

XLIII Konferencja
**Statystyka
Matematyczna**

Będlewo

4 – 8 Grudnia, 2017



Organizatorzy

Międzynarodowe Centrum Matematyczne im. Stefana Banacha
Instytut Matematyczny PAN
Komisja Statystyki Matematycznej Komitetu Matematyki PAN
Instytut Podstaw Informatyki PAN
Wydział Matematyki i Nauk Informatycznych PW

Sponsorzy

Międzynarodowe Centrum Matematyczne im. Stefana Banacha
Polska Akademia Nauk

Komitet Naukowy

prof. Teresa Ledwina, IM PAN
prof. Jacek Koronacki, IPI PAN
prof. Przemysław Grzegorzewski, IBS PAN/MINI PW

Komitet Organizacyjny

prof. Jan Mielniczuk, IPI PAN/MINI PW- przewodniczący
dr Paweł Teisseyre, IPI PAN
mgr Mariusz Kubkowski, MINI PW
mgr Norbert Ryciak, MINI PW/IPI PAN
mgr Waldemar Słonina, IPI PAN

Redaktor

mgr Norbert Ryciak

Spis treści

1	Program konferencji	4
2	Zaproszeni wykładowcy	10
3	Streszczenia	16
4	Lista uczestników	73

Program konferencji

Poniedziałek, 4 Grudnia

8:00	Śniadanie
9:00–9:45	Łukasz Dębowski <i>Praktyczne i teoretyczne problemy statystycznego modelowania języka naturalnego, część I (str. 13)</i>
9:45–10:00	Przerwa
10:00–10:20	Tadeusz Inglot <i>Adaptacyjne testy zgodności z rodziną rozkładów Poissona (str. 29)</i>
10:20–10:40	Ewa Strzałkowska-Kominiak <i>Test dopasowania rozkładu dla danych cenzurowanych bazujący na współczynniku maksymalnej korelacji (str. 58)</i>
10:40–11:00	Grzegorz Wyłupek <i>Adaptacyjny test permutacyjny w problemie k-prób dla danych cenzurowanych (str. 67)</i>
11:00–11:30	Przerwa kawowa
11:30–11:50	Bogdan Ćmiel <i>Adaptacyjna estymacja falkowa funkcji gęstości na sferze (str. 22)</i>
11:50–12:10	Przemysław Grzegorzewski <i>Testy zgodności dla danych przedziałowych (str. 28)</i>
12:10–12:30	Paweł Teisseyre <i>Selekcja zmiennych uwzględniająca koszty w klasyfikacji wieloetykietowej (str. 63)</i>
12:30–12:50	Łukasz Smaga <i>Diagonalne testy typu Walda w ogólnym modelu czynnikowym (str. 56)</i>
13:00	Obiad
15:00–15:45	Tadeusz Bednarski <i>Hugo Dionizy Steinhaus - Droga do współczesnej teorii prawdopodobieństwa (str. 18)</i>
15:45–16:20	Przerwa kawowa
16:20–16:40	Magdalena Szymkowiak <i>Zastosowanie funkcji intensywności starzenia (str. 62)</i>
16:40–17:00	Maria Szpak <i>Optymalne oszacowania całkowitego czasu trwania testu dla rozkładów z rosnącą uogólnioną intensywnością awarii (str. 60)</i>
17:00–17:20	Jakub Wojdyła <i>Praktyczne aspekty konstrukcji $\frac{5}{5}$ pasów ufności w problemie Spektora-Lorda-Willisa (str. 66)</i>
17:20–17:40	Agnieszka Siedlaczek <i>Nieparametryczna estymacja kwantyli (str. 53)</i>
18:00	Kolacja

Wtorek, 5 Grudnia

8:00	Śniadanie
9:00–9:45	Łukasz Dębowski <i>Praktyczne i teoretyczne problemy statystycznego modelowania języka naturalnego, część II (str. 13)</i>
9:45–10:00	Przerwa
10:00–10:20	Szymon Majewski <i>Unadjusted Langevin Monte Carlo jako optymalizacja wypukła (str. 36)</i>
10:20–10:40	Mateusz Wilk <i>Plany optymalne dla predykcji w modelach RCR z błędami heteroskedastycznymi (str. 64)</i>
10:40–11:00	Adam Mieldzioc <i>Aproksymacja macierzy kowariancji za pomocą wstęgowych macierzy Toeplitza (str. 40)</i>
11:00–11:30	Przerwa kawowa
11:30–11:50	Krzysztof Rudaś <i>Estymacja współczynnika wpływu w modelowaniu przyczynowości (str. 50)</i>
11:50–12:10	Norbert Ryciak <i>Analiza wydźwięku z wykorzystaniem sieci neuronowych i drzewiastej reprezentacji zdań (str. 52)</i>
12:10–12:30	Agnieszka Piliszek <i>Regresyjna charakteryzacja rozkładów Gamma i Kummera (str. 47)</i>
12:30–12:50	Mariusz Kubkowski <i>Selekcja zmiennych w źle wyspecyfikowanym modelu binarnym (str. 34)</i>
13:00	Obiad
15:00–15:45	Łukasz Dębowski <i>Praktyczne i teoretyczne problemy statystycznego modelowania języka naturalnego, część III (str. 13)</i>
15:45–16:20	Przerwa kawowa
16:20–16:40	Tomasz Górecki <i>Wybór zmiennych w klasyfikacji dla wielowymiarowych danych funkcjonalnych (str. 32)</i>
16:40–17:00	Maria Skupień <i>Metody funkcjonalnej analizy danych w detekcji uszkodzeń maszyn przemysłowych (str. 55)</i>
17:00–17:20	Anna Dudek <i>Metody bootstrapu blokowego dla okresowych szeregów czasowych (str. 23)</i>
17:20–17:40	Monika Mokrzycka <i>Aproksymacja macierzy kowariancji iloczynem Kroneckera z wykorzystaniem entropijnej funkcji straty (str. 43)</i>
18:00	Kolacja
19:00	Posiedzenie komisji

Środa, 6 Grudnia

8:00	Śniadanie
9:00–14:00	Wycieczka
14:00	Obiad
15:00–15:45	Małgorzata Bogdan <i>Wybrane zagadnienia analizy "szerokich" zbiorów danych, część I (str. 11)</i>
15:45–16:20	Przerwa kawowa
16:20–16:40	Wojciech Rejchel <i>Selekcja cech w oparciu o kryterium GIC z karą LASSO w modelach wysokowymiarowych (str. 49)</i>
16:40–17:00	Konrad Furmańczyk <i>Selekcja zmiennych w wysokowymiarowym modelu liniowym za pomocą metody Lasso i procedur testowania wielu hipotez (str. 25)</i>
17:00–17:20	Agnieszka Sołtys <i>DMRnet - pakiet środowiska R do wybierania zmiennych ciągłych i łączenia poziomów dla zmiennych jakościowych (str. 57)</i>
17:20–17:40	Piotr Pokarowski <i>Poprawianie Lasso dla wyboru modelu i predykcji (str. 48)</i>
18:30	Uroczysta kolacja

Czwartek, 7 Grudnia

8:00	Śniadanie
9:00–9:45	Małgorzata Bogdan <i>Wybrane zagadnienia analizy "szerokich" zbiorów danych, część II (str. 11)</i>
9:45–10:00	Przerwa
10:00–10:20	Błażej Miasojedow <i>Zbieżność algorytmów stochastycznej aproksymacji dla funkcji niegładkich i niewypukłych (str. 37)</i>
10:20–10:40	Wojciech Niemiro <i>Algorytmy pMCMC z repróbkowaniem poissonowskim (str. 44)</i>
10:40–11:00	Agnieszka Kulawik <i>Dobór stałej ucinającej dla funkcji Hubera wykorzystywanej w odpornej estymacji parametrów (str. 35)</i>
11:00–11:30	Przerwa kawowa
11:30–11:50	Tomasz Rychlik <i>Nierówności dla mieszanek uporządkowanych zmiennych losowych (str. 51)</i>
11:50–12:10	Agnieszka Goroncy <i>Oszacowania wartości oczekiwanych k-tych rekordów z rozkładów o rosnącej uogólnionej intensywności awarii (str. 26)</i>
12:10–12:30	Mariusz Bieniek <i>Charakteryzacje rozkładów dyskretnych przez pojedynczą regresję wartości (str. 19)</i>
13:00	Obiad
15:00–15:20	Katarzyna Filipiak <i>Zastosowania rzutu macierzy na zbiór macierzy o zadanej strukturze w modelach wielowymiarowych (str. 24)</i>
15:20–15:40	Roman Zmyślony <i>Estymacja i testowanie hipotez w wielowymiarowych modelach normalnych (str. 71)</i>
15:40–16:00	Mariusz Grządział <i>O nieujemnej, minimalnie obciążonej estymacji kwadratowej w modelu regresji liniowej (str. 27)</i>
16:00–16:30	Przerwa kawowa
16:30–16:50	Joanna Karłowska-Pik <i>Zastosowanie metod eksploracji danych do predykcji morfologii włosów i koloru oczu na podstawie materiału genetycznego (str. 31)</i>
16:50–17:10	Karol R. Opara <i>Optymalne planowanie pomiarów ruchu z wykorzystaniem programowania całkowitoliczbowego (str. 45)</i>
17:10–17:30	Jacek Bojarski <i>Stochastyczna metoda redukcji obciążeń szczytowych w systemie elektroenergetycznym (str. 20)</i>
18:00	Kolacja

Piątek, 8 Grudnia

8:00	Śniadanie
8:45–9:30	Małgorzata Bogdan <i>Wybrane zagadnienia analizy "szerokich" zbiorów danych, część III (str. 11)</i>
9:30–9:45	Przerwa
9:45–10:05	Magdalena Alama-Bućko <i>Obszary ufności dla parametrów położenia i skali z rozkładów Gumbela i Burra oparte na k-tych rekordach (str. 17)</i>
10:05–10:25	Krzysztof Jasiński <i>Rekurencyjne zależności pomiędzy momentami k-tych rekordów z populacji dyskretnej (str. 30)</i>
10:25–10:45	Piotr Sulewski <i>Badanie wpływu sposobu wyznaczania dystrybuanty doświadczalnej na moc testu Kołmogorowa-Smirnowa (str. 59)</i>
10:45–11:15	Przerwa kawowa
11:15–11:35	Andrzej Michalski <i>Badanie efektywności estymatorów jądrowych (str. 38)</i>
11:35–11:55	Aleksander Zaigrajew <i>O prawdopodobieństwach inwersji większości w systemach z dwuetapowym głosowaniem (str. 69)</i>
11:55–12:15	Jan Mielniczuk <i>O pomiarze siły interakcji (str. 41)</i>
12:30	Obiad

Zaproszeni wykładowcy



Małgorzata Bogdan uzyskała dyplom magistra z matematyki (1992) i doktora nauk matematycznych (1996, specjalność statystyka matematyczna) na Politechnice Wrocławskiej i habilitację z nauk technicznych (2009, specjalność informatyka) w Instytucie Podstaw Informatyki Polskiej Akademii Nauk. Pracuje jako profesor nadzwyczajny w Instytucie Matematyki Uniwersytetu Wrocławskiego. W roku 2000 przebywała na stypendium post-doktoranckim Fundacji Nauki Polskiej na Uniwersytecie Stanu Waszyngton, a w latach 2012/2013 na stypendium Fulbrighta na Uniwersytecie Stanforda. Pracowała jako profesor wizytujący na Uniwersytecie Stanu Waszyngton, Uniwersytecie Purdue i Uniwersytecie Stanforda w USA, Uniwersytecie Wiedenskim w Austrii i Uniwersytecie Angers we Francji. W latach 2007 i 2008 otrzymała “Women for Math Science Award” od Instytutu Matematyki Politechniki w Monachium. Jest przewodniczącą Komisji Zastosowań Matematyki przy Komitecie Matematyki PAN i członkiem Komitetów Edytorskich “Statistics” i “Scandinavian Journal of Statistics”. Głównym obszarem jej zainteresowań naukowych jest redukcja wymiaru w dużych zbiorach danych i zastosowania do analizy danych genetycznych.

Wybrane zagadnienia analizy ”szerokich” zbiorów danych

Małgorzata Bogdan

Uniwersytet Wrocławski

Przez ”szeroki” zbiór danych rozumiemy zbiór, w którym liczba zmiennych jest większa od liczby obserwacji. Analiza takich zbiorów zwykle polega na identyfikacji nisko-wymiarowego modelu, który dobrze opisuje proces generujący dane. W trakcie wykładów omówimy kilka podstawowych zagadnień związanych z redukcją wymiaru - wielokrotne testowanie, identyfikację istotnych zmiennych w modelach liniowych, identyfikację struktury gaussowskiego modelu graficznego, analizę składowych głównych i analizę skupień w podprzestrzeniach. Omówimy klasyczne i nowe metody podejścia do tych zagadnień, które zilustrujemy wynikami teoretycznymi i wynikami symulacji komputerowych.



Łukasz Dębowski otrzymał stopień magistra fizyki teoretycznej na Uniwersytecie Warszawskim w 1994, stopień doktora nauk matematycznych w zakresie informatyki w Polskiej Akademii Nauk w 2005 oraz stopień doktora habilitowanego nauk matematycznych w zakresie informatyki w Polskiej Akademii Nauk w 2015. Odbył wizyty badawcze w Instytucie Lingwistyki Formalnej i Stosowanej na Uniwersytecie Karola w 2001, w Santa Fe Institute w 2002 oraz w School of Computer Science and Engineering na Uniwersytecie Nowej Południowej Walii w 2006. Ponadto odbył staż podoktorski w Centrum Wiskunde & Informatica od 2008 do 2009 i był profesorem wizytującym w Department of Advanced Information Technology na Uniwersytecie Kyushu w 2015. Obecnie zatrudniony jest jako profesor nadzwyczajny w Instytucie Podstaw Informatyki Polskiej Akademii Nauk. Jego zainteresowania naukowe obejmują teorię informacji oraz statystyczne modelowanie języka naturalnego.

Praktyczne i teoretyczne problemy statystycznego modelowania języka naturalnego

Łukasz Dębowski

Instytut Podstaw Informatyki PAN

Statystyczne modelowanie języka to problem przypisania określonej wartości prawdopodobieństwa dowolnie długim wypowiedziom w języku naturalnym (np. angielskim lub polskim). Problem ten znajduje praktyczne zastosowania między innymi w automatycznym rozpoznawaniu mowy, klawiaturach predykcyjnych telefonów komórkowych i maszynowym tłumaczeniu. Wyzwaniem dla tych zastosowań jest sformułowanie głębszej teorii matematycznej oraz poprawa skuteczności stosowanych heurystyk. W trzyczęściowym wykładzie opowiem o historii problemu i inżynierskich podejściach do jego rozwiązywania oraz o próbach identyfikacji niektórych własności języka naturalnego jako procesu stochastycznego i inspirowanych tymi próbami dwóch problemach matematycznych. Pierwszym z problemów matematycznych jest powiązanie logarytmiczno-potęgowego tempa wzrostu maksymalnego powtórzenia w tekście z potęgowym tempem wzrostu entropii bloku liter. Drugim z problemów matematycznych jest powiązanie potęgowego wzrostu liczby różnych słów w tekście z własnością mocnej nieergodyczności procesu stochastycznego generującego tekst i potęgowym wzrostem liczby niezależnych faktów opisywanych w tekście.

Literatura

- [1] F. Jelinek, 1997, *Statistical Methods for Speech Recognition*, The MIT Press.
- [2] C. D. Manning, H. Schütze, 1999, *Foundations of Statistical Natural Language Processing*, The MIT Press.
- [3] R. Köhler, G. Altmann, R. G. Piotrowski, 2005, *Quantitative Linguistik. Ein internationales Handbuch / Quantitative Linguistics. An International Handbook*, Walter de Gruyter.

[4] Ł. Dębowski, 2016, *Maximal Repetition and Zero Entropy Rate*, URL: <http://arxiv.org/abs/1609.04683>

[5] Ł. Dębowski, 2017, *Is Natural Language Strongly Nonergodic? A Stronger Theorem about Facts and Words*, URL: <http://arxiv.org/abs/1706.04432>

Streszczenia

Obszary ufności dla parametrów położenia i skali z rozkładów Gumbela i Burra oparte na k -tych rekordach.

**Magdalena Alama-Bućko¹ i
Aleksander Zaigrajew²**

¹ Instytut Matematyki i Fizyki, Uniwersytet Technologiczno-Przyrodniczy, Bydgoszcz
² Wydział Matematyki i Informatyki, Uniwersytet Mikołaja Kopernika, Toruń

Niech $x = (x_1, \dots, x_n)$ będzie próbą z rozkładu $P_\theta, \theta = (\theta_1, \theta_2)$, gdzie $\theta_1 \in R$ jest parametrem położenia, $\theta_2 > 0$ parametrem skali. W pracy [2] została zaprezentowana konstrukcja obszaru ufności dla wektora parametrów θ oparta na kombinacji liniowej dwóch statystyk pozycyjnych. Teraz rozważamy sytuację, gdy zamiast kombinacji dwóch statystyk pozycyjnych stosujemy kombinację dwóch statystyk rekordowych rzędu k . Powyższą metodę stosujemy do estymacji parametrów z rozkładów Gumbela i Burra.

Literatura

- [1] I. Malinowska, P. Pawlas, D. Szynal (2005), *Estimation of the parameters of Gumbel and Burr distributions in terms of k -th record values*, Appl. Math. 32, 375–393.
- [2] A. Zaigraev, M. Alama-Bućko (2013), *On optimal choice of order statistics in large samples for the construction of confidence regions for the location and scale*, Metrika. Vol. 76, pp. 577-593.

Hugo Dionizy Steinhaus - Droga do współczesnej teorii prawdopodobieństwa

Tadeusz Bednarski

Uniwersytet Wrocławski

Rozwój teorii prawdopodobieństwa, zapoczątkowany przez Gerolamo Cardano i wiązany z pośmiertną publikacją jego dzieła „Book on Games of Chance” jest nieustanną, trwającą niemal 400 lat interakcją pomiędzy matematycznymi wynikami teorii i rosnącym oczekiwaniem z obszaru fizyki, biologii i ekonomii. To także trwające niemal stulecie próby ujednoczenia opisu prawdopodobieństw: klasycznego, geometrycznego i tzw. przeliczalnego, zakończone w roku 1933 słynną pracą Kołmogorowa.

Głównym celem wykładu jest analiza udziału H. Steinhausa w tym intrygującym procesie, określenie jego punktu widzenia co do idei aksjomatyzacji, a także wskazanie istotnej roli innych polskich matematyków w rozwoju teorii prawdopodobieństwa.

Literatura

- [1] E. Borel, *Les probabilités dénombrables et leurs applications arithmétiques*, M. Rend. Circ. Matem. Palermo, 1909.
- [2] H. Steinhaus, *Les probabilités dénombrables et leur rapport ? la théorie de la mesure*, Fund. Math. 4 286–310, 1923.
- [3] A. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin, 1933
- [4] J. von Plato, *Creating Modern Probability. Its Mathematics, Physics and Philosophy in Historical Perspective*, Cambridge Univ. Press, 1994
- [5] G. Shafer G. and V. Vovk, *The Sources of Kolmogorov’s Grundbegriffe*, Statistical Science. Vol. 21, 2006.

Charakteryzacje rozkładów dyskretnych przez pojedynczą regresję wartości rekordowych

Mariusz Bieniek

Instytut Matematyki UMCS, Lublin

Jednym z podstawowych problemów statystyki jest wyznaczenie nieznanego rozkładu prawdopodobieństwa pewnej wielkości losowej na podstawie danych jej obserwacji. Dla danego ciągu obserwacji pewnej wielkości losowej obserwacja jest nazywana wartością rekordową, jeżeli jest ona większa niż wszystkie poprzednie obserwacje. Dla rozkładów dyskretnych ma również sens rozważanie słabych wartości rekordowych, czyli obserwacji, które są większe lub równe niż wszystkie poprzednie obserwacje. W referacie zostanie omówione nowe podejście do problemu identyfikacji dyskretnych rozkładów prawdopodobieństwa przy użyciu tylko jednej funkcji regresji wartości rekordowych. Korzystając z ich własności Markowa pokażemy, że pojedyncza funkcja regresji (słabych) rekordów wyznacza pewien rozkład dyskretny jednoznacznie wtedy i tylko wtedy, gdy odpowiednie równanie różnicowe (lub układ równań różnicowych) ma dokładnie jedno rozwiązanie. Otrzymany rezultat zostanie zastosowany do wyprowadzenia znanych oraz nowych charakterystyk rozkładów dyskretnych.

Stochastyczna metoda redukcji obciążeń szczytowych w systemie elektroenergetycznym

Grzegorz Benysek, Jacek Bojarski, Marcin Jarnut, Robert Smoleński, Szymon Wermiński

Wydział Matematyki, Informatyki i Ekonometrii
Instytut Inżynierii Elektrycznej
Uniwersytet Zielonogórski

W referacie przedstawione będą wyniki badań zespołowych nad wykorzystaniem systemu rozproszonej automatyki Decentralized Active Demand Response (DADR) do redukcji szczytów w krzywej dziennego obciążenia polskiego systemu elektroenergetycznego. Zagadnienie to jest szczególnie ważne z punktu widzenia podnoszenia efektywności pracy systemu, ale również jego niezawodności, z tego też względu dąży się do redukcji mocy o 1 GW w okresie szczytowym co stanowi około 4% wartości mocy szczytowej. Spośród znanych metod obniżania obciążeń szczytowych można wymienić m.in. rozwiązania taryfowe, jednak nie zawsze przynoszą one pożądaną skutek.

Opracowany algorytm stochastyczny redukuje szczytowe zużycie energii bez negatywnego oddziaływania na system elektroenergetyczny (nie występują skoki i oscylacje mocy).

Literatura

- [1] G. Benysek, J. Bojarski, M. Jarnut, and R. Smoleński, *Decentralized active demand response (DADR) system for improvement of frequency stability in distribution network*, Electric Power Systems Research Vol. 134 (2016), 80–87.
- [2] G. Benysek, J. Bojarski, R. Smoleński, M. Jarnut, and S. Wermiński, *Application of stochastic decentralized active demand response (DADR) system for load frequency control*, IEEE Transactions on Smart Grid Vol. PP (2016), no. 99, 1–10.

[3] S. Wermiński, M. Jarnut, G. Benysek, and J. Bojarski, *Demand side management using DADR automation in the peak load reduction*, Renewable and Sustainable Energy Reviews Vol. 67 (2017), 998–1007.

Adaptacyjna estymacja falkowa funkcji gęstości na sferze

**Bogdan Ćmiel, Natalia Jarzębkowska,
Karol Dziedziul**

AGH Kraków, PG Gdańsk

Przedstawimy adaptacyjny estymator funkcji gęstości na d -wymiarowej sferze jednostkowej \mathbb{S}^d ($d \geq 2$), używając odpowiednio skonstruowanej kraty falkowej. Kratę tą otrzymujemy przez przekształcenie systemu falkowego na \mathbb{R}^d przez odpowiednie operatory stereograficzne. Dowodzimy, że nasz estymator osiąga optymalne tempo zbieżności na pewnych klasach funkcji typu Biesowa poprzez adaptację do nieznanego parametru gładkości. Prezentujemy również metodę implementacji estymatora oraz jego działanie w eksperymencie numerycznym.

Literatura

- [1] P. Auscher, G. Weiss, M. V. Wickerhauser, *Local sine and cosine bases of Coifman and Meyer and the construction of smooth wavelets*, Wavelet Anal. Appl., 2, Academic Press, Boston, MA, 1992
- [2] M. Bownik, K. Dziedziul, *Smooth Orthogonal Projections on Sphere*, Constructive Approximation vol. 41 (1), 23-48, 2015
- [3] F. Dai, *Characterizations of function spaces on the sphere using frames*, Transactions of the American Mathematical Society, vol. 359 (2), 567-589, 2007
- [4] F. Dai, Y. Xu, *Approximation theory and harmonic analysis on spheres and balls*, Springer-Verlag, New York, 2013
- [5] P. I. Lizorkin, and Kh. P. Rustamov, *Nikol'skij-Besov spaces on the sphere in connection with approximation theory*, Trudy Matematicheskogo Instituta im. VA Steklova, vol. 204, 172-200, 1993
- [6] F. Narcowich, P. Petrushev, J. Ward, *Decomposition of Besov and Triebel-Lizorkin spaces on the sphere*, Journal of Functional Analysis, vol. 238 (2), 530-564, 2006

Metody bootstrapu blokowego dla okresowych szeregów czasowych

Anna Dudek

Akademia Górniczo-Hutnicza

Omówione zostaną dwie metody bootstrapu blokowego, które mogą być stosowane w analizie okresowych szeregów czasowych. Są to: Uogólniona Metoda Bootstrapu Bloków Sezonowych (Generalized Seasonal Block Bootstrap (GSBB)) oraz Rozszerzona Metoda Bloków Ruchomych (Extension of Moving Block Bootstrap (EMBB)). Główną zaletą techniki GSBB jest fakt, iż zachowuje ona strukturę okresową zawartą w danych. W rezultacie stosunkowo łatwo można otrzymać zgodne estymatory wielu charakterystyk z dziedziny czasu oraz z dziedziny częstotliwości szeregów okresowych. Jednakże, aby móc użyć GSBB, trzeba znać długość okresu. Jeśli wielkość ta nie jest znana lub rozważany sygnał powstał jako połączenie dwóch komponent okresowych o niewspółmiernych długościach okresu, analiza danych może być prowadzona przy wykorzystaniu metody EMBB. Podczas prezentacji omówione zostaną wady i zalety obu technik oraz przedstawione zostaną wyniki dotyczące ich zgodności dla podstawowych charakterystyk szeregów okresowo skorelowanych.

Literatura

- [1] A.E. Dudek, J. Leśkow, E. Paparoditis, D. Politis (2014), *A generalized block bootstrap for seasonal time series*, J. Time Ser. Anal., 35, 89-114.
- [2] A.E. Dudek (2018), *Block bootstrap for periodic characteristics of periodically correlated time series*, Journal of Nonparametric Statistics - w druku

Zastosowania rzutu macierzy na zbiór macierzy o zadanej strukturze w modelach wielowymiarowych

Katarzyna Filipiak

Instytut Matematyki, Politechnika Poznańska

Filipiak i Klein (2017) scharakteryzowali najlepsze przybliżenie (w sensie normy Frobeniusa) dowolnej macierzy iloczynem Kroneckera dwóch macierzy, przy czym jeden z czynników iloczynu ma strukturę kompletnej symetrii lub procesu autoregresji I-go rzędu. Celem referatu jest przedstawienie zastosowań wyznaczonej aproksymacji w teorii wielopoziomowych modeli wielowymiarowych. Pokażemy, że najlepsze przybliżenie w sensie normy Frobeniusa może służyć do regularyzacji macierzy kowariancji, do estymacji nieznannej macierzy kowariancji (nawet w przypadku danych wysokopoziomowych, w których wielkość próby jest mniejsza od liczby badanych cech), jak również jako miara odległości między hipotezami w zagadnieniu testowania hipotez o strukturze kowariancyjnej.

Referat jest częściowo dofinansowany z działalności statutowej nr 04/43/DSPB/0088.

Literatura

[1] K. Filipiak and D. Klein (2017), *Approximation with a Kronecker product structure with one component as compound symmetry or autoregression*, Submitted

Selekcja zmiennych w wysokowymiarowym modelu liniowym za pomocą metody Lasso i procedur testowania wielu hipotez

Konrad Furmańczyk

Wydział Zastosowań Informatyki i Matematyki SGGW

W referacie zostanie przedstawiona dwustopniowa procedura selekcji zmiennych w wysokowymiarowym modelu liniowym gdy liczba predyktorów p znacznie przewyższa liczbę obserwacji n . W pierwszym kroku wykonana zostanie procedura LASSO, a w drugim kroku procedura stepdown testowania wielu hipotez. Podamy wynik teoretyczny dotyczący oszacowania prawdopodobieństwa prawidłowego wyboru modelu oraz wyniki symulacyjne.

Literatura

- [1] K. Furmańczyk, *On some stepdown procedures with application to consistent variable selection in linear regression*, Statistics, (2015), **49**, 614-628
- [2] K. Furmańczyk and W. Rejchel, *Stepdown procedures with the Lasso for high dimensional linear model selection*, preprint, (2017)

Oszacowania wartości oczekiwanych k -tych rekordów z rozkładów o rosnącej uogólnionej intensywności awarii

Agnieszka Goroncy

Wydział Matematyki i Informatyki,
Uniwersytet Mikołaja Kopernika w Toruniu

Praca dotyczy górnych oszacowań wartości oczekiwanych k -tych wartości rekordowych centrowanych względem średniej z populacji. Rozważany jest przypadek, gdy rekordy pochodzą z klasy rozkładów o rosnącej uogólnionej intensywności awarii, które definiuje się za pomocą porządków wypukłych dystrybuant. Szczególnie ważnymi przykładami takich rozkładów są rozkłady o rosnącej gęstości (ang. increasing density, ID) oraz rosnącej intensywności awarii (increasing failure rate, IFR). Prezentowane oszacowania uzyskane zostały metodą rzutowania, wyrażone są w jednostkach skali opartych na odchyleniu standardowym rozkładu wyjściowego.

Literatura

- [1] M. Bieniek, (2007), *Projection mean-variance bounds on expectations of k th record values from restricted families*, *Commun. Statist. — Theor. Meth.*, **36**, 679–692.
- [2] A. Goroncy, *Optimal upper bounds on expected k th record values from IGFR distributions*, w przygotowaniu.
- [3] A. Goroncy, T. Rychlik (2015), *Optimal bounds on expectations of order statistics and spacings from nonparametric families of distributions generated by convex transform order*, *Metrika*, **78**, 175–204.

O nieujemnej, minimalnie obciążonej estymacji kwadratowej w modelu regresji liniowej

Mariusz Grządziel

Uniwersytet Przyrodniczy we Wrocławiu

Rozważany jest problem nieujemnej estymacji funkcji parametrycznej $\gamma(\beta, \sigma) = \beta' H \beta + h \sigma^2$ w modelu regresji liniowej $\mathcal{M}\{y, X\beta, V(\sigma) = \sigma^2 I\}$, gdzie H jest nieujemnie określoną macierzą i $h \geq 0$. W pracy Gnota i in. [1] zaproponowano podejście, w którym γ jest estymowana przy użyciu formy kwadratowej $y' A y$, gdzie A jest macierzą nieujemnie określoną spełniającą pewne kryteria optymalności.

Obliczanie macierzy A może stanowić wyzwanie. Podczas referatu zostaną porównane różne podejścia do zagadnienia wyznaczenia tej macierzy rozważane w pracy [2].

Literatura

- [1] S. Gnot, G. Trenkler i R. Zmysłony, *Nonnegative minimum biased quadratic estimation in the linear regression models*, J. Multivariate Analysis 54 (1995), 113–125
- [2] M. Grządziel, *On nonnegative minimum biased quadratic estimation in the linear regression models*, Praca zgłoszona do druku

Testy zgodności dla danych przedziałowych

Przemysław Grzegorzewski

Wydział Matematyki i Nauk Informacyjnych
Politechniki Warszawskiej,
Instytut Badań Systemowych Polskiej Akademii Nauk

Testy zgodności należą do podstawowych narzędzi wnioskowania statystycznego i analizy danych. Terminem tym obejmuje się zarówno procedury służące do badania zgodności rozkładu próbki z pewnym rozkładem hipotetycznym, jak i do oceny, czy rozkłady dwóch (lub większej liczby) próbek różnią się istotnie.

W ostatnim czasie dużym zainteresowaniem cieszą się metody wnioskowania na podstawie danych przedziałowych. Za pomocą przedziałów można bowiem w stosunkowo prosty sposób modelować brak precyzji, niepewność wynikającą z braku informacji, fluktuacje mierzonej wielkości itp.

W referacie zostaną zaprezentowane metody uogólniania klasycznych testów zgodności dla pojedynczej próbki i dla dwóch próbek na przypadek danych przedziałowych, z uwzględnieniem dwóch perspektyw postrzegania takich danych: epistemicznej i ontycznej (por. [1,2,3]).

Literatura

- [1] I. Couso, D. Dubois, *Statistical reasoning with set-valued information: Ontic vs. epistemic views*, International Journal of Approximate Reasoning 55 (2014), 1502–1518.
- [2] P. Grzegorzewski, *Testing goodness-of-fit with interval-valued data*, w: Proceedings of the 3rd International Symposium on Fuzzy Sets - Uncertainty Modelling ISFS 2017, U. Bentkowska i poz. (red.), Rzeszów 2017, s. 9-15.
- [3] P. Grzegorzewski, *The Kolmogorov goodness-of-fit test for interval-valued data*, w: 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2017).

Adaptacyjne testy zgodności z rodziną rozkładów Poissona

Tadeusz Inglot

Wydział Matematyki Politechniki Wrocławskiej

W 1997 roku we wspólnej pracy z W. Kallenbergiem i T. Ledwiną zaproponowaliśmy ogólną konstrukcję adaptacyjnych testów zgodności z rodziną parametryczną rozkładów ciągłych na prostej. Okazuje się, że ta konstrukcja daje się zastosować także do rodzin rozkładów dyskretnych. Trochę nieoczekiwanie, w przypadku rodziny rozkładów Poissona, dla której znanych jest kilka bardzo dobrych testów (np. Klar, 1999), otrzymujemy równie dobre testy i do tego czułe nie tylko na ‘typowe’ ale i ‘nie-typowe’ odstępstwa od tej rodziny. Potwierdza to znaną własność asymptotycznej optymalności adaptacyjnych testów wynikowych dla szerokiego spektrum alternatyw. W referacie przedstawimy zarówno konstrukcję jak i własności empiryczne nowych testów.

Literatura

- [1] T. Inglot, W. C. M. Kallenberg, T. Ledwina, *Data driven smooth tests for composite hypotheses*, The Annals of Statistics
- [2] B. Klar, *Goodness-of-fit tests for discrete models based on the integrated distribution function*, Metrica

Rekurencyjne zależności pomiędzy momentami k -tych rekordów z populacji dyskretnych

Krzysztof Jasiński

Wydział Matematyki i Informatyki,
Uniwersytet Mikołaja Kopernika w Toruniu

Stosunkowo niedawno, bo w 2005 roku okazało się, że funkcjonujące w literaturze definicje k -tych rekordów, uznawane za równoważne, w rzeczywistości takie nie są właśnie w przypadku dyskretnym. Stąd pojawiła się konieczność uporządkowania terminologii poprzez wprowadzenie trzech nierównoważnych definicji: mocnych k -tych rekordów, k -tych rekordów, słabych k -tych rekordów. Naszym celem będzie uzyskanie rekurencyjnych zależności pomiędzy momentami k -tych rekordów z populacji dyskretnych w przypadku każdej z wymienionych wyżej grup. Uprościmy otrzymane wyniki w sytuacji, gdy obserwacje mają rozkład geometryczny.

Literatura

[1] Jasiński, K., 2017., *Relations for product moments and covariances of k th records from discrete distributions.*, *Metrika*, w recenzji.

Zastosowanie metod eksploracji danych do predykcji morfologii włosów i koloru oczu na podstawie materiału genetycznego

Joanna Karłowska-Pik

Wydział Matematyki i Informatyki,
Uniwersytet Mikołaja Kopernika w Toruniu

Predykcja morfologii włosów oraz koloru oczu na podstawie materiału DNA jest wykorzystywana w kryminologii w celu ustalenia możliwego wyglądu nieznanego sprawcy przestępstwa. W różnych populacjach, także europejskich, istnieją statystycznie istotne różnice w morfologii włosów i kolorze oczu, dlatego budowa modeli predykcyjnych musi się odbywać w odniesieniu do konkretnej populacji. Celem niniejszego referatu jest zaprezentowanie próby budowy takich modeli w oparciu o wytypowane polimorfizmy pojedynczych nukleotydów oraz płęć z wykorzystaniem różnych metod eksploracji danych, w tym regresji logistycznej, drzew i lasów losowych oraz sieci neuronowych, dla populacji polskiej.

Literatura

- [1] E. Pośpiech, J. Karłowska-Pik, M. Marcińska et al., *Evaluation of the predictive capacity of DNA variants associated with straight hair for Europeans*, Forensic Science International: Genetics, 19, pp. 280-288 (2015)
- [2] E. Pośpiech, J. Karłowska-Pik, B. Ziemkiewicz et al., *Further evidence for population differences in the effect of DNA markers and gender on eye colour prediction in forensics*, International Journal of Legal Medicine, 130 (4), pp. 923-934 (2016)

Wybór zmiennych w klasyfikacji dla wielowymiarowych danych funkcjonalnych

**Tomasz Górecki,
Miroslaw Krzyśko,
Waldemar Wołyński**

Wydział Matematyki i Informatyki,
Uniwersytet im. Adama Mickiewicza w Poznaniu

Rozważać będziemy nową metodę wyboru zmiennych w zagadnieniu klasyfikacji obiektów opisanych przez wektorowy proces losowy postaci $\{\mathbf{X}(t) = (X_1(t), \dots, X_p(t))', t \in I\}$ (Ramsay & Silverman (2005), Horváth & Kokoszka (2012)). Omawiana metoda wyboru zmiennych jest metodą redukcji wymiarowości mającą na celu zastąpienie pierwotnego wektorowego procesu losowego $\mathbf{X}(t)$, przez proces losowy o możliwie niskim wymiarze $(X_{(1)}(t), \dots, X_{(d)}(t))'$ z zachowaniem porównywalnego poziomu błędu klasyfikacyjnego. Wykorzystywać będziemy różne algorytmy klasyfikacyjne dostosowane do wielowymiarowych danych funkcjonalnych. Zaproponowana metoda redukcji wymiarowości bazuje na funkcjonalnej odległości korelacyjnej (Székely & Rizzo (2009, 2012)) i jest modyfikacją procedury podanej w pracy Konga i innych (2015). Działanie zaproponowanej metody wyboru zmiennych zilustrowane zostanie z wykorzystaniem rzeczywistych zbiorów danych.

Literatura

- [1] L. Horváth, P. Kokoszka (2012), *Inference for Functional Data with Applications*, Springer. New York.
- [2] J. Kong, S. Wang, G. Wahba (2015), *Using distance covariance for improved variable selection with application to learning genetic risk models*, *Statistics in Medicine* 34, 1708–1720.
- [3] J.O. Ramsay, B.W. Silverman (2005), *Functional Data Analysis*, Second Edition, Springer. New York.
- [4] G.J. Székely, M.L. Rizzo (2009), *Brownian distance covariance*, *Annals*

of Applied Statistics 3(4), 1236–1265.

[5] G.J. Székely, M.L. Rizzo (2012), *On the uniqueness of distance covariance*, Statistical Probability Letters 82(12), 2278–2282.

Selekcja zmiennych w źle wyspecyfikowanym modelu binarnym

Mariusz Kubkowski, Jan Mielniczuk

Politechnika Warszawska, Instytut Podstaw Informatyki PAN

Rozważmy próbę losową $(X_i^{(n)}, Y_i^{(n)}) \in \mathbb{R}^{p_n} \times \{0, 1\}$ dla $i = 1, \dots, n$ taką, że $\mathbb{P}(Y_i^{(n)} = 1 | X_i^{(n)}) = q(X_i^{(n)})$, gdzie $q : \mathbb{R}^{p_n} \rightarrow (0, 1)$ i problem dopasowania modelu logistycznego $\mathbb{P}(Y^{(n)} = 1 | X^{(n)}) = (1 + \exp(-(\beta^T X)))^{-1}$ dla $\beta \in \mathbb{R}^{p_n}$. W przypadku, gdy model binarny został błędnie wyspecyfikowany, problem selekcji zmiennych może dotyczyć identyfikacji nośnika s_0^* wektora β_0^* , będącego parametrem rzutu Kullbacka-Leiblera modelu binarnego na rodzinę modeli logistycznych. W referacie zostanie omówione działanie i zgodność procedury dwukrokowej, w której w pierwszym kroku dokonywane jest wstępne odsianie zmiennych metodą Lasso i utworzenie ze zbioru tych zmiennych rodziny hierarchicznej, a w następnym kroku wybór zbioru predyktorów minimalizujących kryterium GIC (ang. *Generalized Information Criterion*) dla rodziny hierarchicznej utworzonej w pierwszym kroku.

Literatura

- [1] M. Kubkowski, J. Mielniczuk, *Active sets of predictors for misspecified logistic regression*, Statistics, 2017
- [2] M. Kubkowski, J. Mielniczuk, *Selection consistency of two-step selection for misspecified binary model*, manuskrypt, 2017

Dobór stałej ucinającej dla funkcji Hubera wykorzystywanej w odpornej estymacji parametrów

Agnieszka Kulawik

Instytut Matematyki, Uniwersytet Śląski w Katowicach

Stefan Zontek

Wydział Matematyki, Informatyki i
Ekonometrii, Uniwersytet Zielonogórski

Na przykładzie modelu przesunięcia przedstawiona zostanie pewna idea estymacji stałej ucinającej występującej w definicji funkcji Hubera. Można to uzyskać stosując metodę największej wiarogodności w nowo zdefiniowanym modelu, w którym stała ucinająca traktowana jest jako parametr zakłócający. Rozważania teoretyczne zostaną uzupełnione wynikami symulacji komputerowych.

Literatura

[1] P. J. Huber (1981), *Robust Statistics*, Wiley, New York

Unadjusted Langevin Monte Carlo jako optymalizacja wypukła

Szymon Majewski

Instytut Matematyczny PAN

W moim referacie przedstawię nowy sposób patrzenia na algorytmy typu *Unadjusted Langevin Monte Carlo* (ULA). Algorytmy te polegają na dyskretyzacji równania Langevina:

$$dX_t = -\nabla U dt + dW_t,$$

w celu uzyskania Łańcucha Markowa zbiegającego do rozkładu $\pi \propto e^{-U}$. Okazuje się, że można przedstawić algorytmy tego typu jako algorytmy optymalizujące pewien funkcjonal energii na przestrzeni miar probabilistycznych z metryką Wassersteina. Używając tej interpretacji i technik znanych z optymalizacji wypukłej, wyprowadzę ograniczenia nieasymptotyczne na zbieżność standardowego algorytmu ULA i wersji tego algorytmu korzystającej ze stochastycznych gradientów. Zaprezentuję również algorytmy próbkowania dla nieróżniczkowalnych potencjałów, oparte o wyżej wspomnianą reprezentację.

Zbieżność algorytmów stochastycznej aproxymacji dla funkcji niegładkich i niewypukłych

Błażej Miasojedow

Instytut Matematyczny PAN

Wiele współczesnych metod wnioskowania statystycznego sprowadza się do rozwiązania problemu optymalizacyjnego postaci

$$\hat{\theta} = \operatorname{argmin}_{\theta \in \Theta} \{ \ell(\theta) + p(\theta) \} , \quad (1)$$

gdzie ℓ jest gładką funkcją straty, natomiast p jest karą nieróżniczkowalną. W wielu przypadkach funkcje ℓ oraz p mogą być również niewypukłe. Jedną z popularniejszych z obecnie stosowanych metod znajdowania minimum (lokalnego) dla problemu (1) jest iteracyjny algorytm przybliżonego gradientu (*proximal gradient*) zdefiniowany przez

$$\theta_{k+1} = \operatorname{prox}_{\gamma_k p}(x_k - \gamma_k \nabla \ell(x_k)) ,$$

gdzie

$$\operatorname{prox}_{\gamma p}(\theta) = \operatorname{argmin}_{\vartheta} \left\{ p(\vartheta) + \frac{1}{2\gamma} \|\vartheta - \theta\|^2 \right\} ,$$

oraz γ_k jest ciągiem długości kroków. Podczas wykładu przedstawię wyniki dotyczące zbieżności algorytmu przybliżonego gradientu dla stochastycznej wersji algorytmu, w której gradient $\nabla \ell(\theta_k)$ jest zastąpiony przez estymator tej wielkości uzyskany za pomocą metod MCMC.

Badanie efektywności estymatorów jądrowych

Maciej Karczewski, Andrzej Michalski

Katedra Matematyki, Uniwersytet Przyrodniczy we Wrocławiu

Głównym celem tego artykułu jest porównanie efektywności zastosowanych estymatorów jądrowych w odniesieniu do empirycznej funkcji gęstości uzyskanej dla pewnych danych empirycznych (Michalski, 2016). Autorzy wzięli pod uwagę kilka rodzajów jąder: Gaussowskie, Epanechnikova, dwuwagowe i trójwagowe (por. Berlinet i Devroye, 1994) a także kilka metod doboru zakresu pasma: cross-validation, regułę kciuka (rule of thumb) oraz metodę Altmana-Legera (por. Feluch i Koronacki, 1992; Silverman, 1986; Givens i Hoeting, 2005; Wand i Jones, 1995). Dla porównania numerycznych efektów zastosowanych estymatorów jądrowych w relacji do empirycznej gęstości zastosowano metrykę Marczewskiego-Steinhaus (Marczewski i Steinhaus, 1958). Dla oceny różnic zastosowanych metod estymacji jądrowej uzyskane odległości zostały dodatkowo wykorzystane do zbudowania macierzy i dendrogramów podobieństwa rozważanych estymatorów. W pracy przedstawiono również wyniki badań symulacyjnych.

Literatura

- [1] A. Berlinet, L. Devroye, *A comparison of kernel density estimates*, Publications de l'Institut de Statistique de l'Université de Paris, vol. XXXVIII –Fascicule 3, 3–59, 1994.
- [2] W. Feluch, J. Koronacki, *A note on modified cross-validation in density estimation*, Computational Statistics & Data Analysis, 13, 143–151, 1992.
- [3] G. H. Givens, J. A. Hoeting, *Computational Statistics*, New York: Wiley & Sons, 2005.
- [4] E. Marczewski, H. Steinhaus, *On a certain distance of sets and the corresponding distance of functions*, Coll. Mathematicum 6, 319–327, 1958.
- [5] A. Michalski, *The use of kernel estimators to determine the distribution of groundwater level*, Meteorology Hydrology and Water Management, Vol.4, 1, 41–46, 2016.

[6] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, London: Chapman & Hall, 1986.

[7] M. P. Wand, M. C. Jones, *Kernel Smoothing*, London: Chapman & Hall, 1995.

Aproksymacja macierzy kowariancji za pomocą wstęgowych macierzy Toeplitza

Adam Mieldzioc

Uniwersytet Przyrodniczy w Poznaniu

W modelach wielowymiarowych często pojawia się problem estymacji nieznanych macierzy kowariancji o zadanej strukturze. Jedną z takich struktur stanowią wstęgowe macierze Toeplitza, dla których w zbiorze macierzy dodatnio określonych wyznaczmy najlepsze w sensie normy Frobeniusa przybliżenie macierzy kowariancyjnej. Cui et al. (2016) pokazali, że najlepszym przybliżeniem zadanej macierzy kowariancyjnej w zbiorze macierzy Toeplitza jest jej rzut na przestrzeń macierzy Toeplitza. Ponieważ jednak nie wszystkie rzuty zachowują własność nieujemnej określoności pokażemy, że najlepsze przybliżenie zadanej macierzy kowariancyjnej w zbiorze określonych nieujemnie macierzy Toeplitza charakteryzuje rzut na stożek wypukły generowany przez rozważane macierze (por. Ingram and Marsh, 1991). W referacie podana zostanie zarówno metodologia wyznaczania najlepszego przybliżenia jak i odpowiadający jej algorytm.

Prezentowane wyniki powstały we współpracy z K. Filipiak, A. Markiewiczem i A. Sawikowską.

Literatura

- [1] Cui, X., C. Li, J. Zhao, L. Zeng, D. Zhang, and J. Pan (2016), *Covariance structure regularization via Frobenius norm discrepancy*, Linear Algebra Appl. 510, 124–145.
- [2] Ingram, J.M. and M.M. Marsh (1991), *Projection onto convex cones in Hilbert space*, J. Approx. Theory 64, 343–350.

O pomiarze siły interakcji

Jan Mielniczuk

Instytut Podstaw Informatyki PAN

oraz

Wydział Matematyki i Nauk

Informacyjnych Politechniki Warszawskiej

W modelu logistycznym brak interakcji z reguły rozumie się jako brak zależności zmiany logarytmu szansy sukcesu, przy zmianie wartości jednego predyktora, od wartości pozostałych predyktorów. Taka definicja nie obejmuje wielu innych koncepcji interakcji: w szczególności w modelu logistycznym addytywnym mogą występować inaczej rozumiane interakcje. Wprowadzimy częściowy porządek na miarach interakcji: $I_2 \prec I_1$ jeśli $I_1 = 0$ implikuje $I_2 = 0$. Dla sytuacji dwóch predyktorów nominalnych X_1, X_2 o skończonej liczbie wartości i binarnej odpowiedzi Y zanalizujemy związki między interakcją logistyczną IL a informacją interakcyjną II zdefiniowaną jako

$$II(X_1; X_2; Y) := I[(X_1, X_2); Y] - I(X_1; Y) - I(X_2, Y), \quad (3.1)$$

gdzie $I(X_i; Y)$ jest informacją wzajemną między X_i i Y a $I[(X_1, X_2); Y]$ informacją wzajemną między parą (X_1, X_2) a Y . Wiadomo, że $II(X_1; X_2; Y)$ jest równa dywergencji Kullbacka-Leiblera między $P_{X_1, X_2, Y}$ a aproksymacją Kirkwooda \tilde{P}_K związaną z (niekoniecznie probabilistycznym) rozkładem masy

$$\tilde{p}_K(x_i, x_j, y_k) = \frac{p(x_i, x_j)p(x_i, y_k)p(x_j, y_k)}{p(x_i)p(x_j)p(y_k)}. \quad (3.2)$$

Niech η oznacza całkowitą masę \tilde{P}_K . Pokażemy, że jeśli $\eta \leq 1$ to $IL \prec II$. Podamy przykład rodziny rozkładów, dla której obie miary są równoważne. Zastosowania dotyczą wykrywania interakcji między polimorfizmami pojedynczych nukleotydów (SNP). Prace wykonane wspólnie z Pawłem Teisseyre i Marcinem Rdzanowskim.

Literatura

- [1] J. Mielniczuk, M. Rdzanowski, *Use of information measures and their approximations to detect predictive gene-gene interactions* , Entropy, 2017
- [2] J. Mielniczuk, P. Teisseyre, *A deeper look at two concepts of measuring gene-gene interactions: logistic regression and interaction information revisited*, Genetic Epidemiology, 2018

Aproksymacja macierzy kowariancji iloczynem Kroneckera z wykorzystaniem entropijnej funkcji straty

Monika Mokrzycka

Zespół Szkół Komunikacji w Poznaniu

Celem tego referatu jest określenie najlepszego przybliżenia symetrycznej, dodatnio określonej macierzy Ω stopnia n przez dodatnio określony iloczyn Kroneckera macierzy Ψ i Σ , gdzie Ψ i Σ są dowolnymi macierzami symetrycznymi lub jeden ze składników iloczynu Kroneckera, powiedzmy Ψ , ma korelacyjną strukturę kompletnej symetrii, $(1 - \rho)\mathbf{I} + \rho\mathbf{1}\mathbf{1}^\top$, lub autoregresji stopnia pierwszego, $\sum_{i=0} \rho^i (\mathbf{C}^i + \mathbf{C}^{i\top})$, gdzie $\mathbf{C} = (c_{ij})$ oraz $c_{ij} = 1$ dla $j - i = 1$ i 0 w pozostałych przypadkach. Najlepsza aproksymacja oznacza tutaj, że entropijna funkcja straty (por. [1])

$$f(\Omega, \Psi \otimes \Sigma) = \text{tr} \left[\Omega^{-1} (\Psi \otimes \Sigma) \right] - \ln \left| \Omega^{-1} (\Psi \otimes \Sigma) \right| - n$$

jest minimalizowana ze względu na $\Psi \otimes \Sigma$.

Przedstawione wyniki mogą być szeroko stosowane w statystyce wielowymiarowej, na przykład do regularyzacji struktury kowariancyjnej wielowymiarowych danych, do określenia estymatorów struktury kowariancji czy też do określenia odległości między hipotezą zerową i alternatywną w zagadnieniu testowania hipotez o strukturach kowariancyjnych.

Prezentowane wyniki powstały we współpracy z K. Filipiak, D. Klein i A. Markiewicz.

Literatura

[1] Lin L., Higham N.J., Pan J. (2014), *Covariance structure regularization via entropy loss function*, Computational Statistics and Data Analysis 72, 315–327.

Algorytmy pMCMC z repróbkowaniem poissonowskim

Wojciech Niemiro

Instytut Matematyki Stosowanej i Mechaniki,
Uniwersytet Warszawski

Algorytmy pMCMC (particle Markov chain Monte Carlo) zostały wprowadzone w pracy Andrieu i in. (2010). Łączą idee markowowskich algorytmów Monte Carlo (MCMC) i sekwencyjnych algorytmów Monte Carlo (SMC). Zasadnicza idea jest oparta na konstrukcji “filtru cząsteczkowego”, sekwencyjnego powtarzania kroków ważenia, repróbkowania i propagacji. Symuluje się ewolucję N “cząstek”, których trajektorie przybliżają rozkład filtracji dla ukrytego łańcucha Markowa. Większość znanych algorytmów wykorzystuje ustaloną liczbę N i repróbkowanie z rozkładu wielomianowego. Nasze algorytmy wykorzystują repróbkowanie z rozkładu Poissona i używają losowej, zmiennej liczby cząstek. Motywacje tej nowości są dwie. Po pierwsze, algorytmy z repróbkowaniem poissonowskim łatwiej jest implementować w sposób równoległy. Po drugie, są one lepiej dostosowane do sytuacji, gdy rozważamy ukryty proces Markowa z czasem ciągłym.

Referat prezentuje wyniki wspólnej pracy z Błażem Miasojedowem i Michałem Startkiem.

Literatura

- [1] C. Andrieu, A. Doucet, R. Holenstein, *Particle Markov chain Monte Carlo methods*, Journal of the Royal Statistical Society B
- [2] B. Miasojedow, W. Niemiro, M. Startek, *Particle MCMC algorithms with Poisson resampling: paralelization and continuous time models*, w przygotowaniu (wstępna wersja: [arXiv:1707.01660v1](https://arxiv.org/abs/1707.01660v1))

Optymalne planowanie pomiarów ruchu z wykorzystaniem programowania całkowitoliczbowego

Karol R. Opara

Instytut Badań Systemowych PAN

Natężenie i struktura ruchu drogowego stanowią podstawowe informacje wykorzystywane przez zarządy dróg do planowania remontów i zabiegów utrzymaniowych. Pomiary w skali całej sieci można prowadzić z wykorzystaniem przenośnych liczników przewożonych co jeden lub dwa tygodnie w kolejne miejsca. W referacie omówione zostanie zadanie przygotowania planu tego typu pomiarów. Można je sformalizować w postaci trzech powiązanych ze sobą problemów optymalizacyjnych. Ich trudność polega na konieczności uwzględnienia znacznej liczby wymagań prawnych, technicznych i organizacyjnych. Omówiony zostanie matematyczny model zagadnienia oraz metoda jego rozwiązania bazująca na programowaniu całkowitoliczbowym. Metoda ta pozwala na dokładne wyznaczenie optymalnego rozwiązania. Rozważania zostaną zilustrowane na przykładzie planowania pomiarów w 5000 punktów za pomocą 350 urządzeń w niemieckim kraju związkowym Badenia-Wirtembergia.

Literatura

- [1] K. Opara i J. Zieliński, *Quality assurance of road traffic data*, Materiały konferencyjne World Conference on Pavement and Asset Management, Baveno, Włochy, 12-16 czerwca 2017.
- [2] BMVI, *Richtlinien für die Straßenverkehrszählung im Jahre 2015 auf den Bundesfernstraßen*, Po niemiecku: Wytyczne dla pomiarów ruchu na drogach federalnych w 2015 roku. Niemieckie Federalne Ministerstwo Transportu i Infrastruktury Cyfrowej, 2014.
- [3] M. Mayer-Kreitz, R. Bettermann, M. Skakuj, G. Karzelek i H. Ziegler, *Zählstellenverwaltung und Ergebnisdarstellung temporärer Messungen des Straßenverkehrs mit der Online-Plattform VERA*, Po niemiecku:

Zarządzanie licznikami natężenia ruchu drogowego oraz udostępnianie wyników pomiarów z wykorzystaniem platformy online VERA, Straßenverkehrstechnik, vol. 59, nr 2, 2015.

Regresyjna charakteryzacja rozkładów Gamma i Kummera

Agnieszka Piliszek

Wydział Matematyki i Nauk Informatycznych PW

W 2016 Hamza i Vallois zauważyli następującą własność rozkładów gamma i Kummera: jeśli X i Y są niezależnymi zmiennymi losowymi o rozkładach gamma i Kummera, odpowiednio, z odpowiednimi parametrami, to zmienne losowe $U = \frac{Y}{1+X}$ i $V = X(1+U)$ są niezależne. W referacie opowiem o charakteryzacji regresyjnej rozkładów gamma i Kummera opartej na powyższej własności. Została ona opisana w [Piliszek, Wesołowski 2018]. Zwróć uwagę na użytą technikę zmiany miary, która pozwoliła na znaczne osłabienie założeń.

Wyniki wspólne z prof. J. Wesołowskim.

Literatura

- [1] A. Piliszek, J. Wesołowski, *Change of measure technique in characterizations of the Gamma and Kummer distributions*, Journal of Mathematical Analysis and Applications
- [2] M. Hamza, P. Vallois, *On Kummer's distribution of type two and a generalized beta distribution*, Statistics & Probability Letters

Poprawianie Lasso dla wyboru modelu i predykcji

Piotr Pokarowski

Uniwersytet Warszawski

Lasso, czyli estymator ℓ_1 -penalizowanej straty jest popularną metodą dopasowania oszczędnego/rzadkiego modelu do danych wysokiego wymiaru. Quasi-wypukłe penalizacje jak SCAD lub MCP dokładniej aproksymują ℓ_0 -penalizowaną stratę, czyli GIC i poprawiają nieuniknione obciążenie Lasso. Zaprezentuję alternatywną metodę SS (Screening–Selection) poprawiania Lasso dla ogólnej klasy modeli predykcyjnych zawierającej normalny model liniowy i model logistyczny. Metoda zaczyna się od Lasso dla ustalonej kary, porządkuje wartości bezwzględne współczynników Lasso i wybiera model z małej zagnieżdżonej rodziny. Pokażę górne oszacowanie na prawdopodobieństwo błędu selekcji SS oraz wyniki symulacji na syntetycznych i realnych danych. Z teoretycznego punktu widzenia SS jest bardziej konstruktywny od quasi-wypukłych penalizacji. W praktyce, w porównaniu do implementacji MCP, implementacja SS prowadzi do dokładniejszej predykcji lub mniejszego modelu. Praca wspólna z A. Prochenką, M. Frejem, W. Rejchelem i J. Mielniczukiem.

Selekcja cech w oparciu o kryterium GIC z karą LASSO w modelach wysokowymiarowych

Wojciech Rejchel

Uniwersytet Mikołaja Kopernika w Toruniu
Uniwersytet Warszawski

Wybór modelu (wybór cech istotnych) jest dużym wyzwaniem, zwłaszcza w przypadku wysokowymiarowym, to znaczy w sytuacji, gdy liczba obserwowanych cech jest (znacznie) większa niż liczność próby. W wielu problemach (w genetyce, biologii) znalezienie (małego) zbioru cech istotnych jest równie ważne (a nieraz bardziej) jak poprawna estymacja czy predykcja. Jednym z algorytmów, które próbują poradzić sobie z tym zagadnieniem, jest SOS [1], który oparty jest na minimalizacji funkcji wiarygodności z karą LASSO oraz kryterium GIC. W przypadku uogólnionych modeli liniowych procedura ta jest zgodna w wyborze cech istotnych. W referacie wykażę, że własność ta jest również spełniona w modelach obejmujących, między innymi, regresję kwantylową czy maszyny wektorów podpierających.

Literatura

[1] P. Pokarowski, J. Mielniczuk, *Combined l_1 and greedy l_0 penalized least squares for linear model selection*, Journal of Machine Learning Research

Estymacja współczynnika wpływu w modelowaniu przyczynowości

Krzysztof Rudaś

Politechnika Warszawska

Modelowanie przyczynowości staje się coraz bardziej istotnym problemem w zagadnieniach biznesowych i medycznych. Wyobraźmy sobie sytuację, w której chcemy zastosować terapię na pewnej grupie pacjentów. Naszym celem jest wyznaczenie jednostek, którym rzeczywiście ta terapia pomoże (są zdrowi i nie wyzdorwieliby bez niej). Podstawowym problemem jest dysponowanie tylko połową pożądaną informacją na temat pacjenta (jeśli został poddany terapii to nie mamy informacji o jego stanie w przypadku gdyby nie zastosowano leczenia).

Problem przyczynowości można przedstawić dzieląc dane na zbiór kontrolny i testowy, a następnie stosując założenia regresji liniowej do każdej z grup. W prezentacji zostaną przedstawione dwa powszechnie stosowane podejścia modelowe, a także trzecie będące próbą ich połączenia. Przedstawione zostaną ich podstawowe własności. Za pomocą symulacji pokazane zostanie, w jakich sytuacjach dany estymator daje najlepszą jakość predykcji.

Literatura

[1] M. Jaśkowski, S. Jaroszewicz, *Uplift modeling for clinical trial data.*, ICML 2012 Workshop on Machine Learning for Clinical Data Analysis, Edinburgh, Scotland, June 2012

Nierówności dla mieszanek uporządkowanych zmiennych losowych

Tomasz Rychlik

Instytut Matematyczny PAN

Porównujemy dwie mieszaniny dowolnego stochastycznie uporządkowanego zbioru dystrybuant względem dowolnie ustalonych rozkładów mieszających. Przy założeniu, że rozkład jednej mieszaniny jest znany, wyznaczamy optymalne dolne i górne oszacowania dystrybuanty drugiej mieszanki i wyznaczamy pojedyncze rodziny uporządkowanych dystrybuant, które osiągają te oszacowania jednostajnie w każdym punkcie osi rzeczywistej. Następnie przedstawiamy dokładne górne i dolne oszacowania różnic wartości oczekiwanych mieszanek wyrażone w różnych jednostkach skali. Ogólne wyniki są zilustrowane przykładami.

Analiza wydźwięku z wykorzystaniem sieci neuronowych i drzewiastej reprezentacji zdań

Norbert Ryciak

Wydział Matematyki i Nauk Informatycznych PW

Analiza wydźwięku (ang. sentiment analysis) jest jednym z głównych nurtów badawczych w obszarze przetwarzania języka naturalnego. Z reguły zagadnienie to formalizuje się jako problem klasyfikacji. Istotnym aspektem tego zadania jest reprezentacja danych - w rozważanym przypadku pojedynczego zdania. Podstawowe podejście polegające na wyreprezentowaniu zdania jako wektor cech (np. model bag-of-words) ma zasadniczą wadę - traci informację o strukturze zdania, która często ma istotne znaczenie. Dlatego w praktyce wykorzystuje się również modele uwzględniające strukturę tekstu. Przykładem są rekurencyjne sieci neuronowe, które przyjmują na wejściu ciągi i dokonują klasyfikacji na podstawie analizy sekwencji obiektów z uwzględnieniem kolejności (np. słów w zdaniu). Kolejnym krokiem w rozwoju metod analizy sentymentu jest wykorzystanie lingwistycznej struktury tekstu - reprezentacja zdań w postaci drzewa (zależnościowego lub składniowego). Okazuje się, że wykorzystanie informacji o strukturze zdań w ten sposób daje bardzo dobre wyniki. W referacie omówię pewien model sieci neuronowej działającej na strukturze drzewiastej - Tree Structured Long Short Term Memory Network i przedstawię wyniki otrzymane na przykładzie danych polskojęzycznych.

Literatura

[1] Kai Sheng Tai, Richard Socher, Christopher D. Manning (2015), *Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks*, CoRR

Nieparametryczna estymacja kwantyli

Alicja Jokił-Rokita¹, Agnieszka Siedlaczek²

¹Wydział Matematyki, Politechnika Wrocławska

²Instytut Matematyki i Informatyki, Uniwersytet Opolski

Problem nieparametrycznej estymacji kwantyli rozkładu pojawia się w wielu dziedzinach, np. w hydrologii, finansach. W literaturze statystycznej można znaleźć wiele podejść do tego problemu. Przegląd estymatorów, bazujących na tzw. *plotting position* można znaleźć w artykule Hyndmana i Fana (1996). Natomiast w serii artykułów Zielińskiego (np. Zieliński, 1999, 2006, 2012) rozważany jest problem wyznaczenia niezmienniczych estymatorów kwantyli, optymalnych według różnych kryteriów (zobacz również Rychlik, 2012).

W prezentacji przedstawimy wybrane nieparametryczne estymatory kwantyli oparte na nieparametrycznych estymatorach dystrybuanty rozkładu. Zaproponujemy ciągły i ściśle rosnący nieparametryczny estymator dystrybuanty rozkładu, którego odwrotność prowadzi do estymatora funkcji kwantylowej posiadającej wymieniane w literaturze pożądane własności oraz do zgodnych i asymptotycznie normalnych estymatorów kwantyli. Przedstawimy również wyniki symulacji komputerowych, których celem było porównanie błędów średniokwadratowych, w przypadku małych rozmiarów prób i różnych rozkładów, zaproponowanego estymatora i dziewięciu innych estymatorów. Zaprezentujemy również wyniki analizy danych rzeczywistych.

Literatura

- [1] Hyndman R. J. and Fan Y. (1996), *Sample quantiles in statistical packages*, The American Statistician, 50(4):361–365.
- [2] Rychlik T. (2012), *Prace Ryszarda Zielińskiego o nieparametrycznych estymatorach kwantyli i ich zastosowaniu w statystyce odpornej*, Mathematica Applicanda, 40(2):65–82.
- [3] Zieliński R. (1999), *Best equivariant nonparametric estimator of quantile*, Statistics and Probability Letters, 45:79–84.

- [4] Zieliński R. (2006), *Small-sample quantile estimators in a large non-parametric model*, Communications in Statistics – Theory and Methods, 35(7):1223–1241.
- [5] Zieliński R. (2012), *Optimal estimation of high quantiles in a large nonparametric model*, Applicationes Mathematicae, 39(2):137–142.

Metody funkcjonalnej analizy danych w detekcji uszkodzeń maszyn przemysłowych

Maria Skupień

Instytut Matematyki, Uniwersytet Pedagogiczny w Krakowie

Sygnały wibracyjne próbkowane z bardzo dużą częstotliwością (20 kHz) stanowią podstawowe źródło informacji o zachowaniu maszyny przemysłowej. Niemniej jednak duża wymiarowość zbioru danych i jego silne zaszumienie powodują trudności w wykrywaniu częstotliwości charakterystycznych dla określonego lokalnego uszkodzenia. Narzędzia analizy spektralnej, np. spektrogramy, nie są wystarczające by odróżnić IFB - informacyjne pasmo częstotliwości od pozostałej części sygnału. Stąd, jako odpowiednik dyskretnych metod, proponujemy funkcjonalne podejście w rozwiązaniu owego problemu. Na nowy zbiór danych składają się ciągłe funkcje losowe. Stosujemy techniki analizy danych funkcjonalnych (FDA), na przykład analizy funkcjonalnych składowych głównych (FPC) w celu redukcji wymiarowości. Następnie zaproponujemy nową miarę IFB w oparciu o współczynniki FPC i postaramy się wybrać, spośród zredukowanych modeli, najbardziej optymalny. Zweryfikujemy nasze podejście, wykorzystując rzeczywiste dane generowane przez łożyska kół maszyn przemysłowych. Zestaw danych udostępniony jest dzięki konglomeratowi miedzi KGHM.

Literatura

- [1] L. Horváth, P. Kokoszka, *Inference for Functional Data with Applications*, Springer-Verlag, New York etc., 2012.
- [2] J. Obuchowski, A. Wyłomańska, R. Zimroz, *Selection of informative frequency band in local damage detection in rotating machinery*, *Mechanical Systems and Signal Processing* 48 (2014) 138 - 152
- [3] J. Obuchowski, A. Wyłomańska, R. Zimroz, *Selection of informative frequency band based on the local maxima method in application to bearing diagnostics*, 12th International Technical Systems Degradation Conference

Diagonalne testy typu Walda w ogólnym modelu czynnikowym

Łukasz Smaga

Wydział Matematyki i Informatyki UAM

W pracy rozważamy zagadnienie testowania hipotez w ogólnym modelu czynnikowym, w którym nie zakłada się ani homoskedastyczności ani konkretnego rozkładu błędów losowych. Szczególnymi przypadkami rozpatrywanego modelu są jedno- i dwukierunkowa klasyfikacja czy też model analizy hierarchicznej. Do testowania hipotez, proponujemy statystyki testowe będące modyfikacjami statystyki testowej typu Walda, w której macierz wagową zastąpimy pewnymi macierzami diagonalnymi. Rozkłady statystyk testowych przy założeniu prawdziwości hipotezy zerowej przybliżane są za pomocą rozkładów granicznych oraz permutacyjnych. Udowodnimy zgodność nowych procedur testowych przy ustalonej hipotezie alternatywnej. Wyniki teoretyczne uzupełnimy poprzez rezultaty badań symulacyjnych. W badaniach tych porównujemy empiryczne moce nowych procedur testowych z mocami permutacyjnego testu typu Walda zaproponowanego w pracy Pauly i in. (2015).

Literatura

- [1] E.Y. Chung and J.P. Romano, *Exact and asymptotically robust permutation tests*, Annals of Statistics
- [2] F. Konietschke, A.C. Bathke, S.W. Harrar, M. Pauly, *Parametric and nonparametric bootstrap methods for general MANOVA*, Journal of Multivariate Analysis
- [3] M. Pauly, E. Brunner, F. Konietschke, *Asymptotic permutation tests in general factorial designs*, Journal of the Royal Statistical Society B
- [4] Ł. Smaga, *Diagonal and unscaled Wald-type tests in general factorial designs*, Electronic Journal of Statistics

DMRnet - pakiet środowiska R do wybierania zmiennych ciągłych i łączenia poziomów dla zmiennych jakościowych

Agnieszka Sołtys

Wydział Matematyki, Informatyki i Mechaniki
Uniwersytetu Warszawskiego

W referacie opowiem o uogólnieniu zwyczajnej selekcji zmiennych ciągłych w regresji liniowej i logistycznej na selekcję czynnikową, która umożliwia łączenie poziomów w faktorach. Zaprezentuję na przykładach główne funkcje pakietu środowiska R o nazwie DMRnet, który jest implementacją zarówno algorytmu SOSnet dla zmiennych ciągłych, jak i DMRnet dla zmiennych ciągłych i jakościowych. Pakiet DMRnet jest dostępny na CRAN pod adresem:

<https://cran.r-project.org/web/packages/DMRnet/index.html>.

SOSnet jest rozwinięciem algorytmu SOS (Screening, Ordering, Selection) opisanego w artykule P. Pokarowski et al., 2015, a DMRnet rozwinięciem algorytmu DMR (Delete or Merge Regressors) opisanego w artykule A. Maj-Kańska et al., 2015.

Literatura

[1] P. Pokarowski, J. Mielniczuk (2015), *Combined l_1 and greedy l_0 penalized least squares for linear model selection*, Journal of Machine Learning Research

[2] A. Maj-Kańska, P. Pokarowski, A. Prochenka (2015), *Delete or merge regressors for linear model selection*, Electronic Journal of Statistics

Test dopasowania rozkładu dla danych cenzurowanych bazujący na współczynniku maksymalnej korelacji

Ewa Strzałkowska-Kominiak

Politechnika Krakowska, Instytut Teleinformatyki

Współczynnik maksymalnej korelacji ρ^+ to górne ograniczenie Frecheta-Hoeffdinga współczynnika korelacji bazujące na dystrybuancie $H^+(x, y) = \min(F_1(x), F_2(y))$. Współczynnik maksymalnej korelacji stanowi miarę zgodności dwóch rozkładów, o dystrubuantach F_1 i F_2 , ponieważ $\rho^+ = 1$ wtedy i tylko wtedy, gdy $F_1 = F_2$ z dokładnością do parametrów skali i położenia i był już stosowany w tematyce testowania przez Fortiana i Grané (2003). W moim referacie przedstawię test zgodności bazujący na ρ^+ w kontekście danych cenzurowanych. Udowodnię asymptotyczne własności statystyki testowej. Dodatkowo, zaprezentuję wyniki symulacji komputerowych dotyczących mocy testu oraz jego porównania z najbardziej popularnym testem dla tego typu danych z Akritas (1988).

Literatura

- [1] Akritas, M. G. (1988), *Pearson-Type Goodness-of-Fit test: The Univariate Case*, Journal of the American Statistical Association 83, 222-230.
- [2] Fortiana, J., Grané (2003), *Goodness-of-Fit Tests Based on Maximum Correlations and Their Orthogonal Decompositions*, J. Royal Statistical Society. Series B 65, 115-126.

Badanie wpływu sposobu wyznaczania dystrybuanty doświadczałnej na moc testu Kołmogorowa-Smirnowa

Piotr Sulewski¹

Instytut Matematyki, Akademia Pomorska, Słupsk

W literaturze statystycznej istnieją różne sposoby na wyznaczanie dystrybuanty doświadczałnej. Symulacje komputerowe pokazały, że sposób liczenia dystrybuanty doświadczałnej ma istotny wpływ na moc testu Kołmogorowa-Smirnowa w wariancie Lillieforsa. Rozkładem testowym został złożony rozkład normalny $CND(a_1, b_1, a_2, b_2, \omega)$, gdzie a_1, a_2 są parametrami położenia, b_1, b_2 są parametrami skali, ω jest parametrem udziału. Moc testu wyznaczono przy poziomie istotności $\alpha = 0.1$ i liczebności próby $n = 10, 12, \dots, 30$.

Symulacje Monte Carlo zostały przeprowadzone dla czterech grup wartości parametrów: $ZRN(0, 1, 2, 3, 0.7)$ - skośność i eksces kurtoza dodatnie, $ZRN(0, 1, -2, 3, 0.7)$ - skośność ujemna, $ZRN(0, 1, 0, 3, 0.7)$ - eksces kurtoza dodatnia, $ZRN(0, 1, 4, 1, 0.5)$ - eksces kurtoza ujemna. W artykule zaproponowano miarę podobieństwa między rozkładem normalnym a złożonym rozkładem normalnym. Artykuł pokazuje, że sposób liczenia dystrybuanty empirycznej zależy od charakteru rozkładu nienormalnego, który chce uchodzić za normalny.

Literatura

- [1] N. Smirnov (1948), *Table for estimating the goodness of fit of empirical distributions*, Annals of Mathematical Statistics 19:279-281.
- [2] Z. Drezner, O. Turel, D. Zerom (2009), *A modified Kolmogorov-Smirnov test for normality*, MPRA Paper No. 14385.
- [3] N.M. Razali, Y.B. Wah (2011), *Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests*, Journal of Statistical Modeling and Analytics 2(1):21-33.

¹Podziękowania dla Prof. A. Drapelli za pomoc przy powstaniu tego artykułu.

Optymalne oszacowania całkowitego czasu trwania testu dla rozkładów z rosnącą uogólnioną intensywnością awarii

Maria Szpak

Uniwersytet Marii Curie-Skłodowskiej w Lublinie

Załóżmy, że w pewnym eksperymencie mamy n identycznych elementów, gdzie czas życia każdego z nich jest opisany przez niezależne zmienne losowe X_1, X_2, \dots, X_n o jednakowym rozkładzie z dystrybuantą F , wartością oczekiwaną μ i wariancją σ^2 . Przy prowadzeniu takiego testu w pewnym momencie eksperymentator może być zainteresowany w oszacowaniu tego, jak długo ten test będzie trwał, ponieważ z każdą jednostką czasu pracy dowolnego elementu mogą być związane jakieś koszty (albo zyski) proporcjonalne do czasu życia tego elementu. Czasy kolejnych awarii są opisane przez statystyki porządkowe $X_{1:n} \leq \dots \leq X_{n:n}$ z próby (X_1, \dots, X_n) . Po i -tej awarii w czasie $X_{i:n}$ w eksperymencie zostanie $n - i$ elementów, które będą działać do momentu kolejnej awarii $X_{i+1:n}$, czyli przez $X_{i+1:n} - X_{i:n}$ jednostek czasu. Dlatego dla pewnego $1 \leq r < n$ określamy całkowity czas testu po r -tej awarii przez

$$T_{r,n} = \sum_{i=r+1}^n (n - i + 1)(X_{i:n} - X_{i-1:n}) = \sum_{i=r+1}^n X_{i:n} - (n - r)X_{r:n}$$

Poszukujemy optymalnych oszacowań dla wartości oczekiwanej $T_{r,n}$ dla pewnego $1 \leq r < n$ w ograniczonych rodzinach rozkładów czasu życia. W poprzednich pracach rozważających ten problem zostały wyprowadzone optymalne oszacowania dla całkowitego czasu trwania testu w klasie wszystkich rozkładów, a także w klasie rozkładów z malejącą uogólnioną intensywnością awarii względem uogólnionych rozkładów Pareto. W referacie przedstawimy wyniki dla rozkładów z **rosnącą** uogólnioną intensywnością awarii, w szczególności oszacowania dla rozkładów z **rosnącą gęstością** i dla rozkładów z **rosnącą intensywnością awarii**. Oprócz wyników teoretycznych przedstawimy przykładowe wartości liczbowe uży-

skanych oszacowań i porównamy je z odpowiednimi oszacowaniami dla rozkładów z malejącą uogólnioną intensywnością awarii.

Zastosowanie funkcji intensywności starzenia

Magdalena Szymkowiak

Politechnika Poznańska

Praca poświęcona jest charakteryzacjom jednowymiarowych nieujemnych zmiennych losowych absolutnie ciągłych. W teorii niezawodności zmienne te wykorzystywane są do modelowania czasu życia elementów i układów. Omówione w pracy rozkłady, związane z rozkładem Weibulla, można łatwiej charakteryzować za pomocą intensywności starzenia niż intensywności uszkodzeń.

Ponadto w pracy pokazano zastosowanie intensywności starzenia dla danych generowanych oraz danych rzeczywistych pełnych i uciętych.

Literatura

- [1] D'Agostino, R.B., Stephens, M.A. (1986), *Goodness-of-fit techniques*, Marcel Dekker, New York.
- [2] Murthy, D. N. P., Xie, M., Jiang, R. (2004), *Weibull models*, Hoboken: John Wiley & Sons, Inc.
- [3] Szymkowiak, M., *Characterizations of distributions through aging intensity*, IEEE Transactions on Reliability (after the first revision, submitted 21.09.2017)

Selekcja zmiennych uwzględniająca koszty w klasyfikacji wieloetykietowej

Paweł Teisseyre

Instytut Podstaw Informatyki PAN

Wybór zmiennych objaśniających jest istotnym problemem w zadaniu klasyfikacji wieloetykietowej. W ostatnich latach zaproponowano wiele metod, jednak istniejące rozwiązania zakładają, że koszty związane z uzyskaniem wartości poszczególnych zmiennych są takie same. W pewnych sytuacjach to założenie może być nieadekwatne. Na przykład w diagnostyce medycznej uzyskanie wartości zmiennych może się wiązać z wykonaniem bardzo drogich badań diagnostycznych. Wówczas może się okazać że lepiej wybrać zmienne o nieco gorszej mocy predykcyjnej, ale akceptowalnych kosztach. W referacie przedstawię nową metodę selekcji zmiennych, która uwzględnia informację o ich kosztach. Zaproponowane podejście opiera się o wykorzystanie łańcuchów klasyfikatorów oraz penalizowanej regresji logistycznej, z odpowiednio dobraną karą, zależną od kosztów zmiennych. Pokażę również adaptacyjną wersję algorytmu, w której kary związane z poszczególnymi zmiennymi zmieniają się podczas dopasowania kolejnych modeli w łańcuchu. Eksperymenty przeprowadzone na rzeczywistych danych pokazują, że obie metody dają obiecujące wyniki.

Literatura

- [1] P. Teisseyre, *CCnet: Joint multi-label classification and feature selection using classifier chains and elastic net regularization*, Neurocomputing, vol. 235, strony 98-111, 2017.
- [2] P. Teisseyre, *Cost-sensitive classifier chains: selecting low-cost features in multi-label classification*, manuskrypt, 2017.

Plany optymalne dla predykcji w modelach RCR z błędami heteroskedastycznymi

Mateusz Wilk^{1,2}, Aleksander Zaigrajew¹

¹WMIi UMK, ²IP PAN

Rozważamy model regresji liniowej z losowymi parametrami i błędami heteroskedastycznymi, tzn. model opisany równaniem

$$y(x) = \beta_1 + \beta_2 x + e(x), \quad x \in [a, b] \cup \chi,$$

gdzie $[a, b] \subset \mathbb{R}$ jest obszarem planowania, χ jest obszarem predykcji takim, że $[a, b] \cap \chi = \emptyset$, $\beta = [\beta_1 \ \beta_2]^T$ jest wektorem losowym nieskorelowanym z $e(x)$, $\mathbb{E}[e(x)] = 0$ oraz $\text{cov}[e(x)] \propto g(x)$, gdzie $g : \mathbb{R} \rightarrow \mathbb{R}$ jest ustaloną funkcją dodatnią. Zakładamy również, że będziemy obserwować generowane przez ten model, nieskorelowane ze sobą wartości $y(x_i)$ w różnych od siebie punktach $x_i \in [a, b]$, $i = 1, \dots, n$.

Badamy problem wyznaczenia planów optymalnych dla predykcji wartości $y(x)$ w punktach $x \in \chi$. Każdy ciągły k -punktowy plan ζ jest dyskretną miarą probabilistyczną przyjmującą wartości $\omega_i \geq 0$ w punktach $x_i^\zeta \in [a, b]$, $i = 1, \dots, k$. Rozpatrujemy przy tym stochastyczne kryteria optymalności opisywane przy pomocy poniższych funkcji:

- $\Phi_1(\zeta) = \mathbb{P}(\max_{x \in \chi} |\hat{y}(x) - y(x)| < \epsilon)$,
- $\Phi_2(\zeta) = \min_{x \in \chi} \mathbb{P}(|\hat{y}(x) - y(x)| < \epsilon)$,
- $\Phi_3(\zeta) = \mathbb{P}(\int_{\chi} (\hat{y}(x) - y(x))^2 dx < \epsilon)$,

gdzie przez $\hat{y}(x)$ oznaczamy predykcje. Mówimy, że plan ζ jest Φ_i -optymalny, jeżeli maksymalizuje on funkcję Φ_i dla dowolnego $\epsilon > 0$, $i = 1, 2, 3$. W trakcie referatu przyglądamy się bliżej ograniczeniom, jakie muszą zostać nałożone na funkcję g , aby utrzymane zostały klasyczne rezultaty.

Literatura

- [1] Chen, R.-B., Wong, W. K., & Li, K.-Y. (2008), *Optimal minimax designs over a prespecified interval in a heteroscedastic polynomial model*, *Statistics & Probability Letters*, 78(13), 1914-1921.
- [2] Liski, E.P., Mandal, N.K., Shah, K.R., Sinha, B.K. (2002), *Topics in Optimal Design.*, *Lecture Notes in Statistics*: 163. Springer-Verlag, New York.

Praktyczne aspekty konstrukcji pasów ufności w problemie Spektora-Lorda-Willisa

**Bogdan Ćmiel,
Zbigniew Szkutnik
Jakub Wojdyła**

Wydział Matematyki Stosowanej AGH w Krakowie

Omówione zostaną praktyczne aspekty konstrukcji nieparametrycznych pasów ufności dla intensywności procesu Poissona w stereologicznym problemie Spektora-Lorda-Willisa, który z matematycznego punktu widzenia jest problemem odwrotnym. Środek analizowanych pasów ufności wyznaczony jest przez odpowiedni estymator typu jądowego. Pokazane zostanie, że procedura wyboru parametru wygładzania metodą Goldenshlugera-Lepkiego (Goldenshluger i Lepski, 2011) gwarantuje adaptacyjność estymatora na rozpatrywanej rodzinie klas typu Sobolewa. Zaprezentowane zostaną wyniki badań symulacyjnych, w których porównano (w kontekście konstrukcji pasów ufności) metodę Goldenshlugera-Lepkiego wyboru parametru wygładzania z procedurą opartą o minimalizację ryzyka empirycznego estymatora. Ponadto, omówione zostaną inne, istotne z praktycznego punktu widzenia, aspekty związane z konstrukcją pasów w skończonych próbach.

Literatura

[1] A. Goldenshluger and O. Lepski, *Bandwidth selection in kernel density estimation: oracle inequalities and adaptive minimax optimality*, The Annals of Statistics

Adaptacyjny test permutacyjny w problemie k -prób dla danych cenzurowanych

Grzegorz Wyłupek

Instytut Matematyczny, Uniwersytet Wrocławski

Rozważamy klasyczny problem k -prób, $k \geq 2$, dla danych cenzurowanych. Na podstawie zbioru niekompletnych obserwacji

$$(X_{li}, \Delta_{li}) = (\min\{X_{li}^o, U_{li}\}, 1(X_{li}^o \leq U_{li})), \quad i = 1, \dots, n_l, \quad l = 1, \dots, k,$$

gdzie X_{li}^o są niezależnymi czasami przeżycia, a U_{li} niezależnymi czasami cenzurowania pochodzącymi z populacji o ciągłych dystrybuantach F_l i G_l , odpowiednio, będziemy testować

$$\mathcal{H} : F_1(x) = F_2(x) = \dots = F_k, \quad \text{dla wszystkich } x \geq 0,$$

$$\mathcal{A} : F_r(x) \neq F_s(x), \quad \text{dla pewnych } 1 \leq r < s \leq k, \quad \text{oraz pewnego } x \geq 0,$$

w obecności nieskończenie wymiarowych parametrów zakłócających G_1, \dots, G_k .

Historia problemu liczy już blisko pół wieku. A literatura na ten temat jest niezwykle bogata. W sytuacji, gdy $k > 2$, na uwagę zasługują, między innymi, następujące artykuły: Breslow (1970), Martínez-Cambor (2010), Chen i inni (2016), Liu i Yin (2017).

W referacie przedstawimy nowe rozwiązanie powyższego problemu, przedyskutujemy uzyskane wyniki teoretyczne oraz zaprezentujemy reprezentatywne wyniki przeprowadzonych badań symulacyjnych, w których porównano nowe rozwiązanie z popularnymi procedurami.

Literatura

- [1] Breslow, N. (1970), *A generalized Kruskal-Wallis test for comparing K samples subject to unequal patterns of censorship*, *Biometrika*, 57, 579-594.
- [2] Chen, Z., Huang, H., Qiu, P. (2016), *Comparison of multiple hazard rate functions*, *Biometrics*, 72, 39-45.

- [3] Liu, Y., Yin, G. (2017), *Partitioned log-rank tests for the overall homogeneity of hazard rate functions*, Lifetime Data Analysis, 23, 400-425.
- [4] Martínez-Cambor, P. (2010), *Comparing k -independent and right censored samples based on the likelihood ratio*, Computational Statistics, 25, 363-374.

O prawdopodobieństwach inwersji większości w systemach z dwuetapowym głosowaniem

Aleksander Zaigrajew

Nicolaus Copernicus University, Toruń

Serguei Kaniovski

Austrian Institute of Economic Research (WIFO), Vienna

Systemy z niebezpośrednim głosowaniem, takie jak np Kolegium Wyborcze USA, są podatne na inwersję większości, czyli na sytuacje, w których wynik wyborów, określony większością głosów elektorów, nie jest poparty większością głosów ogółu wyborców. Uzyskaliśmy prawdopodobieństwo inwersji większości w modelu dwuetapowego głosowania z trzema stanami o różnej liczebności oraz ustaliliśmy dolne i górne oszacowanie tego prawdopodobieństwa w ogólnym przypadku więcej niż trzech stanów. Okazuje się, że prawdopodobieństwo inwersji większości jest Schur-wypukłą funkcją względem liczebności stanów, co oznacza, że im bardziej są nierówne stany, tym większe jest prawdopodobieństwo wystąpienia inwersji. To, czy własność Schur-wypukłości ma miejsce też w ogólnym przypadku, pozostaje na razie przypuszczeniem.

Literatura

- [1] M. R. Feix, D. Lepelley, V. Merlin, J.-L. Rouet (2004), *The probability of conflicts in a U.S. presidential type election*, Economic Theory, 23, 227-257.
- [2] M. R. Feix, D. Lepelley, V. Merlin, J.-L. Rouet (2011), *Three ways to compute accurately the probability of the referendum paradox*, Mathematical Social Sciences 62, 28-33.
- [3] K. May (1948), *Probabilities of certain election results*, Amer. Math. Monthly, 55, 203-209.

[4] O. de Mouzon, T. Laurent, M. Le Breton, D. Lepelley (2016), *The IAC probability of a divided verdict in a simple U.S. presidential type election*, Toulouse School of Economics. Working paper 16 (671).

Estymacja i testowanie hipotez w wielowymiarowych modelach normalnych

Roman Zmyślony

Uniwersytet Zielonogórski

Własności algebr Jordana (Jordan 1934) lub kwadratowych podprzestrzeni (Seely 1977, Zmyślony 1980) będą omawiane z punktu widzenia statystycznych zastosowań do wnioskowania w jedno- i wielowymiarowych modelach normalnych. Zostaną zaprezentowane zarówno estymacja, jak i testowanie hipotez. Szczególne przypadki losowych efektów w modelu wielowymiarowym z blokowo-symetryczną macierzą kowariancji (BCS) dla podwójnie wielowymiarowych obserwacji (m wymiarowy wektor obserwacji ponownie mierzony ze względu u lokacji bądź punktów czasowych), co jest wielowymiarowym uogólnieniem symetrycznej macierzy kowariancji dla wielowymiarowych danych (Rao (1945, 1953)).

Literatura

- [1] Arnold, S. F., *Linear Models with Exchangeably Distributed Errors*, Journal of the American Statistical Association Volume 74, 1979 - Issue 365. 1979.
- [2] Covas R., Mexia J.T. , Zmyślony R., *Lattices in Jordan Algebra*, Linear Algebra and Its Applications 432 (2010), 2679-2690. 2010
- [3] Drygas, H., *The Coordinate-Free Approach to Gauss-Markov Estimation*, Berlin, Heidelberg: Springer. 1970
- [4] Jordan, P., Neumann, von, J. and Wigner, E., *On an algebraic generalization of the quantum mechanical formalism*, The Annals of Mathematics, 35(1), 29-64. 1934
- [5] Rao, C. R., *Familial correlations or the multivariate generalizations of the intraclass correlation*, Current Science Volume 14, Pages 66-67. 1945
- [6] Rao, C. R., *Discriminant functions for genetic differentiation and selection*, Sankhyā Volume 12, Pages 229-246. 1953
- [7] Roy, A., Leiva, R., Ivan Žežula and Daniel Klein, *Testing the equality of*

mean vectors for paired doubly multivariate observations in blocked compound symmetric covariance matrix setup, Journal of Multivariate Analysis, 137, 50-60. 2014

Lista uczestników

1. **dr Magdalena Alama-Bućko**
Uniwersytet Technologiczno-Przyrodniczy w Bydgoszczy
mbucko@utp.edu.pl
2. **prof. Tadeusz Bednarski**
Uniwersytet Wrocławski
tadeusz.bednarski@uwr.edu.pl
3. **dr hab. Mariusz Bieniek**
Instytut Matematyki UMCS
mariusz.bieniek@umcs.lublin.pl
4. **dr hab. Małgorzata Bogdan**
Uniwersytet Wrocławski
malgorzata.bogdan@uwr.edu.pl
5. **dr Jacek Bojarski**
Wydział Matematyki, Informatyki i Ekonometrii, Uniwersytet Zielonogórski
j.bojarski@wmie.uz.zgora.pl
6. **prof. Tadeusz Caliński**
Uniwersytet Przyrodniczy w Poznaniu
calinski@up.poznan.pl
7. **dr Bogdan Ćmiel**
Akademia Górniczo-Hutnicza
cmielbog@gmail.com
8. **dr hab. Łukasz Dębowski**
Instytut Podstaw Informatyki PAN
ldebowsk@ipipan.waw.pl

9. **dr Anna Dudek**
Akademia Górniczo-Hutnicza
aedudek@agh.edu.pl
10. **dr hab. Katarzyna Filipiak**
Politechnika Poznańska
katarzyna.filipiak@put.poznan.pl
11. **dr hab. Konrad Furmańczyk**
Wydział Zastosowań Informatyki i Matematyki SGGW
konfur@wp.pl
12. **dr Barbara Gołubowska**
Instytut Podstawowych Problemów Techniki PAN
bgołub@ippt.pan.pl
13. **dr hab. Tomasz Górecki**
Wydział Matematyki i Informatyki, Uniwersytet im. Adama Mickiewicza
w Poznaniu
tomasz.gorecki@amu.edu.pl
14. **dr Agnieszka Goroncy**
Wydział Matematyki i Informatyki, Uniwersytet Mikołaja Kopernika w
Toruniu
gemini@mat.umk.pl
15. **dr Mariusz Grządziel**
Uniwersytet Przyrodniczy we Wrocławiu
mariusz.grzadzziel@upwr.edu.pl
16. **dr hab. Przemysław Grzegorzewski**
Wydział Matematyki i Nauk Informatycznych Politechniki Warszawskiej
pgrzeg@mini.pw.edu.pl
17. **prof. Tadeusz Inglot**
Wydział Matematyki Politechniki Wrocławskiej
Tadeusz.Inglot@pwr.edu.pl
18. **dr Krzysztof Jasiński**
Uniwersytet Mikołaja Kopernika w Toruniu, Wydział Matematyki i Infor-
matyki
krzys@mat.umk.pl
19. **dr Joanna Karłowska-Pik**
Wydział Matematyki i Informatyki, Uniwersytet Mikołaja Kopernika w
Toruniu
joanka@mat.umk.pl

20. **mgr Katarzyna Kobylańska**
Uniwersytet Warszawski
kz277838@students.mimuw.edu.pl
21. **dr hab. Wasyl Kowalczuk**
Instytut Podstawowych Problemów Techniki PAN
vkoval@ippt.pan.pl
22. **prof. Mirosław Krzyśko**
Wydział Matematyki i Informatyki UAM
mkrzysko@amu.edu.pl
23. **mgr Mariusz Kubkowski**
Politechnika Warszawska
M.Kubkowski@mini.pw.edu.pl
24. **dr Agnieszka Kulawik**
Instytut Matematyki, Uniwersytet Śląski w Katowicach
Agnieszka.Kulawik@us.edu.pl
25. **prof. Teresa Ledwina**
Instytut Matematyczny PAN
ledwina@impan.pl
26. **dr Piotr Majerski**
AGH Akademia Górniczo-Hutnicza w Krakowie
majerski@agh.edu.pl
27. **mgr Szymon Majewski**
Instytut Matematyczny PAN
smajewski@impan.pl
28. **dr hab. Marek Męczarski**
Instytut Ekonometrii, Szkoła Główna Handlowa w Warszawie
mecz@sgh.waw.pl
29. **dr Błażej Miasojedow**
Instytut Matematyczny PAN oraz Instytut Matematyki Stosowanej i Me-
chaniki Uniwersytetu Warszawskiego
bmiasojedow@impan.pl
30. **dr hab. Andrzej Michalski**
Katedra Matematyki, Uniwersytet Przyrodniczy we Wrocławiu
apm.mich@gmail.com
31. **mgr Adam Mieldzioc**
Uniwersytet Przyrodniczy w Poznaniu
adam.mieldzioc@mail.up.poznan.pl

32. **prof. Jan Mielniczuk**
Instytut Podstaw Informatyki PAN i Wydział Matematyki i Nauk Informatycznych PW
miel@ipipan.waw.pl
33. **mgr Monika Mokrzycka**
Zespół Szkół Komunikacji w Poznaniu
monika_mokrzycka@poczta.fm
34. **prof. Wojciech Niemi**
Wydział Matematyki Stosowanej i Mechaniki, Uniwersytet Warszawski
wniemi@gmail.com
35. **dr Karol Opara**
Instytut Badań Systemowych PAN
karol.opara@ibspan.waw.pl
36. **mgr Luiza Pańczyk**
Instytut Matematyki UMCS
luiza.panczyk@o2.pl
37. **mgr Agnieszka Piliszek**
Wydział Matematyki i Nauk Informatycznych PW
a.piliszek@mini.pw.edu.pl
38. **dr hab. Piotr Pokarowski**
Uniwersytet Warszawski
pokar@mimuw.edu.pl
39. **mgr Łukasz Rajkowski**
Wydział Matematyki, Informatyki i Mechaniki, Uniwersytet Warszawski
l.rajkowski@mimuw.edu.pl
40. **dr Wojciech Rejchel**
Uniwersytet Warszawski
wrejchel@gmail.com
41. **dr Ewa Eliza Rożko**
Instytut Podstawowych Problemów Techniki PAN
erozko@ipt.pan.pl
42. **mgr Krzysztof Rudaś**
Politechnika Warszawska
krzysztof.rudas@gmail.com
43. **prof. Tomasz Rychlik**
Instytut Matematyczny PAN
trychlik@impan.pl

44. **mgr Norbert Ryciak**
Politechnika Warszawska
n.ryciak@mini.pw.edu.pl
45. **mgr Agnieszka Siedlaczek**
Uniwersytet Opolski
asiedlaczek@math.uni.opole.pl
46. **mgr Maria Skupień**
Uniwersytet Pedagogiczny w Krakowie
marysia.skupien@gmail.com
47. **dr Łukasz Smaga**
Wydział Matematyki i Informatyki, Uniwersytet im. Adama Mickiewicza
ls@amu.edu.pl
48. **dr Agnieszka Sołtys**
Uniwersytet Warszawski
ap220756@mimuw.edu.pl
49. **dr Ewa Strzałkowska-Kominiak**
Politechnika Krakowska, Instytut Teleinformatyki
estrzalkowska-kominiak@pk.edu.pl
50. **dr Piotr Sulewski**
Instytut Matematyki, Akademia Pomorska. Słupsk
piotr.sulewski@apsl.edu.pl
51. **mgr Maria Szpak**
Instytut Matematyki UMCS
shpak.maria@gmail.com
52. **dr Magdalena Szymkowiak**
Politechnika Poznańska
magdalena.szymkowiak@put.poznan.pl
53. **dr Paweł Teisseyre**
Instytut Podstaw Informatyki PAN
teisseyrep@ipipan.waw.pl
54. **mgr Mateusz Wilk**
Wydział Matematyki i Informatyki UMK
mtszwlk@mat.umk.pl
55. **mgr Jakub Wojdyła**
Wydział Matematyki Stosowanej AGH
jwojdyła@agh.edu.pl

56. **dr Grzegorz Wyłupek**
Instytut Matematyczny, Uniwersytet Wrocławski
wylupek@math.uni.wroc.pl
57. **dr hab. Aleksander Zaigrajew**
Uniwersytet M.Kopernika w Toruniu
alzaig@mat.umk.pl
58. **prof. Roman Zmyślony**
Uniwersytet Zielonogórski
rzmyslony@op.pl