On the information properties of working used systems using dynamic signature*, Applied Stochastic Models in Business and Industry, 37 (2021), 318–341.

Positive-Unlabeled Regression: learning from partially labeled quantitative outcomes

Paweł Teisseyre, Jan Mielniczuk

Faculty of Mathemathics and Information Science, Warsaw University of Technology
Institute of Computer Science, Polish Academy of Sciences, Poland

Language of the presentation: English

We study a novel machine learning problem: Positive-Unlabeled Regression (PUR), which extends the classical Positive-Unlabeled (PU) learning framework to regression setting with a non-negative, discrete target variable. This setting arises naturally in applications where only some positive quantitative outcomes are reported, while the absence of a label may either indicate a true zero outcome or an unreported positive value. Applications include predicting the number of diseases a patient may have, the number of complications following an illness, or the severity level of a disease. We formalize the PUR problem and highlight the limitations of naive approaches that either use reported target variable or discard unlabeled data. To account for the inherent bias in such strategies, we propose two principled methods. The first is based on calibration of regression estimates using posterior probabilities from classical PU learning. The second builds on an empirical risk minimization framework, restating the target risk as a weighted function dependent on the instance-specific propensity score. We demonstrate both theoretically and empirically that our approaches yield improved performance over standard baselines.

Półnadzorowane niejednorodne ukryte modele Markowa

Filip Wichrowski, Katarzyna Kaczmarek-Majer

Instytut Badań Systemowych, Polska Akademia Nauk, Newelska 6, 01-447 Warszawa

Język prezentacji: polski

Ukryte modele Markowa (HMM) są popularnym narzędziem do analizy szeregów czasowych, stosowanym w wielu kluczowych zagadnieniach – od rozpoznawania mowy po bioinformatykę. Ich atrakcyjność wynika przede wszystkim z prostoty, przejrzystości, elastyczności i osadzenia na solidnych fundamentach matematycznych; jednocześnie, nie są to modele pozbawione wad. W swoim najbardziej podstawowym sformułowaniu wykorzystują jedynie uczenie nienadzorowane, które w praktyce uniemożliwia wykorzystanie jakiejkolwiek dodatkowej informacji na temat ukrytych stanów w procesie Markowa. Ponadto, zakładają one geometryczny rozkład czasu przebywania w stanie, co istotnie ogranicza ich zdolność do modelowania procesów o innej, potencjalnie bardziej złożonej strukturze.

Celem niniejszej pracy jest jednoczesne zaadresowanie obu wspomnianych problemów. Po pierwsze, w celu umożliwienia bardziej elastycznego modelowania dynamiki przejść między ukrytymi stanami, wykorzystujemy niejednorodny ukryty model Markowa (IHMM), w którym macierz przejść zależy również od czasu już spędzonego w danym stanie. Częściowe nadzorowanie realizowane jest natomiast poprzez wprowadzenie macierzy wag, wykorzystywanej do skalowania prawdopodobieństw emisji obserwacji w danym stanie. Taka modyfikacja bezpośrednio ogranicza maksymalny możliwy czas trwania stanu $d^{\max}$ i wpływa na identyfikowalność macierzy przejścia dla czasów trwania stanu $d > d^{\max}$. Z tego też powodu, proponujemy modelowanie macierzy przejścia dla $d > d^{\max}$ z wykorzystaniem brzegowych oczekiwanych liczb przejść pomiędzy stanami.

Walidacja zaproponowanego podejścia przeprowadzana jest w oparciu o dane symulowane, różniące się między innymi rozkładami prawdopodobieństw dla czasów trwania stanów. Uzyskane wyniki wskazują na większą moc predykcyjną w porównaniu do podejścia nienadzorowanego oraz podejścia częściowo nadzorowanego opartego na klasycznych ukrytych modelach Markowa.

Literatura

[1] Huang, X.D., Jack M.A., *Semi-continuous hidden Markov models for speech signals*, Computer Speech & Language, 3 (1989), 239-251.

[2] Ramesh P., Wilpon J., *Modeling state durations in hidden markov models for automatic speech recognition*, IEEE International Conference on Acoustics, Speech, and Signal Processing, (1992), 381-384.

[3] Li J., Lee J-Y., Liao L., *A new algorithm to train hidden Markov models for biological sequences with partial labels*, BMC Bioinformatics, 22 (2021), 162.

The project *ExplainMe: Explainable Artificial Intelligence for Monitoring Acoustic Features extracted from Speech (FENG.02.02-IP.05-0302/23)* is carried out within the First Team programme of the Foundation for Polish Science co-financed by the European Union under the European Funds for Smart Economy 2021-2027 (FENG).

A bootstrap test for the one-sided two-sample right-censored model

Grzegorz Wyłupek

University of Wrocław

Language of the presentation: English

The paper poses and solves a novel testing problem on detection of stochastic ordering of two survival curves when data is right-censored. The null hypothesis asserts the lack of the ordering while the alternative expresses its existence. The main focus is put on controlling the power function of the test outside the alternative. As a result, the asymptotic Type I error of the constructed solution is smaller than or equal to the fixed significance level α on the whole set where the stochastic ordering does not hold. A finite sample application of the test exploits bootstrap. The conducted simulation study shows that the Type I error is satisfactorily controlled under finite sample sizes and that the new solution competes well with the best and the most popular tests. A real data analysis demonstrates nice behaviour of the new test in practice.

References

[1] Burke M. D., Csörgő S., Horváth L., *A correction to and improvement of 'Strong approximations of some biometric estimates under random censorship'*, Probability Theory and Related Fields, 79 (1988), 51-57.

[2] Chang H-W., McKeague I. W., *Empirical likelihood based tests for stochastic ordering under right censorship*, Electronic Journal of Statistics, 10 (2016), 2511-2536.

[3] Ditzhaus M., Pauly M., *Wild bootstrap logrank tests with broader power functions for testing superiority*, Computational Statistics and Data Analysis, 136 (2019), 1-11.

[4] Fleming T. R., Harrington D. P., *Counting Processes and Survival Analysis*, John Wiley & Sons, 1991.

[5] Gehan E. A., *A generalized Wilcoxon test for comparing arbitrarily singly censored samples*, Biometrika, 52 (1965), 203-223.

[6] Gill R. D., *Censoring and stochastic integrals*, Statistica Neerlandica, 34 (1980), 124-124.

[7] Mantel N., *Evaluation of survival data and two new rank order statistics arising in its consideration*, Cancer Chemotherapy Reports, 50 (1966), 163-170.

[8] Pepe M. S., Fleming T. R., *Weighted Kaplan-Meier statistics: A class of distance tests for censored survival data*, Biometrics, 45 (1989), 497-507.

[9] Uno H., Tian L., Claggett B., Wei L. J., *A versatile test for equality of two survival functions based on weighted differences of Kaplan-Meier curves*, Statistics in Medicine, 34 (2015), 3680-3695.

[10] Wyłupek G., *A bootstrap test for the one-sided two-sample right-censored model*, under review (2025).

Lista uczestników

1. **prof. Małgorzata Bogdan**
   Uniwersytet Wrocławski

2. **dr Jacek Bojarski**
   Instytut Matematyki, Uniwersytet Zielonogórski

3. **dr hab. Agata Boratyńska**
   Szkoła Główna Handlowa SGH, Kolegium Analiz Ekonomicznych

4. **mgr inż. Adam Przemysław Chojecki**
   Wydział Matematyki i Nauk Informacyjnych Politechniki Warszawskiej

5. **dr Bogdan Ćmiel**
   Akademia Górniczo-Hutnicza w Krakowie

6. **mgr Iza Danielewska**
   Politechnika Warszawska

7. **dr hab. Anna Dudek**
   Akademia Górniczo-Hutnicza w Krakowie, Aix-Marseille University

8. **dr hab. inż. Katarzyna Filipiak**
   Politechnika Poznańska

9. **prof. Jean-Marc Freyermuth**
   Aix-Marseille University

10. **dr hab. Konrad Furmańczyk**
    Instytut Informatyki Technicznej SGGW

11. **prof. Lesław Gajek**
    Politechnika Łódzka

12. **prof. Aldo William Medina Garay**
    Université Paris Nanterre

13. **mgr Bartłomiej Gibas**
    AGH Akademia Górniczo-Hutnicza,
    Wydział Informatyki, Elektroniki i Telekomunikacji

14. **dr hab. Agnieszka Goroncy**
    Uniwersytet Mikołaja Kopernika w Toruniu

15. **prof. Przemysław Grzegorzewski**
    Politechnika Warszawska, Wydział Matematyki i Nauk Informacyjnych,
    Instytut Badań Systemowych PAN

16. **prof. Szymon Jaroszewicz**
    Instytut Podstaw Informatyki PAN,
    Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska

17. **dr Krzysztof Jasiński**
    Wydział Matematyki i Informatyki, Uniwersytet Mikołaja Kopernika w Toruniu

18. **dr Cristian Felipe Jiménez Varón**
University of York

19. **dr Joanna Karłowska-Pik**
Uniwersytet Mikołaja Kopernika w Toruniu

20. **prof. Jacek Koronacki**
Prof. em. Instytutu Podstaw Informatyki PAN

21. **dr hab. Wasyl Kowalczuk**
Instytut Podstawowych Problemów Techniki Polskiej Akademii Nauk

22. **mgr Anna Kozak**
Politechnika Warszawska

23. **prof. Rafał Kulik**
University of Ottawa

24. **dr hab. Łukasz Lenart**
Uniwersytet Ekonomiczny w Krakowie

25. **prof. Francyelle de Lima Medina**
Université Paris Nanterre

26. **mgr Karolina Łukasiewicz**
AGH University of Krakow, Universite Paris-Nanterre

27. **mgr Zuzanna Maciszewska**
AGH w Krakowie, Universite Paris Nanterre

28. **mgr Bartosz Majewski**
AGH Akademia Górniczo-Hutnicza

29. **mgr Katarzyna Mamla**
Politechnika Warszawska

30. **prof. Augustyn Markiewicz**
Uniwersytet Przyrodniczy w Poznaniu

31. **prof. Jan Mielniczuk**
Instytut Podstaw Informatyki PAN,
Wydział Matematyki i Nauk Informacyjnych, PW

32. **dr Adam Mieldzioc**
Politechnika Poznańska

33. **dr inż. Monika Mokrzycka**
Instytut Genetyki Roślin PAN

34. **mgr Dominika Mosur**
Aix Marseille University

35. **prof. Wojciech Niemiro**
Uniwersytet Warszawski, Uniwersytet Mikołaja Kopernika w Toruniu

36. **dr Piotr Nowak**
Uniwersytet Wrocławski. Wydział Prawa, Administracji i Ekonomii

37. **prof. Hernando Ombao**
KAUST King Abdullah University of Science and Technology, Saudi Arabia

38. **mgr Filip Pazio**
Politechnika Warszawska

39. **mgr Sylwester Piątek**
Politechnika Wrocławska

40. **mgr Jakub Pięta**
Institut de Mathématique de Marseille, Aix Marseille Université

41. **dr Łukasz Rajkowski**
Uniwersytet Warszawski

42. **dr hab. Wojciech Rejchel**
Uniwersytet Mikołaja Kopernika w Toruniu

43. **dr Krzysztof Rudaś**
IPI PAN, MINI PW

44. **dr Marcin Rudź**
Instytut Matematyki Politechniki Łódzkiej

45. **mgr Jakub Rutkowski**
AGH Akademia Górniczo-Hutnicza

46. **prof. Tomasz Rychlik**
Instytut Matematyczny PAN

47. **mgr Jakub Sadowy**
Uniwersytet Mikołaja Kopernika w Toruniu, Wydział Matematyki i Informatyki,
Katedra Statystyki Matematycznej i Eksploracji Danych

48. **mgr Katarzyna Szczerba**
University of Luxembourg

49. **dr Anna Szczepańska-Alvarez**
Katedra Metod Matematycznych i Statystycznych,
Uniwersytet Przyrodniczy w Poznaniu

50. **dr Piotr Szulc**
danetyka.com

51. **dr hab. Magdalena Szymkowiak**
Politechnika Poznańska

52. **dr hab. Paweł Teisseyre**
Politechnika Warszawska

53. **mgr inż. Filip Wichrowski**
Instytut Badań Systemowych PAN

54. **dr Katarzyna Woźnica**
Politechnika Warszawska, Instytut Badań Systemowych

55. **dr Grzegorz Wyłupek**
Uniwersytet Wrocławski, Instytut Matematyczny

56. **mgr Milena Zacharczuk**
Politechnika Warszawska, Wydział Matematyki i Nauk Informacyjnych

57. **mgr inż. Tomasz Zieliński**
Uniwersytet Mikołaja Kopernika w Toruniu,
Szkoła Doktorska Nauk Ścisłych i Przyrodniczych, Ośrodek Analiz Statystycznych

58. **prof. Roman Zmyślony**
Prof. em. Uniwersytetu Zielonogórskiego

Spis treści