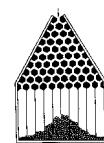


ROBUST 2024
BARDĚJOV
8. – 13. září 2024



Sborník abstraktů



Organizátoři ROBUSTu 2024 by tímto chtěli poděkovat všem, kteří pomohli při přípravě a realizaci celé akce. Velmi důležitá pro nás byla morální i věcná podpora ČMS JČMF, ČStS, KPMS MFF UK, RSJ a SSDS. Neméně důležitá byla podpora řady pracovišť, která především umožnila účast mnoha studentů, jakož udělení cen za nejlepší vystoupení studentů a/nebo doktorandů.

Nejdůležitější pro zdárný průběh ROBUSTu však bylo úsilí všech účastníků, které věnovali jak přípravě a prezentaci svých vystoupení, tak vytvoření skutečně „robustní atmosféry“. Všem děkují a na setkání na příštím Robustu se těší **organizátoři**.

RSJ



Úvodem několik historických mezníků z kroniky města Bardějov

Bardějov Nejstarší písemnou zmínkou o Bardějovu je zápis v Ipatijevské kronice z roku 1241. Již v té době se jednalo o důležité město na významné obchodní cestě mezi Černým a Baltickým mořem. Není proto divu, že již v roce 1320 dostalo město rozsáhlá privilegia a příkaz vybudovat hradby. Od roku 1365 zde sídlil i kat. Svobodným královským městem se Bardějov stává v roce 1376. Město postupně rostlo, a podle soupisu daní z roku 1437 bylo uvnitř hradeb 517 domů a žilo zde více než 3000 svobodných ončanů, 64 řemesel, 51 cechů a 146 řemeslnických mistrů. V 16. století zde vznikla humanistická škola, která se stala centrem kultury a vzdělanosti pro široké okolí. V 17. století lze díky válkám a opakujícím se požárům vysledovat postupný úpadek města. Částečný rozkvet pak lze vypočítat až konce 19. století. Na druhé straně, slabý průmysl a nevyvinuté polnohospodářství vedlo na přelomu 19. a 20. století k masivní vlně vystěhovalectví, především do Ameriky. Největší rozkvet nastal po druhé světové válce, kdy vznikla řada nových továren, jež do Bardějova přilákaly mnoho lidí. Po roce 1989 bohužel mnohé průmyslové podniky zanikly. Historické jádro města bylo již v roce 1950 prohlášeno památkovou rezervací a postupně rekonstruováno. V roce 2000 bylo historické jádro a soubor budov Synagogy zapsáno do Světového kulturního dědictví UNESCO.



Slovensko je krajina, která je známá nejenom krásnou přírodou, ale i významnými kulturními památkami. Mezi ně bezpochyby patří východoslovenské dřevěné kostely. Z urbanistického hlediska tvoří nádherné lokální dominanty krajiny. Již v roce 1967 bylo 27 z nich prohlášeno Národní kulturní památkou. V roce 2008 bylo osm z nich zapsáno také do seznamu UNESCO. Ve středu 11. září 2024 se proto bude konat výlet po **dřevěných chrámech východoslovenských Karpat**, jež jsou zařazeny do světového dědictví UNESCO.

Vydáme se také po stopách Karpatsko-Dukelské operace 1944, což byla útočná operace sovětských a československých vojsk na severovýchodním Slovensku na podzim 1944, která měla spojit povstalecké síly SNP se sovětskými armádami. Byla jednou z největších bitev druhé světové války. Operace byla součástí Východokarpatské operace. Od 8. září do 27. října 1944 zde zahynuly desítky tisíc sovětských a tisíce československých bojovníků (Čechů, Slováků, Rusínů, a dalších), jakož i jiných národností. Šlo o příslušníky vojenských jednotek obnovujícího se Českosloveska.

Friday I. Agu	
<i>Exploring truncated distributions from the Schröter family ditributions</i>	1
Jaromír Antoch	
<i>Modelling pollution over and around open-cut coal mine Libouš</i>	1
Filip Bočinec	
<i>Polopriestorová hĺbka lokácie a disperzie: Robustnosť a minimax optimalita</i>	1
Tomáš Cipra	
<i>Dynamic MEWMA projections of large dimensional conditional covariance matrices</i>	2
Adéla Czolková	
<i>Statistická analýza dvourozměrných hustot v Bayesových prostorech</i>	2
Zdeněk Fabián	
<i>On information and uncertainty</i>	2
Kamila Fačevicová	
<i>Detekce prukových odlehlých hodnot ve vícefaktorových kompozičních datech</i>	2
Andrej Gajdoš	
<i>Metóda abstraktného Fourierovho rozvoja pre neparametrický odhad difúzneho koeficientu</i>	3
Laura Hajzoková	
<i>Variacia v genetických regulačných sieťach a ich stabilita v závislosti od objemu systému</i>	3
Eliška Hálová	
<i>Problém čtyř bodů</i>	4
Jan Hanousek	
<i>Husté nuly</i>	4
Anna Hejlová	
<i>Když na rozvařenou bramboru nasadíte regresní model</i>	4
František Hendrych	
<i>Differential equations driven by Besov-Orlicz paths</i>	4
Zdeněk Hlávka	
<i>Mnohorozměrné permutační p-hodnoty</i>	5
Daniel Hlubinka	
<i>Statistické testy pro funkcionální data</i>	5
Tomáš Hobza	
<i>Statistický model dopravního proudu na neřízené křižovatce</i>	5
Martin Hrba	
<i>Bootstrapping not independent and not identically distributed data</i>	6
Šárka Hudecová	
<i>Testy pro vícerozměrná eliptická data</i>	6
Marie Hušková	
<i>Some results on detection of changes in panel data</i>	6
Tomáš Chupáň	
<i>Aplikace pro preprocessing, imputaci a analýzu proteomických dat</i>	6
Daniela Jarušková	
<i>Naše zkušenosti s použitím MCMC pro odhadování parametrů v nelineární regresi</i>	7
Jan Janoušek	
<i>Bagging and regression trees in individual claims reserving</i>	7
Paulína Jašková	
<i>Analýza párových logpodielov v kontexte vysoko-dimenzionálnych dát</i>	7
Monika Kalátová	
<i>Dvojúrovňové modely v úlohách teórie portfólia</i>	7
Iva Karafiátová	
<i>Testování nezávislosti orientací</i>	8
Daniel Klein	
<i>Testovanie kovariančnej štruktúry v mnohorozmernom lineárnom modeli</i>	8
Jan Koláček	
<i>Modelování spolehlivosti Li-Ion baterií pro elektricky poháněná letadla</i>	8
Miloš Kopa	
<i>Multivariate probabilistic forecasting of electricity prices with trading applications</i>	8
Aneta Kostárová	
<i>Models for integer-valued time series</i>	9
Roman Králik	
<i>Beer photodegradation development</i>	9
David Kraus	
<i>Analýza funkcionálních dat a cenzorování</i>	9
Matúš Maciak	
<i>Functional profiles and predictions</i>	9
Ján Mačutek	
<i>Lexikálna kompaktnosť textov modelovaná diskretným rozdelením pravdepodobnosti</i>	10

Markéta Makarová	
<i>Rozšíření WGCNA pro skupinové porovnávání</i>	10
Michaela Marččková	
<i>Bayesovské modely aukcí</i>	10
Patrícia Martinková	
<i>Analýza položek vícepoložkových měření</i>	10
Tomáš Masák	
<i>Phase variation of Covid-19 point process</i>	11
Monika Matoušková	
<i>Robustní případ úlohy rozvrhování s náhodným zpožděním a heterogenními stroji</i>	11
Ivan Mizera	
<i>Štatistické metódy pre nepravidelne pozorované časové rady</i>	11
Tomáš Mrkvička	
<i>False discovery rate envelopes</i>	12
Stanislav Nagy	
<i>Polopriestorová hĺbka pre smerové dáta</i>	12
Ivana Pavlů	
<i>Funkcionální metoda hlavních komponent pro hustoty z agregovaných dat</i>	12
Zbyněk Pawlas	
<i>Testy nezávislosti ve stochastické geometrii</i>	13
Michal Pešta	
<i>Changing intensities and band bootstrap</i>	13
Martina Petráková	
<i>Vlastnosti Poissonovy-Laguerrovy mozaiky s neomezenými kótami</i>	14
Ondřej Pokora	
<i>Modelování pomocí degradačních procesů</i>	14
Tomáš Pompa	
<i>Metoda podpurných vektorů pro funkcionální data</i>	14
Zuzana Prášková	
<i>Použití bootstrapu v panelových datech</i>	14
Michaela Prokešová	
<i>Časoprostorové shlukové bodové procesy - nehomogenita, separabilita a odhady</i>	15
Lukáš Račko	
<i>Stability in games with random payoffs</i>	15
Hedvika Ranošová	
<i>Symetrie funkcionálních dat</i>	15
Martin Romaňák	
<i>Multiple changepoint detection in tensor data</i>	15
Samuel Rosa	
<i>Optimálny dizajn, grafy a siete</i>	16
Kristína Sakmarová	
<i>Employing random effects in variance components modelling</i>	16
Pál Somogyi	
<i>Optimálne návrhy pre viacrozmerné modely a algoritmus mREX</i>	16
Gábor Szűcs	
<i>Bonusovo-malusové systémy v poistení motorových vozidiel</i>	17
Jan Vávra	
<i>Odhadování názorových pozic politiků z jejich proslovů</i>	17
Petr Vejmelka	
<i>CDR a jeho využití v rámci clusterování škod v neživotním pojištění</i>	17
Ondřej Vencálek	
<i>Statistika v době post-covidové</i>	17
Ján Veselý	
<i>Diskrétné dynamické systémy s náhodnými impulzmi</i>	18
Ondřej Vozár	
<i>Vícekritériální hodnocení technik znáhodněného dotazování pro odhady populačního průměru</i>	18
Gejza Wimmer	
<i>Kalibrácia dvoch závaží s použitím referenčného závažia</i>	19
Viktor Witkovský	
<i>Poznámka k testovaniu rozdielu mediánov nezávislých populácií</i>	19
Ivan Žezula	
<i>Generalization of matrix mean testing in model with special variance structures</i>	20

FRIDAY I. AGU¹, JÁN MAČUTEK²**Exploring truncated distributions from the Schröter family distributions**¹MI SAV, Štefánikova 49, SK – 814 73 Bratislava²FNSI CPU, Dept of Mathematics, Andreja Hlinku 1, SK – 949 Nitra

agu1@uniba.sk, jmacutek@yahoo.com

Traditional methods for computing insurance payouts often rely on convolution to determine aggregate claim amounts. However, this approach faces challenges, especially when dealing with a large volume of claims and claim severity. In response, the Schröter distribution family has emerged as a promising alternative, featuring a fixed positive probability at zero. This family offers a more accurate depiction of both the number of claims and claim severity, particularly in discrete claim scenarios. Nonetheless, while advantageous, the Schröter distribution family may not fully capture the nuances of claim amounts if they are truncated. This is a common practice in insurance where only events leading to claims are of interest, with the smallest observed claim amount set to one.

Truncated distributions play an important role in risk theory for modeling both claim severity and the intervals between successive claims. They serve as indispensable tools for insurers and actuaries, aiding in the assessment of potential losses' frequency and severity, thus facilitating policy pricing and risk management. This research delves into truncated Schröter distributions, employing a blend of truncated probability distribution and convolution techniques. Specifically, we investigate convolutions involving both truncated and non-truncated variants of distributions from the Schröter family, with the aim of enhancing our comprehension of truncated claim amounts within the realm of insurance.

Acknowledgments: This study was supported by projects VEGA 2/0120/24 (Ján Mačutek/Friday I. Agu) and APP0515 (Friday I. Agu).

JAROMÍR ANTOCH, ELIŠKA CÉZOVÁ**Modelling pollution over and around open-cut coal mine Libouš**

MFF UK, KPMS, Sokolovská 83, CZ – 186 75 Praha 8

FS ČVUT, ÚKaČS, Technická 4, CZ – 160 00 Praha 6

antoch@karlin.mff.cuni.cz

Open-cut coal mine Libouš is located in the NW of Czech Republic (5025'18.562"N, 1320'36.026"E). It is approximately 5 km long and 4 km wide, the shape of mine cavity resembling letter L lying down. Its biggest depth is around 200 metres. The mine is surrounded by the Ore Mountains on N and NW, town Chomutov on NE, town Kadan on SW, and several villages on N and NE. From W, S and E side it is surrounded within a few kilometers by at terrain covered by unmown grass, large water reservoir Nechanice is located nearby to S.

Because of mentioned towns and villages situated downstream of the prevailing wind directions (wind comes most frequently from W, NW and N), the objective of our study was to understand the transport and dispersion of dust particles from mining activities. To that purpose we tried to determine the main ow field within the coal mine cavity and its downstream vicinity with respect to prevailing wind direction(s).

Based on the analysis of the data covering basic meteorological characteristics as well as pollution measured during 2012-2014, we can conclude that an impact of the complex terrain on passive pollutant dispersion with respect to the prevailing wind direction is important. The results also reveal that not only cavity shape and deepness, but also the surroundings orography has influence on the flow patterns, hence on ventilation, within the area of interest.

FILIP BOČINEC**Polopriestorová hĺbka lokácie a disperzie: robustnosť a minimax optimalita**

MFF UK, KPMS, Sokolovská 83, CZ – 186 75 Praha 8

filipbocinec@gmail.com

Klasická polopriestorová hĺbka lokácie je nástrojom neparametrickej štatistiky mnohorozmerných dát, ktorá umožňuje usporiadať body vzhľadom k danému pravdepodobnostnému rozdeleniu. Najhlbší bod potom môže slúžiť ako mnohorozmerný medián, teda odhad lokačného parametru rozdelenia. Analogickým konceptom pre matice je polopriestorová hĺbka disperzie, ktorá umožňuje usporiadať pozitívne definitné matice vzhľadom k danému pravdepodobnostnému rozdeleniu. Táto hĺbka analogicky indukuje odhad disperzie pomocou najhlbšej matice. Tento príspevok sa zaoberá vlastnosťami uvedených odhadov z hľadiska robustnosti a minimax optimality.

TOMÁŠ CIPRA, RADEK HENDRYCH**Dynamic MEWMA projections of large dimensional conditional covariance matrices**

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

cipra@karlin.mff.cuni.cz

Dynamic predictions of large dimensional conditional covariance matrices are considered. Since numerically simple prediction methods are usually recommended in the context of multivariate conditional covariances (covolatilities), one prefers in this paper the multivariate EWMA (exponentially weighted moving average) processes, which are estimated recursively. Various modifications of recursive MEWMA projections are suggested to improve the quality of covolatility projections. Moreover, combinations of multiple types of MEWMA predictions are also included. An extensive numerical study compares particular types of covolatility projections employing various criteria and tests.

ADÉLA CZOLKOVÁ, KAREL HRON**Statistická analýza dvourozměrných hustot v Bayesových prostorech**

PřF UPOL, KMAAM, 17. listopadu 12, CZ–771 46 Olomouc

ivana.pavlu@upol.cz

V tomto příspěvku budou popsány a dále rozvinuty dosavadní poznatky v této oblasti (rozklad interakční části hustoty, ekvivalence různých formulací funkcionální analýzy hlavních komponent u nezávislé části hustoty pro diskretizovaná data), a všechny teoretické postupy budou navíc ilustrovány na reálném datovém souboru.

ZDENĚK FABIÁN**On information and uncertainty**

ÚI AV ČR, Pod Vodárenskou Věží 2, 18207 Praha 8

zdenek@cs.cas.cz

Recently has been proposed description of continuous random variables in terms of the corresponding score random variables, expressing the relative influence of the outcome of a random experiment on the typical value of the probability distribution under consideration. This is the key tool for identifying functions describing relative density of information and uncertainty of the distribution. The mean information is a variant of the Fisher information, mean uncertainty can be used as a new measure of variability of distributions.

KAMILA FAČEVICOVÁ¹, PETER ROUSSEEUW², KAREL HRON¹**Detekce prvkových odlehlých hodnot ve vícefaktorových kompozičních datech**¹PřF UPOL, KMAAM, 17. listopadu 12, CZ–771 46 Olomouc²Department of Mathematics, KU Leuven, Belgie

kamila.facevicova@upol.cz

Podobně jako je tomu v případě reálných dat, i analýza kompozičních dat může být ovlivněna přítomností odlehlých hodnot. Narozdíl od klasických odlehlých pozorování, ta kompoziční se spíše než netypickými hodnotami proměnných vyznačují vybočujícími hodnotami poměrů mezi nimi a jejich detekce pomocí standardních metod aplikovaných na původní data tedy není vhodná. V příspěvku se budeme zabývat problematikou detekce prvkových odlehlých hodnot, která je již částečně vyřešena pro vektorová (jednofaktorová) kompoziční data. My se zaměříme na obecnější problém, kdy jednotlivá (kompoziční) pozorování reprezentují relativní strukturu definovanou dvěma nebo více faktory a mají tedy podobu tabulek nebo vícerozměrných matic. Úloha detekce odlehlých hodnot v těchto strukturách tak mimo detekci jednotlivých prvků a celých pozorování přirozeně zahrnuje i detekci vybočujících kategorií určujících faktorů. Efektivita metody pro detekci všech těchto typů anomálií bude představena prostřednictvím simulační studie a praktické použití metody bude následně demonstrováno na empirických datech.

ANDREJ GAJDOŠ, JOZEF KISEĽÁK, BARBORA KLEMOVÁ**Metóda abstraktného Fourierovho rozvoja pre neparametrický odhad difúzneho koeficientu s využitím vo financiách**

PF UPJŠ, ÚMAT, Jesenná 5, SK – 040 01, Košice 1

andrej.gajdos@upjs.sk

Volatilita (difúzny koeficient) je kľúčovým parametrom modelovania vo finančnej matematike [1]. Dôležitá je okrem iného v hedgingových stratégiách (rôzne stochastické modely volatility ako napr. Blackov-Scholesov a jeho modifikácie), ďalej pri manažmente rizika (tzv. VaR metodológia), pri predikciách (ARCH-GARCH modely) ale tiež hrá významnú úlohu pri optimálnom výbere portfólia. Rôzne metódy výpočtu volatility sú popísané v odbornej literatúre [2, 3, 4]. Tento difúzny koeficient môže byť odhadnutý parametrickým i neparametrickým prístupom. Keďže volatilita sa mení v čase, neparametrické metódy sa pri jej výpočte zameriavajú na malé časové okná, pričom sa využívajú vysokofrekvenčné dáta.

V rámci tohto príspevku prezentujeme nami navrhnutého odhadcu difúzneho koeficientu založeného na neparametrickom prístupe s využitím abstraktného Fourierovho rozvoja. Pritom predpokladáme, že vývoj ceny aktíva je popísaný náhodným procesom nazývaným Brownov semi-martingal a spĺňa príslušnú stochastickú diferenciálnu rovnicu s neznámym difúznym koeficientom, ktorý sa snažíme odhadnúť. Uvádžeme i niektoré základné vlastnosti zavedeného odhadcu, spolu s jeho výhodami či nevýhodami oproti inému existujúcemu odhadcovi podobného typu [5, 6]. V našom príspevku ilustrujeme využitie odhadcu volatility pre vylepšenie tzv. HAR modelu skonštruovaného na reálnych finančných dátach.

Reference

- [1] Buff, R. (2002). *Uncertain Volatility Models - Theory and Application*. Springer.
- [2] Andersen, T., Bollerslev, T. and Diebold, F. (2002). Parametric and nonparametric volatility measurement. In *Handbook of Financial Econometrics*. (L. P. Hansen and Y. Ait-Sahalia, eds.) North-Holland, Amsterdam.
- [3] Kanatani, T. (2004). Integrated volatility measuring from unevenly sampled observations. *Economics Bulletin* 3, no. 36, 1–8.
- [4] Barrucci, Renó, R. (2002). On measuring volatility of diffusion processes with high frequency data. *Economics Letters* 74, 371–378.
- [5] Malliavin, P. and Mancino, M. E. (2002). Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics*, 6(1):49–61.
- [6] Malliavin, P. and Mancino, M. E. (2009). A Fourier transform method for nonparametric estimation of multivariate volatility. *Ann. Statist.*, 37(4):1983–2010.

Podakovanie Táto práca bola podporená Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-21-0216 a APVV-21-0369 ako aj Vedeckou grantovou agentúrou VEGA na základe Zmluvy č. VEGA 1/0585/24.

LAURA HAJZOKOVÁ**Variancia v genetických regulačných sieťach a ich stabilita v závislosti od objemu systému**

ÚM SAV, Dúbravská cesta 9, SK – 841 04 Bratislava

FMFI UKo, KAMŠ, Mlynská dolina F1, SK – 842 48 Bratislava

laura.hajzokova@savba.sk

Expresia génov je heterogénna medzi geneticky identickými bunkami a dynamicky fluktuuje v ich vnútri. Vzájomné interakcie génov formujú tzv. genetické regulačné siete, ktoré majú prirodzene stochastický charakter. V tomto kontexte modelujeme kaskádu proteínov ako Markovovský proces a sledujeme jej stabilitu v závislosti od architektúry siete.

Na rozdiel od deterministického modelovania genetických regulačných sietí, v stochastickom prístupe sa vyskytuje variabilita v dôsledku zavedenia parametra veľkosti systému.

Pri veľkých systémoch, v zmysle počtu molekúl, systém vykazuje takmer deterministické správanie. Avšak mnohé chemické reakcie prebiehajú na tzv. nano-škále, t.j. pri nízkych koncentráciách, kde deterministický prístup zvyčajne zlyhá. V závislosti od objemu systému skúmame stabilný stav syntetickej regulačnej kaskády a prostredníctvom stochastických simulácií sa snažíme presne určiť oblasti zmeny stability.

ELIŠKA HÁLOVÁ**Problém čtyř bodů**

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

eliska.halova@gmail.com

Tento příspěvek se věnuje známé matematické úloze, která nese název problém čtyř bodů. Ptá se na pravděpodobnost, že čtyři náhodně zvolené body v rovině vytvoří konvexní čtyřúhelník. Řešení ale není jednoznačné, liší se podle volby pravděpodobnostního rozdělení bodů. V příspěvku se zabýváme třemi různými volbami rozdělení, a to spojitým rovnoměrným, diskrétním rovnoměrným a dvourozměrným normálním rozdělením, přičemž předpokládáme, že body jsou navzájem nezávislé.

Jan Hanousek**Husté nuly**

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

hanousek@karlin.mff.cuni.cz

Tento výzkum se zaměřuje na speciální typ dat časových řad, kde je významný podíl hodnot rovných nule. Cílem je vytvořit statistický model, který přesně zachycuje chování těchto dat. Prostřednictvím zkoumání stávajících teorií o GARCH a MEM modelech jsou navrženy nové modely spolu s odvozením jejich důležitých teoretických vlastností. Pro posouzení jejich účinnosti jsou tyto modely testovány na reálných datech. Toto hodnocení odhaluje, že každý model má své vlastní silné a slabé stránky. Celkové výsledky jsou však nadějně, prokazují platnost modelů a jejich využitelnost v praxi a otevírají dveře pro další výzkum v této oblasti.

ANNA HEJLOVÁ**Když na rozvařenou bramboru nasadíte regresní model**

TF ČZU, KMF, Kamýcká 169, CZ–165 00 Praha

hejlova@tf.czu.cz

Tento příspěvek vede k zamyšlení nad tím, jaká náhodná data mohou být statisticky vyhodnocována a následně interpretována. Jako příklad jsou uvedena data získaná z testování rozvařivosti brambor, která jsou ovlivněna řadou náhodných vlivů. Nově vyvinutá metoda je založená na sledování hmotnosti vzorků v průběhu vaření, při níž jsou data odečítána a zapisována ručně. Z varných křivek jsou poté odvozeny parametry, které jsou vztaženy k dalším vlastnostem bramborových hlíz. Bohužel, ani samotný zkoumaný materiál zdaleka není homogenní. Mohou taková data o něčem průkazně vypovídat?

FRANTIŠEK HENDRYCH**Differential equations driven by Besov-Orlicz paths**

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

hendrychfrantisek@karlin.mff.cuni.cz

Stochastic processes, such as Brownian motion, are often incorporated into differential equations to accommodate non-systematic errors or uncertainties in the model. These processes typically exhibit sample paths that are nowhere differentiable. Consequently, the concept of an integral, and consequently that of a solution to the equation, must be approached differently from a path-by-path perspective. One common approach involves constructing the integral as a limit of a Riemann-type sum within the L^2 space. This holds true for both the Itô and Stratonovich integrals.

When considering the standard Brownian motion B , the solution map

$$S : B(\omega) \mapsto Y(\omega) \quad \text{where } Y \text{ solves the Itô SDE } dY_t = f(Y_t)dB_t,$$

is known as the Itô map. This map is measurable but lacks continuity, regardless of the norm used to equip the space of realizations of B . To address this issue, rough path theory offers a solution. It suggests that the Itô solution map can be decomposed into a measurable map Ψ and a continuous solution map \hat{S} as

$$B(\omega) \xrightarrow{\Psi} (B, \mathbb{B})(\omega) \xrightarrow{\hat{S}} Y(\omega).$$

This decomposition breaks the procedure into two steps. In the first step, which is probabilistic, one constructs the iterated integral \mathbb{B} of the path B (e.g., in the Itô or Stratonovich sense). This step, encapsulated in the map Ψ ,

is referred to as a rough path lift. In the second step, represented by the map \hat{S} , the analytical aspect comes into play. This step takes both the path B and its (iterated) integral \mathbb{B} as input and solves the now deterministic equation for that input.

The standard rough path theory is constructed for drivers with Hölder continuous paths, resulting in Hölder continuity of the paths of the solution. However, there exist more appropriate function spaces for all the canonical examples of drivers in stochastic differential equations, such as Besov spaces or Besov-Orlicz spaces. Interestingly, it has been observed that if we confine ourselves to these spaces, the solution also resides within these smaller spaces. The discussion will focus on investigating this concept further.

ZDENĚK HLÁVKA, DANIEL HLUBINKA, ŠÁRKA HUDECOVÁ

Mnohorozměrné permutační p -hodnoty

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

hlavka@karlin.mff.cuni.cz

Permutační testy se často úspěšně používají v situacích s neznámým nebo komplikovaným nulovým rozdělením testové statistiky. V mnohorozměrném případě, například pokud simultánně používáme několik testů, lze mnohorozměrný permutační test založit na takzvané kombinační funkci, tj. na vhodné jednorozměrné transformaci vektoru jednorozměrných (částečných) permutačních p -hodnot [3]. Volba kombinační funkce, kterých byla navržena celá řada, přitom může výsledek testu výrazně ovlivnit.

V našem příspěvku představíme postup založený na optimálním transportu, který umožňuje mnohorozměrnou permutační p -hodnotu vypočítat pomocí přirozeně definovaných mnohorozměrných kvantilů [1]. Navrženou metodu můžeme použít buď na vektor p -hodnot nebo přímo na mnohorozměrnou testovou statistiku. Výslednou mnohorozměrnou p -hodnotu lze rozložit na příspěvky jednotlivých složek a při zamítnutí nulové hypotézy tak můžeme výsledek testu snadno interpretovat.

Navržený postup představíme na klasickém k -výběrovém problému, kde můžeme mnohorozměrnou testovou statistiku složit z několika t -testů a na testování shody rozdělení funkcionálních dat [2], kde dvourozměrnou testovou statistiku získáme zkombinováním testu shody středních funkcí a testu shody kovariančních operátorů. V simulační studii zkoumáme volbu počtu permutací, některé technické detaily i výpočetní náročnost.

Reference

- [1] Hlávka, Z., Hlubinka, D., Hudecová, Š. (2023): Multivariate quantile-based permutation tests with application to functional data, preprint arXiv:2311.04017
- [2] Hlávka, Z., Hlubinka, D., Koňasová, K. (2021): Functional ANOVA based on empirical characteristic functionals. *J. Multivariate Anal.* **189** 104878.
- [3] Pesarin, F. (2001): *Multivariate permutation tests: with applications in biostatistics*. J. Wiley.

DANIEL HLUBINKA

Statistické testy pro funkcionální data

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

hlubinka@karlin.mff.cuni.cz

Na rozdíl od konečněrozměrných náhodných vektorů se u funkcionálních náhodných veličin nemůžeme spolehnout na hustotu rozdělení vůči nějaké univerzální σ -konečné míře. V příspěvku se proto podíváme na využití charakteristických funkcionálů pro testování některých hypotéz o funkcionálních datech. Zejména testy typu ANOVA, testy nezávislosti a subnezávislosti a naznačíme i další souvislosti.

TOMÁŠ HOBZA

Statistický model dopravního proudu na neřízené křižovatce a jeho využití pro odhad kapacity

FJFI ČVUT, Katedra matematiky, Trojanova 13, CZ–120 00 Praha 2

tomas.hobza@fjfi.cvut.cz

Tento příspěvek je věnován pravděpodobnostnímu modelu dopravního proudu na neřízené křižovatce tvaru T a jeho možnému využití pro výpočet kapacity křižovatky. Použitá statistická metoda vychází z obecně uznávaného postupu formulovaného Wernerem Sieglochtem, v němž je odhad kapacity ovlivněn tzv. akceptačním pravidlem, kdy řidič vozidla na vedlejší komunikaci akceptuje/odmítá mezeru (časový odstup) mezi dvěma po sobě jedoucími vozidly na hlavní komunikaci. V tomto příspěvku je toto stochastické schéma přeformulováno do jazyka náhodných

veličin. To umožňuje ověřit starší přístupy, kterými různí autoři odhadovali pravděpodobnosti přijetí mezery a další související statistické charakteristiky, včetně samotné kapacity. Provedená analýza přesvědčivě ukazuje, že některé tradiční regresní techniky používané v praktických aplikacích Sieglochovy metody mohou fatálně selhat. Z tohoto důvodu je navržena nová metodika založená na aplikaci moderních regresních metod. Přínos navrženého modelu a metodiky spočívá v zásadním zpřesnění odhadu kapacity, jak je demonstrováno aplikací na empirické soubory dat.

Martin Hrba

Bootstrapping not independent and not identically distributed data

MFF UK, KPMS, Sokolovská 83, CZ – 186 75 Praha 8

hrba@karlin.mff.cuni.cz

Classical normal asymptotics could bring serious pitfalls in statistical inference, because some parameters appearing in the limit distributions are unknown and, moreover, complicated to estimated (from a theoretical as well as computational point of view). Due to this, plenty of stochastic approaches for constructing confidence intervals and testing hypotheses cannot be directly applied. Bootstrap seems to be a plausible alternative. A methodological framework for bootstrapping not independent and not identically distributed data is presented together with theoretical justification of the proposed procedures. Among others, bootstrap laws of large numbers and central limit theorems are provided. The developed methods are utilized in insurance and psychometry.

ŠÁRKA HUDECOVÁ¹, MIROSLAV ŠIMAN²

Testy pro vícerozměrná eliptická data založené na geometrickém přístupu ke konstrukci pořadí a znamének

MFF UK, KPMS, Sokolovská 83, CZ – 186 75 Praha 8

ÚTIA AV ČR, Pod Vodárenskou věží 4, CZ – 182 00 Praha 8

hudecova@karlin.mff.cuni.cz

Pořadové testy pro vícerozměrná data tvoří důležitou třídu statistických testů. Pro eliptická rozdělení se většinou využívají tzv. Mahalanobisova pořadí a znaménka, která však vyžadují konzistentní odhad škálovací matice. To může způsobovat problém např. v případě rozdělení s těžšími chvosty. Z tohoto důvodu byl v literatuře navržen koncept pořadí a znamének založený na geometrickém přístupu, konkrétně na relativním počtu určitého typu nadrovin. Nevýhodou původního konceptu je velká výpočetní náročnost, která nedovoluje výpočet pro vyšší dimenze. V tomto příspěvku ukážeme možná řešení, která daný geometrický koncept zjednodušují při zachování asymptotických vlastností příslušných testů.

MARIE HUŠKOVÁ

Some results on detection of changes in panel data

MFF UK, KPMS, Sokolovská 83, CZ – 186 75 Praha 8

huskova@karlin.mff.cuni.cz

Various results for detection of changes in panel data models are presented. Particularly, the focus is on panel regression models with cross-sectional dimension N based on T observations. Both tests and estimators are discussed. Also independent panels and the case of mild cross-sectional dependence are studied. Theoretical results are accompanied by simulations and illustrative applications to real data sets. The talk is based on several papers with several coauthors.

TOMÁŠ CHUPÁŇ

Aplikace pro preprocessing, imputaci a analýzu proteomických dat

PrF UPOL, KMAAM, 17. listopadu 12, CZ – 771 46 Olomouc

tomas.chupan@seznam.cz

Cílem příspěvku je představit novou Shiny aplikaci *proteoME*, která slouží jako interaktivní nástroj pro preprocessing, imputaci a analýzu proteomických dat. Po importu datové sady nabízí základní vizualizaci dat, jejich transformaci a normalizaci. Následujícím krokem je agregace na úroveň vzorků, při níž má uživatel k dispozici kromě několika možností postupu také vizualizace usnadňující volbu parametrů. Agregovaná data je také možné filtrovat dle procenta chybějících hodnot na několika úrovních, příp. chybějící hodnoty imputovat jednou z dostupných metod imputace. Samotná analýza takto připravených proteomických dat poskytuje několik možností

testování shody distribuce abundancí proteinů mezi jednotlivými skupinami léčby – lze porovnávat libovolný počet takových skupin. Výsledky analýzy ve formě detailní tabulky s širokou paletou možností exportu, řazení řádků i filtrace hodnot se snaží maximálně vyhovět potřebám uživatele, podobně jako hlavní grafický výstup analýzy – *volcano plot*, který si uživatel taktéž může přizpůsobit dle svých požadavků. Bude nastíněno také ukázkové zpracování syntetické datové sady v aplikaci *proteoME*.

JAN JANOUŠEK

Bagging and regression trees in individual claims reserving

MFF UK, KPMS, Sokolovská 83, CZ – 186 75 Praha 8

janousek@karlin.mff.cuni.cz

This lecture provides a practical exploration of classification and regression trees, alongside bootstrap aggregating, in the context of individual reserving in insurance. Beginning with a thorough overview of theoretical foundations often overlooked in introductory materials, we establish mathematical formalities crucial for understanding. Expanding upon prior research, we integrate regression trees and bagging techniques to enhance the accuracy of reserve estimates, particularly in modeling claim sizes. By applying these methods to insurance data, empirical distributions are derived, enabling the calculation of confidence intervals and quantiles necessary for determining reserves for both the upcoming year and ultimate reserves. Through this examination, attendees gain insights into leveraging machine learning for more effective individual reserving strategies within the insurance sector.

DANIELA JARUŠKOVÁ, MICHAL ZDRAŽIL

Naše zkušenosti s použitím MCMC pro odhadování parametrů v nelineární regresi

FSv ČVUT, KM, Thákurova 7, Praha CZ – 1600 Praha 6

daniela.jaruskova@cvut.cz

Při spolupráci s inženýry naší fakulty se v posledních letech nejčastěji setkáváme s problémem odhadu mnohdimenzionálního parametru v nelineární regresi. Inženýři řeší tuto úlohu metodou MCMC, kde používají klasickou náhodnou procházku a Metropolisův-Hastingsův algoritmus. Vzniká přirozená otázka, zda klasické numerické metody nebo naopak sofistikované metody MCMC nemohou dát v určitém konkrétním problému „lepší“ výsledky.

PAULÍNA JAŠKOVÁ, M. TEMPL, D. LACHMAN, M. BERLAND, K. HRON, J. PALAREA-ALBALADEJO

Analýza párových logpodielov v kontexte vysoko-dimenzionálnych dát

PřF UPOL, KMAAM, 17. listopadu 12, CZ – 771 46 Olomouc

jaskovapaulina@gmail.com

Kompozičné dáta sú dáta nesúce relatívnu informáciu, najjednoduchší spôsob na vyjadrenie kompozičných dát je prostredníctvom párových logpodielov medzi jednotlivými komponentami. Súradnicové systémy logpodielov, často využívané v praxi, ako bilancie a pivotové súradnice, agregujú základnú informáciu spolu a tým zachytávajú vyššiu úroveň závislosti medzi komponentami. Avšak v niektorých prípadoch môže byť výhodné poskytnúť podrobnejšiu reprezentáciu vo forme základných párových logpodielov, v kontexte štatistických analýz týkajúcich sa napríklad redukcie dimenzií, regresnej analýzy alebo klasifikácie. Jednou z hlavných požiadavok v kompozičnej analýze je využívať ortonormálne logpodielové vyjadrenie, avšak táto požiadavka predstavuje náročný cieľ pri práci s vysoko-dimenzionálnymi dátami, nakoľko D zložková kompozícia vedie k $D \cdot (D - 1) / 2$ párovým logpodielom. Na základe využitia teórie Latinských čtvorcov tento príspevok predstavuje algoritmus, ktorý umožňuje reprezentovať všetky párové logpodielové pomocou iba $D - 1$ logpodielových súradnicových systémov. Prezentované metódy sú ilustrované na dvoch prípadových štúdiách.

MONIKA KALATOVÁ, MILOŠ KOPA

Dvojúrovňové modely v úlohách teórie portfólia

MFF UK, KPMS, Sokolovská 83, CZ – 186 75 Praha 8

kalatova@karlin.mff.cuni.cz

Je bežnou praxou, že investor sa nerozhoduje iba na základe in-sample výkonnosti portfólia, ale zaujíma ho aj out-of-sample výkonnosť. Tieto dva prístupy spojíme a vytvoríme dvojúrovňový stochastický optimalizačný problém. Dolná úroveň pozostáva z mean-risk efektívneho optimalizačného problému, kde ako mieru rizika použijeme

podmienenu hodnotu v riziku. Horná úroveň vyberá in-sample efektívne portfólia (výsledky z dolnej úrovne), ktoré maximalizujú out-of-sample výnos.

Reformulácia problému predpokladá diskkrétne pravdepodobnostné rozdelenie, čo vedie k lineárnemu dvojúrovňovému optimalizačnému problému. Pri moving-window analýze pracujeme s dvoma typmi úlohy – premenné v hornej úrovni sú rovnaké pre všetky okná alebo sa líšia. Týmto prístupom nakoniec dostaneme dva lineárne dvojúrovňové problémy s niekoľkými úlohami v dolnej úrovni. Výsledky pre oba prípady budú prezentované na konferencii.

IVA KARAFIÁTOVÁ

Testování nezávislosti orientací

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

karafiatova@karlin.mff.cuni.cz

Předmětem této práce je nezávislost r -tic orientací symetrických objektů v trojrozměrném prostoru. Kromě aplikací v krystalografii a materiálové vědě se takové objekty vyskytují napříč různými vědními obory. Hlavním přínosem této práce je představení nové definice kovariance mezi dvěma orientacemi a konstrukce vícerozměrných asymptotických testů nezávislosti po dvou. Síla navržených testů je porovnána prostřednictvím simulační studie.

DANIEL KLEIN

Testovanie kovariančnej štruktúry v mnohorozmernom lineárnom modeli

PF UPJŠ, ÚM, Jesenná 5, SK–040 11, Košice

daniel.klein@upjs.sk

Experiment je jedným z hlavných nástrojov vedeckého výskumu a jeho výsledkom je zvyčajne súbor nameraných alebo pozorovaných údajov. Pri analýze dát je potrebné vyriešiť dva hlavné problémy – odhadnúť hodnoty neznámych parametrov a testovať hypotézy o hodnotách neznámych parametrov. Prednáška bude zameraná na problematiku testovania neznámych parametrov vo viacúrovňových mnohorozmerných lineárnych modeloch. Prezentované budú odhady neznámych parametrov za určitých špecifických podmienok, najmä za predpokladu špeciálnych štruktúr kovariančnej matice, a tiež testovanie strednej hodnoty a kovariančnej štruktúry.

Podakovanie: Práca bola podporená projektmi APVV-21-0369 a VEGA 1/0585/24.

JAN KOLÁČEK

Modelování spolehlivosti Li-Ion baterií pro elektricky poháněná letadla

PřF MU, ÚMS, Kotlářská 2, CZ–611 37 Brno

kolacek@math.muni.cz

Zkoumáme vlastnosti a indikátory spolehlivosti u dobíjecích Li-Ion baterií, které jsou zdrojem energie pro elektrické letadlo. Tyto baterie byly skladovány po dobu pěti let. Na bateriích byly provedeny zátěžové testy. Zaměříme se na statistické modelování průběhu napětí při těchto vybíjecích procesech. Cílem je simulovat velký počet vybíjecích procesů a nalézt pravděpodobnostní rozdělení prvního dosažení času (FHT), který je definován jako čas, kdy napětí na článku baterie dosáhne kritické hodnoty 2,51 V. Výsledky poté porovnááme s hodnotami ze zátěžových testů pro stejné typy nových Li-Ion baterií, které nebyly skladovány. Pro modelování poklesu napětí používáme techniky analýzy funkcionálních časových řad, protože analyzujeme data s vysokou frekvencí měření s nepatrnou chybou měření, tj. data vykazují funkcionální charakter.

MILOŠ KOPA

Multivariate probabilistic forecasting of electricity prices with trading applications

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

kopa@karlin.mff.cuni.cz

A recently introduced approach is extended to probabilistic electricity price forecasting (EPF) utilizing distributional artificial neural networks, based on a regularized distributional multilayer perceptron (DMLP). We develop this technique for a multivariate case EPF with incorporated dependence. The performance of a fully connected architecture and a LSTM architecture of neural networks are tested. The empirical data application analyzes two day-ahead electricity auctions for the United Kingdom market. This creates the opportunity to buy in the first

auction for lower price and sell in the second for higher price (or vice versa). Utilizing forecasting results, we develop trading strategies with various investors' objectives. We find that, while DMLP shows similar performance compared to the benchmarks, the algorithm is considerably less computationally costly.

ANETA KOSTÁROVÁ

Models for integer-valued time series

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

kostarova.aneta@gmail.com

In this contribution, we deal with models for time series of integer-valued variables. Such series arise in various applications, often as increment series for counts of interest, such as the number of customers, insurance claims, financial transactions, or goals in ice hockey match. The presented model has a GARCH-type structure, with the baseline conditional distribution of the outcome, given the past, is modelled using the Skellam distribution. We provide an introduction to the topic.

ROMAN KRÁLÍK, PETR GABRIEL, JAROMÍR ANTOCH

Beer photodegradation development

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

romankralik5@gmail.com

It is well-known that exposure of beer to light leads to the development of undesirable flavors or even complete spoilage of the beverage. To better assess development of this spoilage, laic sensory panels were organized. On the basis of the panel tests, the level of light beer damage was estimated, along with demonstrating the exponential growth in the ability of lay assessors to recognize damaged beer as its damage increased. Furthermore, the progression of beer damage by natural sunlight and accelerated damage in an accelerator was measured. These measurements were compared, and a time conversion scale between natural and accelerated damage was estimated.

Lastly, predominantly two-hour courses of accelerated absorption development were measured. The measured samples were either completely undamaged before the measurement itself or had been damaged by natural light for a certain period before the measurement. These courses were compared with each other and the observed deviations in damage development between them were described.

DAVID KRAUS

Analýza funkcionálních dat a cenzorování

PrF MU, ÚMS. Kotlářská 2, CZ–611 37 Brno

david.kraus@mail.muni.cz

V příspěvku se budeme bavit o novém typu pozorování funkcionálních dat, kdy hodnoty funkcí jsou pozorovány v diskretních (nepravidelných, často řídkých) argumentech s pozorovacím šumem a navíc mohou být cenzorované. Dozvíme se, jak takovou situaci modelovat, jak odhadovat jednotlivé složky modelu a jak rekonstruovat individuální trajektorie. Metody budou ilustrovány na datech z výzkumu HIV.

MATŮŠ MACIAK¹, IVAN MIZERA², MICHAL PEŠTA¹

Functional profiles and predictions

¹MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

²University of Alberta, Edmonton, Canada

{maciak,pesta}@karlin.mff.cuni.cz, imizera@ualberta.ca

Functional profiles (no matter what it actually means) can be easily used to represent various data structures with a wide range for real-life applications— usually because such functional representations are, typically, quite intuitive and relatively straightforward to understand.

We apply such “functional profiles” idea for the data common and typical in claim reserving problems (and not just exclusively there) and we use so-called run-off data triangles to estimate and predict the overall claims reserve that must be covered by an insurance company. Three competitive techniques are proposed and discussed. The estimation and the prediction are both performed in a fully data-driven manner and the overall reserve is given in terms of the whole distribution obtained via a permutation bootstrap.

Some theoretical justifications are provided and practical implementation issues are addressed. A finite sample evaluation in terms of a full-scale comparison with standard (parametric) reserving techniques is carried on several hundreds of real run-off triangles against known real loss outcomes.

Reference

- [1] Maciak, M., Mizera, I., Pešta, M. (2022). Functional profile techniques for claims reserving. *ASTIN Bulletin* **52(2)**, 449–482.

JÁN MAČUTEK

Lexikálna kompaktnosť textov modelovaná (asi novým) diskretným rozdelením pravdepodobnosti

MI SAV, Štefánikova 49, SK–814 73 Bratislava

NSI CPU, Dept of Mathematics, Andreja Hlinku 1, SK–949 Nitra

jmacutek@yahoo.com

V texte pozostávajúcom z n viet sú dve vety spojené „lexikálnym linkom“, ak obsahujú to isté plnovýznamové slovo. Do úvahy berieme nielen link samotný (ak existuje), ale aj jeho dĺžku. Lexikálna kompaktnosť textu je vážený súčet linkov (kde váhami sú ich dĺžky) vydelený maximálnou možnou hodnotou tohto súčtu. Čitateľ zlomku je diskretnou náhodnou premennou, ktorej rozdelenie a jeho základné vlastnosti budú predstavené v tomto príspevku.

Pod'akovanie Podporované grantom VEGA 2/0120/24.

MARKÉTA MAKAROVÁ

Rozšíření WGCNA pro skupinové porovnávání

PřF MU, ÚMS, Kotlářská 2, CZ–611 37 Brno

468063@mail.muni.cz

Weighted Gene Co-Expression Network Analysis (WGCNA) patří mezi populární metody pro analýzu genomických a proteomických datových souborů. Tato metoda však dosud nebyla plně přizpůsobena pro meziskupinové porovnávání. Některá existující rozšíření sice umožňují párové porovnání, avšak nedovolují analyzovat rozdíly na úrovni jednotlivých genů a proteinů. V tomto příspěvku představujeme nové rozšíření WGCNA, které umožňuje detailnější inferenci pro meziskupinové porovnávání, založené na metodách high-dimension low sample size (HDLSS) statistiky. Naše výsledky demonstrujeme na proteomických datech, která porovnávají profily podkožní a viscerální tukové tkáně u obézních pacientů. Tato rozšíření poskytují hlubší vhled do biologických rozdílů mezi skupinami a mohou přispět k lepšímu porozumění mechanismům biomarkerů.

MICHAELA MARČEKOVÁ

Bayesovské modely aukcí

PřF MU, ÚMS, Kotlářská 2, CZ–611 37 Brno

500000@mail.muni.cz

Príspevok sa zaoberá bayesovskými modelmi priebehu aukcií. Venujeme sa štyrom modelom aukcií. Dva modely *BSAR* a *BSARM*, bayesovský regresný model spektrálnej analýzy a jeho verzia s obmedzením na monotónnosť, využívajú na špecifikovanie apriórneho rozdelenia neparametrickej časti regresnej funkcie gaussovský proces. Na vylepšenie odhadu priebehu aukcií navrhujeme ďalšie dva modely *BSAR-GF* a *BSARM-GF*, v ktorých gaussovský proces nahradzame gaussovským polom. Tieto modely sú navrhnuté tak, aby umožňovali zahrnúť do modelu vyvolávaciu cenu neparametricky a v interakcii s časom.

PATRÍCIA MARTINKOVÁ

Analýza položek vícepoložkových měření

ÚI AV ČR, OSM, Pod Vodárenskou věží 2, CZ–182 07 Praha 8

PedF UK, ÚVRV, Myslíkova 7, CZ–110 00 Praha 1

martinkova@cs.cas.cz

Měření v psychologii, pedagogice a v dalších sociálních vědách jsou často nepřesná a využívají proto většího množství položek či hodnotitelů. V příspěvku pojednáme o statistických metodách pro analýzu položek vícepoložkových měření. Pojednáme o modelech tzv. teorie odpovědi na položku (item response theory, IRT) a jejich vztahu s modely faktorové analýzy a modely zobecněné lineární a nelineární regrese se smíšenými efekty. Dále popíšeme postupný vývoj modelů IRT pomocí empirických charakteristických křivek a zobecněných lineárních a nelineárních modelů (GLNM) s důrazem na didaktickou hodnotu takového přístupu. Nakonec představíme některé nové přístupy k odhadu parametrů a jejich praktickou implementaci.

TOMÁŠ MASÁK**Phase variation of Covid-19 point process**

Institute for Stat and Math, WU Wien, Welthandelsplatz 1, AT-1020 Vienna

tmasak@wu.ac.at

This study investigates the influence of governmental restrictions on the spread of COVID-19 during the early stages of the pandemic, as measured by infection counts. Leveraging daily infection data from each US state and the Oxford Stringency Index, we model these datasets as realizations of a point process and functional data, respectively. The realizations of the latent COVID-19 process are allowed to vary not only in magnitude, but also temporally.

By non-parametrically separating these amplitude and phase variations, we examine the impact of government restrictions on these sources of variability. Employing Wasserstein PCA, we analyze the COVID-19 temporal variability as well as the distribution of restrictions represented by the Stringency Index. Subsequently, we utilize the resulting scores, along with constants representing the overall stringency budget and total COVID infection counts, as inputs for a linear vector-on-vector regression model.

Our findings reveal that promptly-imposed restrictions effectively push the incidence of COVID-19 forward in time. However, we observe no significant effect of these restrictions on the overall infection counts.

MONIKA MATOUŠKOVÁ**Robustní případ úlohy rozvrhování s náhodným zpožděním a heterogenními stroji**

MFF UK, KPMS, Sokolovská 83, CZ – 186 75 Praha 8

matouskova@karlin.mff.cuni.cz

Tento příspěvek pojedná o úlohách rozvrhování s pevně danými intervaly prací. Konce prací mohou být ovlivněny náhodným zpožděním. Zadané práce chceme přiřadit strojům tak, aby rozvrh zůstal přípustný s co největší pravěpodobností. Uvažujeme heterogenní případ, tedy máme více typů prací a strojů. Předpokládáme, že sdružené rozdělení náhodných zpoždění prací je dáno Archimedovskou kopulí. Zajímá nás rozvrh s největší pravděpodobností přípustnosti v nejhorším případě, kdy stresujeme daný podíl marginálních rozdělení zpoždění prací. Tento problém má zajímavou reformulaci, která obsahuje často používanou míru rizika. Ukážeme, jak lze takový problém řešit numericky.

IVAN MIZERA**Štatistické metódy pre nepravidelne pozorované časové rady**

MFF UK, KPMS, Sokolovská 83, CZ – 186 75 Praha 8

University of Alberta, MSS, Edmonton, AB T6G 2G1, Canada

mizera@karlin.mff.cuni.cz, imizera@ualberta.ca

Pod pojmom „štatistická analýza časových radov“ si väčšina z nás absolvovavších štandardné štatistické vzdelanie vybaví sekvenčné dáta, ktoré síce nejde zahrnúť do klasického „iid“ probabilistického modelu — pričom nejde až tak o to posledné „id“, ako o to prvé „i“, užitočná dogma pravdepodobnostnej nezávislosti nie je pre ne realistická — no napriek tomu, čo sa indexovania časom týka, akceptovateľné ako matematické postupnosti abstrahujúce od potenciálnych iregularít časových intervalov medzi jednotlivými pozorovaniami. Ak sa nejaké iregularity vôbec vyskytnú, tak sa z praktického hľadiska považujú za zanedbateľné — v konečnom dôsledku už len index „ t “, nahradzujúci štandardné „ i “ pripomína, že špecifický dôraz je tu na orientácii od minulosti k budúcnosti a nie naopak.

Je treba priznať, že uvedené zjednodušenie bolo rozhodujúcim faktorom, ktorý umožnil, že sa klasická analýza časových radov vyvinula do nesmierne potentného nástroja s obrovskou literatúrou, sofistikovanej metodológie založenej po stránke teoretickej na stochastike diskretných náhodných procesov, a po stránke výpočtovej na dômyselných algoritmoch diskretnéj Fourierovej analýzy. No na druhej strane, časové rady „nepravidelne pozorované“ — a to tak nepravidelne, že sa ani náhodou nedajú vtisnúť do aproximatívnych interpolačných stratégií praktických aplikácií — sú tu, a dožadujú sa patričnej pozornosti.

Je to tak? A ak áno, oddá sa im ju venovať? Stojí to za to? A v neposlednom rade, má to vôbec nádej na úspech? Je dosť zrejmé, že v duchu inkrementálneho charakteru súčasnej vedy prirodzeným východiskom bude už spomenutá bohatá metodológia pre časové rady pozorované pravidelne, ktoré tiež budú akosi „klasickou limitou“ našich výbojov — a bude azda zaujímavé sledovať, ako ďaleko sme sa už v tomto smere dostali, a čo ešte stále je pred nami.

TOMÁŠ MRKVIČKA**False discovery rate envelopes**

EF JČU, KAMI, Studentská 13, CZ–370 05 České Budějovice

mrkvicka.toma@gmail.com

False discovery rate (FDR) is a common way to control the number of false discoveries in multiple testing. There are a number of approaches available for controlling FDR. However, for functional test statistics, which are discretized into m highly correlated hypotheses, the methods must account for changes in distribution across the functional domain and correlation structure. Further, it is of great practical importance to visualize the test statistic together with its rejection or acceptance region. Therefore, based on resampling principles, a graphical envelope that controls FDR and detects the outcomes of all individual hypotheses by a simple rule: the hypothesis is rejected if and only if the empirical test statistic is outside of the envelope, will be presented. Such an envelope offers a straightforward interpretation of the test results, similarly as the recently developed global envelope testing which controls the family-wise error rate.

STANISLAV NAGY¹, RAINER DYCKERHOFF²**Polopriestorová hĺbka pre smerové dáta**¹MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8²Universät zu Köln, Kolín nad Rýnom, Nemecko

nagy@karlin.mff.cuni.cz

Angulárna polopriestorová hĺbka (**ahD**) je prirodzenou modifikáciou oslavovanej polopriestorovej (alebo Tukeyho) hĺbky do kontextu smerových dát. **ahD** umožňuje zavedenie prvkov neparametrickej inferencie, ako sú medián, medzi-quantilové oblasti, alebo poradové štatistiky, pre dáta žijúce na jednotkovej sfére \mathbb{S}^{d-1} . Napriek tomu, že **ahD** bola v literatúre prvýkrát popísaná už v roku 1987, táto hĺbka nikdy nedosiahla širšieho uznania, najmä z dôvodu predpokladanej vysokej výpočetnej náročnosti.

V tomto príspevku sa budeme zaoberať tak výpočtom, ako aj teóriou **ahD**. Použitím špeciálnej projekcie vyjadríme **ahD** na jednotkovej sfére \mathbb{S}^{d-1} ako zovšeobecnenú (euklidovskú) polopriestorovú hĺbku v dimenzii $d - 1$. To nám dovolí vyvinúť rýchle a presné výpočetné algoritmy pre **ahD** v každej dimenzii $d \geq 1$. Ďalej sa budeme zaoberať teoretickými vlastnosťami **ahD**. Ukážeme, že **ahD** má — ako jediná v literatúre — všetky vlastnosti očakávané od hĺbky vhodnej pre smerové dáta. Nakoniec, odvodíme rovnomernú spojitosť/konzistenciu príslušného mediánu pre smerové dáta a centrálnych **ahD**-oblastí, predstavujúcich analógiu kvantilov pre smerové dáta.

Reference

- [1] Dyckerhoff, R., a Nagy, S. (2024). Exact computation of angular halfspace depth. *Preprint*.
- [2] Nagy, S., a Laketa, P. (2024). Theoretical properties of angular halfspace depth. *Bernoulli*, prijaté k publikácii.
- [3] Nagy, S., Demni, H., Buttarazzi, D., a Porzio, G. C. (2024). Theory of angular depth for classification of directional data. *Advances in Data Analysis and Classification*, prijaté k publikácii.
- [4] Small, C. G. (1987). Measures of centrality for multivariate and directional distributions. *The Canadian Journal of Statistics*, **15**, 31–39.

IVANA PAVLŮ, J. MACHALOVÁ, R. TOLOSANA-DELGADO, K. HRON, K.G. van den BOOGAART**Funkcionální metoda hlavních komponent pro pravděpodobnostní hustoty získané z agregovaných dat**

PřF UPOL, KMAAM, 17. listopadu 12, CZ–771 46 Olomouc

ivana.pavlu@upol.cz

Analýza funkcionálních dat [1, 2] umožňuje skrze vhodnou datovou reprezentaci aplikování původně mnoho-rozměrných statistických metod na funkce, včetně pravděpodobnostních hustot. U těchto objektů je třeba brát v potaz relativní charakter obsažené informace, což znemožňuje použití standardních funkcionálních metod, např. funkcionální metody hlavních komponent, pro objekty z L^2 prostoru. Namísto toho je vhodné je chápat jako prvky tzv. Bayesova prostoru \mathcal{B}^2 , díky jehož struktuře Hilbertova prostoru je možné transformovat hustoty do L^2 jako reálné funkce s nulovým integrálem [3, 4]. Pro jejich vhodnou reprezentaci byly zavedeny tzv. kompoziční splajny [5], které nulový integrál splňují - v této práci se navíc pracuje s ortogonalizovanou příslušnou splinovou bází.

Při zpracovávání pravděpodobnostních hustot se běžně setkáváme se situací, kdy jsou podkladová, často diskrétní data, získána agregací původních měření (např. v geovědách, demografii atd.). Takto agregovaná data,

např. v podobě četnostních/proporčních histogramů, jsou potom vstupem pro vytváření vyhlazených hustot. Původní získané hodnoty jsou ale nutně zatížené chybou měření, které může ovlivnit naše vnímání variability v datech a tedy např. i počet a tvar hlavních komponent. Výsledky ‘standardní’ metody hlavních komponent se tedy stávají nepřesnými a mohou mít negativní dopad na následnou interpretaci výsledků.

V tomto příspěvku se proto budeme věnovat odhadu této chyby měření a jejímu zakomponování do funkcionální metody hlavních komponent. Efekt chyby měření bude ukázán jak na simulacích, tak na reálných datech popisujících rozdělení částic v říčních sedimentech.

Reference

- [1] Ramsay, J. and Silverman, B.W.: Functional Data Analysis. Springer, New York (2005).
- [2] Kokoszka, P. and Reimherr, M.: Introduction to Functional Data Analysis. Chapman and Hall/CRC, New York (2017).
- [3] van den Boogaart, K.G. and Egozcue, J.J. and Pawlowsky-Glahn, V.: Bayes linear spaces. *Statistics and Operations Research Transactions*, 34(2): 201-222 (2010).
- [4] van den Boogaart, K.G. and Egozcue, J.J. and Pawlowsky-Glahn, V.: Bayes Hilbert spaces. *Australian & New Zealand Journal of Statistics*, 54(2): 171-194 (2014).
- [5] Machalová, J. and Talská, R. and Hron, K. and Gába, A.: Compositional splines for representation of density functions. *Computational Statistics*, 36: 1031-1064 (2021).

ZBYNĚK PAWLAS

Testy nezávislosti ve stochastické geometrii

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

pawlas@karlin.mff.cuni.cz

Stochastická geometrie je obor, který se zabývá náhodnými geometrickými strukturami. Mezi nejstudovanější modely stochastické geometrie patří prostorový bodový proces (model pro náhodně rozmístěné body v prostoru), proces částic (kolekce náhodných kompaktních podmnožin euklidovského prostoru), náhodně uzavřená množina (model pro oblast, ve které se projevuje nějaká sledovaná událost, může se jednat o sjednocení částic) nebo náhodná mozaika (rozklad prostoru na disjunktní oblasti zvané buňky). Pozorujeme-li dvě případně více realizací nějakého modelu, můžeme se ptát, zda je lze považovat za nezávislé. V přednášce představíme některé postupy pro provedení neparametrického testu nezávislosti.

Všechny uvedené modely lze rozšířit o dodatečné informace (tzv. káty) spojené s jednotlivými body, částicemi nebo buňkami. Pokud by byly káty přiřazeny nezávisle na geometrické strukturu, pak bychom mohli obě složky modelovat zvlášť. Další důležitou statistickou úlohou je proto test nezávislosti kát a příslušné geometrické struktury. I pro tento případ uvedeme některé navržené metody.

MICHAL PEŠŤA

Changing intensities and band bootstrap

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

michal.pestamff.cuni.cz

We deal with a situation such that some phenomena can cause several related events and each event contributes to a different univariate counting process. A collection of these naturally dependent point processes therefore forms a flexible multivariate counting process, where neither stationarity nor independence of interarrival times of the marginal processes are assumed. Our main aim is to detect a structural break of the event's occurrences over time, which means to test whether some (not necessarily all) intensities of the univariate counting processes are subject to change at some unknown time point. We investigate the asymptotic behavior of the test statistic under the null hypothesis as well as under the alternatives. The proposed band bootstrap overcomes the computational curse of dimensionality, while it preserves the dependencies within and between the processes. The validity of this novel resampling technique is proved. As a by-product, a changepoint estimator is introduced and its consistency is provided. Multiple changepoints' detections are designed. The empirical properties are illustrated in a simulation study. The unsupervised data-driven detection procedure is presented through an actuarial problem concerning claims from various insurance lines of business.

MARTINA PETRÁKOVÁ**Some notes about Poisson–Laguerre tessellation with unbounded weights**

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

petrakova@karlin.mff.cuni.cz

The object of our research is the Poisson–Laguerre tessellation, i. e. a random Laguerre tessellation whose generator is a Poisson marked point process with intensity t . We are interested in the asymptotic behaviour (as $t \rightarrow \infty$) of functionals of the tessellation – e. g. the perimeter of the cells, the ratio of volumes of the neighbouring cells – in the case where the weights of the random generator are not uniformly bounded. As it turns out, it is useful to study the behaviour of the distance to the furthest neighbour of a typical point of the point process.

In this contribution we will present some properties of this characteristic, which were derived using the concept of tempered configurations.

ONDŘEJ POKORA**Modelování pomocí degradačních procesů**

PřF MU, ÚMS. Kotlářská 2, CZ–611 37 Brno

pokora@math.muni.cz

Analýza a modelování opotřebení technického systému pomocí diagnostických signálů je jedním ze způsobů hodnocení spolehlivosti a plánování údržby.

Ke stochastickému modelování ve spolehlivosti se využívají degradační procesy. Budeme se věnovat některým takovým modelům, typicky založeným na Lévyho procesech, specificky Wienerově procesu, difuzních procesech nebo gama procesu.

TOMÁŠ POMPA**Metoda podpůrných vektorů pro funkcionální data**

PřF MU, ÚMS. Kotlářská 2, CZ–611 37 Brno

xpompat@math.muni.cz

Klasifikace funkcionálních dat patří mezi základní úlohy moderního odvětví statistiky – analýzy funkcionálních dat, která se zabývá specifickým typem objektů závislých například na časové nebo jiné proměnné. Metoda podpůrných vektorů (ang. *Support vector machines*), jeden z oblíbených nástrojů v klasifikační problematice, je však nejčastěji (podobně jako mnoho jiných metod) definována pouze pro konečně-rozměrná data. V příspěvku nastíníme teoretické pozadí této metody a představíme některé přístupy, jež ji umožňují aplikovat na funkcionální objekty. Tyto přístupy následně porovnáme s jinými metodami určenými pro klasifikaci funkcionálních dat, a to na simulovaných i reálných datech. Pozornost věnujeme zejména datovým souborům, ve kterých se pozorované funkce liší uvnitř jedné klasifikační třídy vertikálním posunutím. Ukážeme, které metody jsou v takové situaci úspěšné, které nikoli, a dále demonstrováme výhodu funkcionálního přístupu využitím derivací pozorovaných křivek ke klasifikaci namísto původních funkcí.

ZUZANA PRÁŠKOVÁ**Použití bootstrapu v panelových datech**

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

praskova@karlin.mff.cuni.cz

Uvažujme panelový autoregresní model prvního řádu s pevnými efekty. Je známo, že OLS odhad autoregresního parametru je vychýlený a vychýlení nemusí zmizet ani asymptoticky při velkém počtu pozorování a panelů. Pro redukci vychýlení lze užít opravené varianty OLS, odhady založené na instrumentálních proměnných (IV) nebo bootstrap. V příspěvku budou prezentovány různé varianty metody bootstrap pro OLS a IV odhady a jejich chování v konečných výběrech.

MICHAELA PROKEŠOVÁ**Časoprostorové shlukové bodové procesy - nehomogenita, separabilita a odhady**

PřF TUL, KMA, Studentská 1402/2, CZ–461 17 Liberec

michaela.prokesova@tul.cz

V příspěvku se budeme zabývat parametrickými modely bodových procesů, které se používají pro modelování časoprostorových bodových vzorků např. v epidemiologii. Konkrétně se jedná o tzv. nehomogenní shot-noise Coxovy procesy. Ty se dobře hodí pro modelování shlukových procesů s nehomogenní (proměnnou) intenzitou prvního řádu, která může být závislá na kovariátách. Po očištění o tuto nehomogenitu prvního řádu nám ale stále zůstává shlukový charakter procesu, tj. body procesu se vzájemně ovlivňují (jsou mezi nimi interakce).

Budeme se zabývat metodou odhadu, která používá minimální kontrast na vhodnou charakteristiku druhého řádu. Vysvětlíme, jaké problémy přináší kombinace nehomogenity a interakcí v procesu, jak jim naše metoda čelí, jaké má asymptotické vlastnosti a pomocí simulační studie prozkoumáme, jak se metoda chová na středně velkých vzorcích.

LUKÁŠ RAČKO**Stability in games with random payoffs**

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

racko@karlin.mff.cuni.cz

The paper deals with non-cooperative games in which the payoff function of its players is influenced by exogenous randomness. The main goal is to provide a general concept of stability in those games because the standard notion of Nash equilibrium is no longer satisfactory. One could find a deterministic equivalent to the game with a random payoff by considering a risk measure and defining a new game with a payoff function adjusted by the risk measure. This, however, causes several problems as the Fundamental Theorem of non-cooperative game theory no longer holds for most of such defined equivalents. Our idea is to loosen the standard concept of the best response to an α -best response which requires the strategy to be the best response only with a certain high probability. Based on this idea we define the α -Nash equilibria and we prove that for every finite game with random payoff non-trivial α -Nash equilibria exist. Moreover, those equilibria characterize equilibria in a broad class of deterministic equivalent games. Finally, we extend the idea of a static game with a random payoff to a game with multiple stages and we show that every finite stochastic game may be represented as a sequential game with a random payoff. In the numerical study, this theory is applied to a management problem of competition of hospitals for vaccines during a pandemic.

HEDVIKA RANOŠOVÁ, DANIEL HLUBINKA**Symetrie funkcionálních dat**

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

ranosova@karlin.mff.cuni.cz

Analýza funkcionálních dat se často zakládá na přebírání a zobecňování postupů mnohorozměrné statistiky, do vektorových prostorů s nekonečnou dimenzí nejsou všechny jednoduše preveditelné. V této práci se zaměřujeme na otázku funkcionální symetrie. V mnohorozměrné statistice (znaménkovou) symetrii můžeme testovat mimo jiné pomocí charakteristické funkce a jejího empirického protějšku. Podobně se u funkcí můžeme opírat o charakteristický funkcionál definovaný na duálu uvažovaného prostoru funkcí. Takto můžeme zkonstruovat test založený na statistice Cramér-von Misesova typu. Použití charakteristického funkcionálu porovnáme s metodou využívající funkcionální hlavní komponenty kombinované s mnohorozměrným testy.

MARTIN ROMAŇÁK**Multiple changepoint detection in tensor data**

MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

romanakm@icloud.com

Tensor data consisting of multivariate discrete-continuous outcomes over the items and across the subjects with longitudinal and cross-sectional dependence are considered. Distributional free detection procedures for changepoints at different locations are proposed, which are of unsupervised learning manner. The bootstrap superstructure is developed in order to overcome computational issues in such a universal setup. The completely data-driven test is presented on real data examples.

SAMUEL ROSA**Optimálny dizajn, grafy a siete**

FMFI UKo, KAMŠ, Mlynská dolina F1, SK – 842 48 Bratislava

samuel.rosa@fmph.uniba.sk

Problematika optimálneho dizajnu experimentu sa venuje výberu experimentálnych podmienok z vopred danej množiny tak, aby bolo možné z experimentu získať čo najviac informácie. V konečnom dôsledku sa táto úloha typicky rieši maximalizáciou zvolenej funkcie vlastných hodnôt informačnej matice (štandardné kritériá ako D-, A- a E-optimalita). V tomto príspevku porozprávame o prepojeniach problému optimálneho dizajnu a teórie grafov (sietí). Dizajny pre isté štatistické modely sa totiž dajú reprezentovať pomocou grafov tak, že informačná matica je vyjadrená pomocou laplaciánu grafu. Takáto reprezentácia umožňuje aplikovať poznatky z oblasti teórie grafov na nachádzanie optimálnych dizajnov. Zaujímavé je, že nezávisle od optimálneho dizajnu v teórii grafov skúmajú tie isté funkcie vlastných hodnôt laplaciánu (zodpovedajúce D-, A-, E-optimalite apod.).

KRISTÍNA SAKMÁROVÁ¹, ARNOŠT KOMÁREK¹, MARTIN OTAVA²**Employing random effects in variance components modelling**¹MFF UK, KPMS, Sokolovská 83, CZ – 186 75 Praha 8²Janssen-Cilag s.r.o., Walterovo nám. 329/1, CZ – 158 00 Praha 5

sakmarova@karlin.mff.cuni.c

The uniformity of dosage units of oral solid dosage (i.e. tablets) refers to the degree of variability of the content of the active pharmaceutical ingredient (API) in the tablet. It is the critical quality attribute, as the amount of the API needs to be tightly controlled in order to deliver to patient the required dose. Both underdosing and overdosing can have severe implications for the patient. Further, stratified content uniformity is often assessed during the process development for a deeper understanding of the variability within the batch, separating location-to-location variability and residual variability. Typical models used for the analysis of the API include fixed run effects model and random run effects model, while both models assume homogenous random location and residual effects. Location and residual errors however do not have to be necessarily iid, which invokes need for population effect and random effect parametrization. For parameter estimation Bayesian methods were used, using brms package. Assuming hierarchical random effects model we provide results for population residual and location-to-location variability among batches.

PÁL SOMOGYI, RADOSLAV HARMAN, SAMUEL ROSA**Optimálne návrhy pre viacrozmerné modely a zovšeobecnený znáhodnený výmenný algoritmus mREX**

FMFI UKo, KAMŠ, Mlynská dolina F1, SK – 842 48 Bratislava

pal.somogyi@fmph.uniba.sk

V príspevku predstavujeme zovšeobecnenie znáhodneného výmenného algoritmu REX ([1]) a venujeme sa algoritmickejmu riešeniu Kieferových problémov optimálneho aproximatívneho navrhovania experimentov, v ktorých elementárne informačné matice, t. j. informačné matice získané z jedného pokusu, môžu mať hodnotu vyššiu ako jedna. Medzi tieto úlohy patria lineárne regresné modely s viacrozmernými pozorovaniami, môžeme však takto formulovať široké spektrum rôznorodých aplikácií. Pomocou nového algoritmu mREX dokážeme nájsť „riedke“ návrhy a pre najpopulárnejšie kritérium D-optimality predstavujeme optimálne výmeny využitím charakteristického polynómu. V numerických štúdiách porovnávame náš nový algoritmus mREX s konkurenčným multiplikatívnym algoritmom a algoritmom YBT na známom probléme Emax modelu s kovariátmi (napr. v [2]) z oblasti klinických skúšok.

Reference

- [1] Harman, R., Filová, L., Richtárik, P. (2020): A Randomized Exchange Algorithm for Computing Optimal Approximate Designs of Experiments, *Journal of the American Statistical Association*, Vol. 115, No. 529, 348-361.
- [2] Magnusdottir, B. T. (2016): Optimal designs for a multiresponse Emax model and efficient parameter estimation, *Biometrical Journal*, Vol. 58, No. 3, 518-534.

GÁBOR SZÚCS, MÁRIA IVANČÍKOVÁ**Bonusovo-malusové systémy v poistení motorových vozidiel**

FMFI UKo, KAMŠ, Mlynská dolina F1, SK – 842 48 Bratislava

Gabor.Szucs@fmph.uniba.sk

Bonusové-malusové systémy (BMS) sú široko využívanými nástrojmi v oblasti poisťovania motorových vozidiel, ktoré slúžia predovšetkým na doladovanie výšky poistného a kvantifikáciu rizikovosti príslušníkov skúmaného poistného kmeňa. Poisťovatelia sa pri takýchto systémoch snažia nastaviť a ponúknuť primerané poistné pre každého poistníka. Klienti (v rámci BMS) sa časom postupne dostávajú na taký stupeň, ktorý zodpovedá ich minulému škodovému priebehu. Idea je taká, aby poistníci s horším škodovým priebehom platili v ďalšom období vyššie poistné, ako klienti s nízkou alebo dokonca nulovou frekvenciou škodovosti. Tento príspevok je venovaný vytvoreniu prehľadu o používaní a atribútoch BMS vo vybraných štátoch sveta. Ďalším cieľom príspevku je predstaviť niektoré matematicko-štatistické ukazovatele, ktoré sa zvyčajne používajú na strane poisťovateľov pri matematickom modelovaní, navrhovaní a porovnávaní tzv. markovovských BMS. Prednáška bude zameraná na indikátory ako napríklad relatívna priemerná stacionárna úroveň poistného v BMS, Loimarantova efektivita BMS, koeficient variácie či miera konvergenencie bonusovo-malusového systému.

Podakovanie: Práca na tomto výskume bola podporená grantmi APVV-20-0311 a VEGA č. 1/0760/22.

JAN VÁVRA**Odhadování názorových pozic politiků z jejich proslovů**

WU Wien, Institute for Statistics and Mathematics, Welthandelsplatz 1, D4, A – 1020 Vienna

vavraj@karlin.mff.cuni.cz

V prednáške bude predstaven bayesovský hierarchický model, ktorý je vystaven na četnostech slovných spojení používaných v databázi proslovů. V těchto tisících dokumentech je nalezeno na desítky diskutovaných témat, u kterých je předpokládán dvojí pohled na věc. V závislosti na tom, jaká spojení používá autor napříč jeho dokumenty, je pro každé z diskutovaných témat odhadnuta jeho názorová pozice na latentní přímce vyjadřující míru jeho souznění s jedním či druhým pohledem. Z aposteriorního rozdělení parametrů modelu (aproximovaného pomocí *variational inference*) lze vyčíst, které téma je nejvíce/nejméně diskutované či názorově bohaté, diskuze na které téma je nejvíce ovlivněna dostupnými informacemi o autorech, dokonce lze nalézt i konkrétní proslovy zodpovědné za extrémní pozice svých autorů. Vše bude ilustrováno na reálných proslovech amerických (případně českých) politiků.

PETR VEJMĚLKA**CDR a jeho využití v rámci clusterování škod v neživotním pojištění**

MFF UK, KPMS, Sokolovská 83, CZ – 186 75 Praha 8

vejmelp@karlin.mff.cuni.cz

V tomto příspěvku představíme využití CDR (Claims Development Result) při rezervování v neživotním pojištění. CDR má klíčovou roli v rámci Evropské směrnice Solvency II a slouží k posouzení schopnosti pojistitele dostát svým závazkům v horizontu jednoho roku. Tuto metriku, která odpovídá skutečným nákladům na pojistná plnění, využijeme v rámci clustrování pojistných škod. Naší snahou je získat co nejpřesnější předpovědi škodních rezerv. Rozdělení škod do více skupin a následná aplikace mnohorozměrných rezervovacích metod může z hlediska predikce pomoci. Navržená metoda je tak alternativou k obvyklejším clusterovacím algoritmům, jako je na např. k-means. V rámci příspěvku bude tato metoda představena a demonstrována jak na reálných, tak generovaných datech.

ONDŘEJ VENCÁLEK, TOMÁŠ FÜRST**Statistika v době post-covidové**

PřF UPOL, KMAAM, 17. listopadu 12, CZ – 771 46 Olomouc

ondrej.vencalek@upol.cz

Navážeme na příspěvek *Statistika v době covidové* z konference Robust 2022. V něm jsme ukázali četné případy špatného použití statistiky při vyhodnocování efektivity vakcín a efektu nefarmakologických opatření.

Během posledních dvou let se problematika šíření onemocnění COVID-19 a efektivity nejrůznějších protiepidemických opatření vytratila z titulních stran mainstreamových médií. Neměla by se však vytratit z popředí zájmu vědecké komunity.

V příspěvku představíme výsledky datových analýz týkajících se efektu očkování proti nemoci COVID-19 na celkovou mortalitu. Diskutovat budeme zejména tzv. *efekt zdravého očkovaného* (Healthy Vaccinee Effect, HVE), viz [1] a [2].

Reference

- [1] Fürst, T., Straka, R., JAnosek, J. Healthy vaccinee effect: a bias not to be forgotten in observational studies on COVID-19 vaccine effectiveness. *Pol Arch Intern Med.* 2024; 134: 16634. doi:10.20452/pamw.16634
- [2] Fürst, T. et al. Does the healthy vaccinee bias rule them all? Association of COVID-19 vaccination status and all cause mortality from an analysis of data from 2.2 million individual health records. *International Journal of Infectious Diseases*, 2024; 142: 106976. ISSN 1201-9712, doi: 10.1016/j.ijid.2024.02.019.

JÁN VESELÝ, JOZEF KOVÁČ, KATARÍNA JANKOVÁ

Diskrétné dynamické systémy s náhodnými impulzmi

FMFI UKo, KAMŠ, Mlynská dolina F1, SK – 842 48 Bratislava

{jan.vesely,jozef.kovac,katarina.jankova}@fmph.uniba.sk

Je známe, že diskretný dynamický systém generovaný funkciou f môže vykazovať rôzne typy správania, a to aj v závislosti od tvaru funkcie f . Nahradením funkcie f impulzovou funkciou g v pevných časoch možno docieľiť napríklad konvergenciu trajektórie k určitému bodu (viď [1]). Takéto systémy sa zvyknú nazývať ako diskretné dynamické systémy s impulzmi. Ak chceme do týchto systémov vniesť náhodnosť, možno to urobiť rôzne. V našom výskume sme uvažovali, že sa impulzová funkcia g bude voliť nie v pevných, ale v náhodných časoch. Zaoberali sme sa podmienkami, ktoré vedia zabezpečiť, že takto vytvorený systém, bude stochasticky stabilný, teda že preňho bude existovať limitné rozdelenie. Výsledky článku [2] o existencii limitného rozdelenia sa dajú priamo použiť v prípade, pokiaľ uvažujeme, že medzi dvomi impulzovými funkciami g môže byť najviac konečne veľa funkcií f . Podarilo sa nám výsledky z článku [2] rozšíriť aj na situáciu, kedy môže byť medzi dvomi funkciami g aj spočítateľne veľa funkcií f . Ďalej z podmienky priemernej kontraktívnosti z článku [3] sa nám podarilo nájsť konkrétnu podmienku pre funkcie f a g , ktorá zaručí existenciu limitného rozdelenia. V prípade, že sú tieto podmienky splnené, uvádzame vyjadrenie limitného rozdelenia diskretného dynamického systému s náhodnými impulzmi, a to pomocou stacionárneho rozdelenia špeciálneho jednoduchého náhodného dynamického systému.

Reference

- [1] Danca M., Fečkan M. and Pospíšil M. (2019). Difference equations with impulses. *Opuscula Mathematica*, 39(1):5–22.
- [2] Díaz L. J. and Matias E. (2018). Stability of the Markov operator and synchronization of Markovian random products. *Nonlinearity*, 31(5):1782-1806.
- [3] Stenflo Ö. (1995). Iterated function systems controlled by a semi-Markov chain. Umeå University. Department of Mathematics.

ONDŘEJ VOZÁR

Vícekritériální hodnocení technik znáhodněného dotazování pro odhady populačního průměru

ČSÚ, Na Padesátém 81, CZ – 110 – 82 Praha 10

FIS VŠE Praha, KSP, Nám. W. Churchilla 3, CZ – 130 67 Praha 3

ondrej.vozar@gov.czso.cz, ondrej.vozar@vse.cz

Rostoucí poptávka po průzkumech zaměřených na kvantitativní charakteristiky (mimo jiné průměr populace) v kontroverzních oblastech, jako jsou korupce, spotřeba návykových látek, daňových úniky, spotřeba drog, vede k zjišťování citlivých proměnných, jako výdaje na drogy nebo nelegální zdroje příjmů, které při přímém dotazování u respondentů nelze zajistit v dostatečné kvalitě. Tím je motivován výzkum technik znáhodněného dotazování pro kvantitativní proměnné. Pro srovnání nově navrhovaných technik s již existujícími navrhuje komplexní multikritériální metodologii hodnocení pro techniky randomizované odpovědi pro průměr populace. Na základě rozsáhlého přehledu v literatuře byly navrženy následující okruhy kritérií: statistické vlastnosti odhadu, implementace a výběr parametrů, zátěž respondentů a důvěryhodnost a ochrana důvěrnosti dat respondentů. Navržená metoda je aplikována na srovnání standardní techniky využívající scramble proměnných a nedávno navrženou techniku dichotomické otázky (Antoch a kol., 2022).

Reference

- [1] Antoch, J., Mola, F., Vozár, O. New Randomized Response Technique for Estimating the Population Total of a Quantitative Variable. *Statistika*, 102(2): 205–227 <https://doi.org/10.54694/stat.2022.11>
- [2] Vozár, O., Marek, L. Multicriteria assessment of randomized response techniques for population mea. *Statistika*, 103(4): 492–503. <https://doi.org/10.54694/stat.2023.32>

GEJZA WIMMER¹, VIKTOR WITKOVSKÝ², JAROSLAV ZŮDA³

Kalibrácia dvoch závaží s použitím referenčného závažia

¹MÚ SAV, Štefániková 49, 814 73 Bratislava, SR

²ÚM SAV, Dúbravská cesta 9, 84104 Bratislava, SR

³ČMI, Okružní 31, 638 00 Brno, ČR

wimmer@mat.savba.sk, witkovsky@savba.sk, jzuda@cmi.cz

Príspevok rieši kalibračnú úlohu sformulovanú na pracovisku Český metrologický institut, Oblastní inspektorát Brno, Oddělení primární etalonáže hmotnosti.

Majme dve závažia M_1 a M_2 a referenčné závažie M_R . Nominálne hmotnosti každého závažia sú 1kg. Porovnaním každých dvoch závaží na komparátore v troch prostrediach so známymi hustotami ρ_1, ρ_2, ρ_3 určte hmotnosti dM_1 (rozdiel medzi nominálnou hodnotou a meranou hodnotou M_1), V_1 (objem meraného závažia M_1 pri teplote 20°C), dM_2 (rozdiel medzi nominálnou hodnotou a meranou hodnotou M_2), V_2 (objem meraného závažia M_2 pri teplote 20°C). Sformulujte matematicko-štatistický (nelineárny) model tejto úlohy a zostrojte iteračný algoritmus na určenie NNLO (najlepších nevychýlených lineárnych odhadov) hľadaných veličín modelu. Celý proces odhadovania realizujeme pomocou programu OEFPIIL.

Pod'akovanie: Práca bola podporená projektmi SK-CZ-RD-21-0109, APVV-21-0216, VEGA 2/0120/24 a interným projektom Českého metrologického institutu podporeným Ministerstvom priemyslu a obchodu ČR č. UTR 24E601201.

VIKTOR WITKOVSKÝ

Poznámka k testovaniu rozdielu mediánov nezávislých populácií pomocou empirických charakteristických funkcií

ÚM SAV, v.v.i., Dúbravská cesta 9, SK – 841 04 Bratislava

witkovsky@savba.sk

Prezentácia je motivovaná rozšírenou, ale mylnou predstavou, s ktorou sa možno stretnúť pri aplikácii štatistických metód v rôznych oblastiach vedy a výskumu, ale aj medzi študentmi matematickej štatistiky, že neparametrický Wilcoxon–Mann–Whitney test (WMW) testuje rovnosť mediánov. Ako je dobre známe, táto predstava je vo veľkej miere nepresná, ako je ukázané napríklad v práci [1], a platí len výnimočných prípadoch, keď distribúcie uvažovaných populácií dvoch skupín sú iba posunuté oproti sebe (t.j., odlišujú sa len v parametri polohy, nie v tvare alebo škále daných distribúcií). Formálne, WMW test testuje nulovú hypotézu $H_0 : p = \Pr(X_1 < X_2) + \Pr(X_1 = X_2)/2 = 0.5$.

Ak je primárnym cieľom analýzy všeobecné testovanie mediánov (napríklad test zhodnosti mediánov), potom WMW test, poprípade aj jeho zovšeobecnená verzia (Kruskal-Wallis test), nie sú spoľahlivé testovacie procedúry, keďže sú citlivé na rozdiely v parametroch škály a symetrie. V tomto prípade ako korektnú metódu možno uvažovať napríklad Moodov test mediánov resp. permutačné testy, ktoré ale taktiež majú svoje obmedzenia.

Táto prezentácia uvažuje ako jednoduchú alternatívu metódu, ktorá je založená na *bootstrap* rozdelení rozdielu výberových mediánov. Novinkou tejto práce je odvodenie presného rozdelenia pomocou charakteristických funkcií. Uvedené metódy a postupy sú implementovateľné pomocou MATLAB balíčka pre charakteristické funkcie [2].

Reference

- [1] Divine, G. W., Norton, H. J., Barón, A. E., Juarez-Colunga, E. (2018). The Wilcoxon–Mann–Whitney procedure fails as a test of medians. *The American Statistician*, 72(3), 278–286.
- [2] Witkovský, V. (2024). CharFunTool: The characteristic functions toolbox (MATLAB). Available at GitHub: <https://github.com/witkovsky/CharFunTool>.

Pod'akovanie: Práca bola podporená projektmi APVV-21-0216, VEGA 2/0096/21 a VEGA 2/0023/22.

IVAN ŽEŽULA, DANIEL KLEIN

Generalization of matrix mean testing in model with special variance structures

PF UPJŠ, ÚMAT, Jesenná 5, SK-04001, Košice 1

ivan.zezula@upjs.sk, daniel.klein@upjs.sk

Matrix multivariate models in practice usually suffer from problem of small sample size. Various special variance structures are used, when possible, to tackle the problem of rank deficiency or numerical instability of the sample variance matrix. One common feature of all these models used in practice is that the variance matrix is decomposed into parts belonging to some commutative quadratic subspace. Thus, almost all previous results can be viewed as special cases of the general model introduced here. We will present some general results concerning mean testing in such model.

Acknowledgment This work was supported by the Slovak Research and Development Agency under the Contract No. APVV-21-0369, and grant VEGA No. 1/0585/24.