

laboratoří a soudních síní. Cílem tohoto revolučního přístupu je nahradit dosud užívanou intuitivní expertní interpretaci naměřených dat či získaných informací plně racionálním, vědecky správným, opakovatelným a přezkoumatelným systémem, který umožňuje správně určit váhu důkazu včetně související nejistoty znaleckých závěrů.

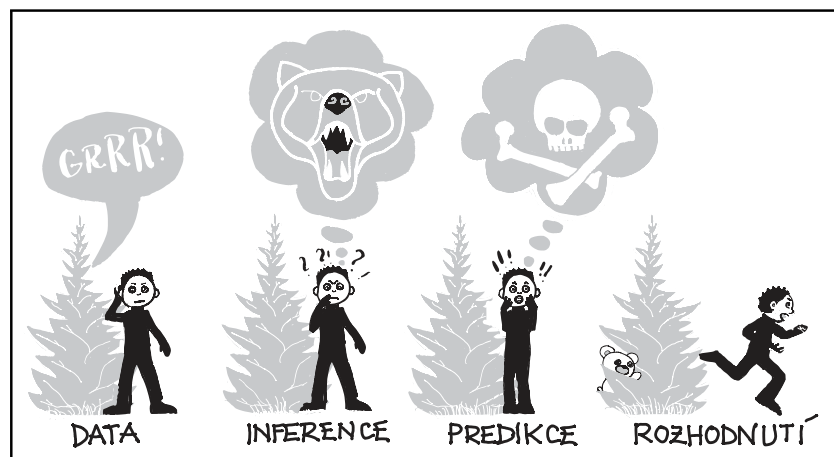
Čtenář se v knize seznámí s **obecnými východisky a postupy bayesovského usuzování a jeho aplikacemi ve forenzních vědách**.

Součástí publikace je **komentovaný český překlad směrnice Evropské sítě forenzních laboratoří (ENFSI), která upravuje oblast znaleckého hodnocení důkazů** a která je formulována na bayesovských principech.

Třetí část knihy obsahuje **dva modelové příklady z oblasti forenzní analýzy DNA a znaleckého zkoumání ručního písma**, které ukazují, jakým způsobem lze směrnici ENFSI promítnout do znalecké praxe.

Kniha je určena každému, kdo ve své praxi naráží na potřebu hodnotit váhu důkazu – na jedné straně tedy znalcům a forenzním expertům, kteří tuto váhu stanovují, a na druhé straně soudcům, státním zástupcům a advokátům, kteří potřebují rozumět tomu, jak znalcem hodnocené důkazy zasadit do kontextu ostatních důkazů.

Kniha má 154 stran, vyšla ve vydavatelství Leges a je k dostání na <https://www.knihyleges.cz/nova-teorie-dkazu-a-akomentovany-cesky-preklad-smernice-enfsi-pro-znalecke-hodnoceni-ve-forenznich-vedach>



SETKÁNÍ ČLENŮ ČSTS

MEETING OF THE CSTS MEMBERS

Výbor společnosti

Proběhla členská schůze České statistické společnosti

Dne 12. 4. 2023 proběhla v prostorech Českého statistického úřadu v Praze členská schůze České statistické společnosti. Schůze se zúčastnilo 26 členů společnosti.

Členskou schůzi zahájil a řídil předseda společnosti Ondřej Vencálek, který v úvodu uvítal účastníky. Následně promluvila Eva Krumpová – 1. místopředsedkyně Českého statistického úřadu (ČSÚ). Po přivítání účastníků hovořila o aktuálním dění v ČSÚ, zejména o realizaci přezkumu (peer review) Národního statistického systému, a dále o probíhajících stavebních úpravách tří budov ČSÚ, včetně sídla úřadu v Praze.

Ondřej Vencálek přednesl *Zprávu o činnosti České statistické společnosti v roce 2022*. Zprávu doplnil fotografiemi z akcí pořádaných ČStS. Následně přečetl *Zprávu o hospodaření České statistické společnosti za rok 2022* vypracovanou Tomášem Löster, který se z účasti na členské schůzi omluvil. Dále pak seznámil přítomné s návrhem rozpočtu na rok 2023. Docentka Blatná podala *Zprávou z revize hospodaření ČStS za rok 2022*.

Volba výboru a předsedy ČStS

Ondřej Vencálek poděkoval za práci ve výboru Petru Volfovi, který se rozhodl znovu do výboru nekandidovat. Poté stručně seznámil účastníky s novou kandidátkou do výboru společnosti, doc. Zuzanou Hübnerovou, působící na VUT v Brně. Členská schůze vybrala pro volbu výboru jako skrutátora Martina Pergela. Hlasovalo celkem 26 členů společnosti. Členy výboru ČStS byli zvoleni (v abecedním pořadí):

- prof. RNDr. Jaromír Antoch, CSc. (MFF UK Praha), 26 hlasů,
- RNDr. Marie Budíková, Dr. (PřF MU Brno), 26 hlasů,
- prof. RNDr. Gejza Dohnal, CSc. (FS ČVUT Praha), 25 hlasů,
- prof. Ing. Jakub Fischer, Ph.D. (FIS VŠE Praha), 26 hlasů,
- doc. RNDr. Kateřina Helisová, Ph.D. (FEL ČVUT Praha), 26 hlasů,
- Ing. Tomáš Hlavsa, Ph.D. (PEF ČZU Praha), 26 hlasů,
- doc. Mgr. Zuzana Hübnerová, Ph.D. (VUT Brno), 26 hlasů,

- Ing. Martina Litschmannová, Ph.D. (FEI VŠB-TU Ostrava), 25 hlasů,
- Ing. Tomáš Löster, Ph.D. (FIS VŠE Praha), 26 hlasů,
- prof. RNDr. Jan Pícek, CSc. (FP TU Liberec), 26 hlasů,
- Mgr. Ondřej Vencálek, Ph.D. (PřF UP Olomouc), 25 hlasů,
- Ing. Ondřej Vozár (ČSÚ, Praha), 26 hlasů a
- RNDr. Marta Žambochová, Ph.D. (FSE UJEP Ústí nad Labem), 25 hlasů.

Předsedou České statistické společnosti byl zvolen Mgr. Ondřej Vencálek, Ph.D. Hlasovalo celkem 26 členů společnosti: 24 hlasů pro.

Změna stanov v článku 8, části (1)

Členská schůze schválila změnu stanov České statistické společnosti v článku 8 týkající se složení předsednictva výboru ČStS. Návrh na změnu předložilo předsednictvo odstupujícího výboru. Změna umožní výboru společnosti zvolit více než jednoho místopředsedu. Motivaci této změny objasnil profesor Jaromír Antoch.

Původní znění prvních dvou vět článku 8 stanov:

(1) Předsednictvo je výkonným orgánem výboru ČStS. Je tvořeno těmito členy výboru: předsedou, místopředsedou, vědeckým tajemníkem, hospodářem a zástupci sekcí.

Nové znění prvních dvou vět článku 8 stanov:

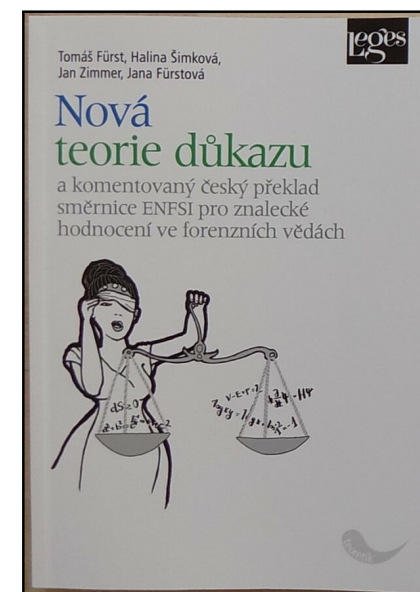
(1) Předsednictvo je výkonným orgánem výboru ČStS. Je tvořeno těmito členy výboru: předsedou, místopředsedy, vědeckým tajemníkem, hospodářem a zástupci sekcí.

Odborná přednáška

V odborné části programu vystoupil doc. RNDr. Zdeněk Karpíšek, CSc., z Ústavu matematiky Fakulty strojního inženýrství VUT v Brně. Téma přednášky bylo fitování diskrétních rozdělení pravděpodobnosti pomocí kvazino-rem.

NOVÁ KNIHA ČLENŮ ČSTS RECENTLY PUBLISHED BOOK

Redakce časopisu



Členové ČStS Halina Šimková a Tomáš Füst (spolu s Janem Zimmerem a Janou Füstovou) v lednu 2022 vydali publikaci „Nová teorie důkazu a komentovaný český překlad směrnice ENFSI pro znalecké hodnocení ve forenzních vědách“ zabývající se mj. znaleckým hodnocením ve forenzních vědách.

Halina Šimková k publikaci uvedla: „Knih není nikdy dost, zvláště když je potřeba něco zgruntu reformovat – třeba celý systém usuzování ve forenzních vědách a hodnocení důkazů soudem. 150 stran ryzí bayesovské logiky (nejen) pro znalce, soudce, státní zástupce a advokáty, ale i pro každého dalšího, kdo chce porozumět tomu, jak ve složitém světě, kde nic není jisté, interpretovat informace, výsledky experimentů či testů a veškerá další data.“

Publikace je prvním česky psaným textem, který uceleně a od samotných základů představuje tzv. **bayesovské hodnocení váhy důkazu** – moderní koncept, který si v posledních letech stále intenzivněji nachází cestu do forenzních

Literatura

- [1] Košťáková, T. (2019): *O složitém jednoduše, aneb, Nebojte se statistiky, nekouše*. Praha: Český statistický úřad, 2019. ISBN 978-80-250-2908-4. cit. 11
- [2] Arlt, J., Arltová, M. a Rublíková, M. (2002): *Analýza ekonomických časových řad s příklady*. 1.vyd. Praha: Vysoká škola ekonomická Praha, 2002. 148 s. ISBN 80-245-0307-7. cit. 13
- [3] *Zpráva o finanční stabilitě 2020/2021*. Praha: Česká národní banka, 2021. ISBN 978-80-88424-02-4. cit. 15
- [4] Česká bankovní asociace k aktuálním datům o odkladu splátek. Tiskové sdělení 7.5.2020. Dostupné z: <https://cbaonline.cz/cba-k-aktualnim-datum-o-odkladu-splatek-kveten>. cit. 15
- [5] Cipra, T. (2006). *Pojistná matematika – teorie a praxe* (2. vydání). Praha: Ekopress. 2006. ISBN 80-86929-11-6. cit. 16
- [6] Mazouch, P., Hulíková, P. a Tesárková, K. (2018). *Kohortní úmrtnostní tabulky v ČR: Metodické aspekty zpracování*. Vysoká škola ekonomická v Praze, Nakladatelství Oeconomica, 2018, 90 s. ISBN 978-80-245-2248-7. Dostupné z: <https://oeconomica.vse.cz/publikace/kohortni-umrtnostni-tabulky-v-cr-metodicke-aspekty-zpracovani> [e-kniha]. cit. 17

ZPRÁVA O ČINNOSTI ČESKÉ STATISTICKÉ SPOLEČNOSTI V ROCE 2022

CZECH STATISTICAL SOCIETY IN 2022: GENERAL REPORT OF ACTIVITIES

Ondřej Vencálek

E-mail: ondrej.vencalek@upol.cz

1. Členská základna

K 31. 12. 2022 měla Česká statistická společnost (ČStS) **203** členů. V průběhu roku 2022 přibýlo 10 nových členů, 2 členům bylo členství ukončeno, celkový počet členů se zvýšil o 8.

2. Členská schůze ČStS

Členská schůze ČStS se uskutečnila dne 3. května 2022 v prostorách ČSÚ v Praze. Zúčastnilo se celkem 23 členů. Byla přečtena a schválena zpráva o činnosti za rok 2021 a zpráva o hospodaření za rok 2021.

Odbornou přednášku na téma *Sčítání lidu 2021 – Mezinárodní kontext a využití administrativních zdrojů dat* pronesl Robert Šanda (ČSÚ).

3. Další akce ČStS

- **Setkání odborné skupiny ČStS k epidemii COVID-19** se uskutečnilo 9. 2. 2022 v prostorech Fakulty managementu VŠE v Jindřichově Hradci. Činnost skupiny rekapitulovali a o aktuálních výzvěch hovořili postupně Ondřej Vencálek, Tomáš Füst a Arnošt Komárek. Akce byla streamována.
- **22. letní škola JČMF Robust** proběhla ve dnech 12. – 17. června 2022 ve Volyni. Bylo prezentováno 54 odborných příspěvků, z toho 5 pozvaných a jeden tutoriál o využití superpočítačů ve statistice. Zúčastnilo se celkem 78 účastníků: 62 z České republiky, 12 ze Slovenska, 2 ze Švýcarska, 1 z USA a 1 z Nigérie.
- **Statistické dny v Telči** se uskutečnily 23. – 25. září 2022 v Univerzitním centru MU Brno v Telči. Celkem 32 účastníků vyslechlo 10 odborných příspěvků.

- **Mikulášský statistický den** byl pořádán dne 6. prosince 2022 v prostorech VŠE v Praze. Akce byla streamována. Zúčastnilo se 12 účastníků v Praze a dalších 24 online. Byly předneseny celkem čtyři odborné příspěvky (tři z oblasti statistiky).

4. Spolupráce s Českým statistickým úřadem

Český statistický úřad je dlouhodobě partnerem ČStS a podporuje její činnost například zajištěním tisku Informačního bulletinu ČStS, poskytnutím sídla společnosti, prostor pro konání členských schůzí ČStS a publicity v časopisech ČSÚ (např. rozhovory s členy v časopise *Statistika a My*, informace o společných akcích) a na sociálních sítích.

5. Členství v mezinárodních organizacích

Naše společnost je od roku 2011 členem Federace evropských národních statistických společností FENStatS (The Federation of European National Statistical Societies). Členové ČStS jsou průběžně informováni o dění ve FEN-StatS. Na členský příspěvek ve výši 200 EUR jsme úspěšně čerpali dotaci prostřednictvím Rady vědeckých společností ČR. Podobně hodláme postupovat i v roce 2023.

6. Členství v Radě vědeckých společností ČR

ČStS je členem Rady vědeckých společností ČR (RVS). Plenární zasedání RVS se konalo 27. dubna 2022 v Praze. ČStS na jednání zastupoval Jakub Fischer. Proběhla volba Výkonného výboru RVS. ČStS je ve výboru zastoupena Jakubem Fischerem. Do RVS byly přijaty další dvě společnosti. Na konci roku byla RVS odevzdána zpráva o činnosti naší společnosti.

7. Další činnost ČStS

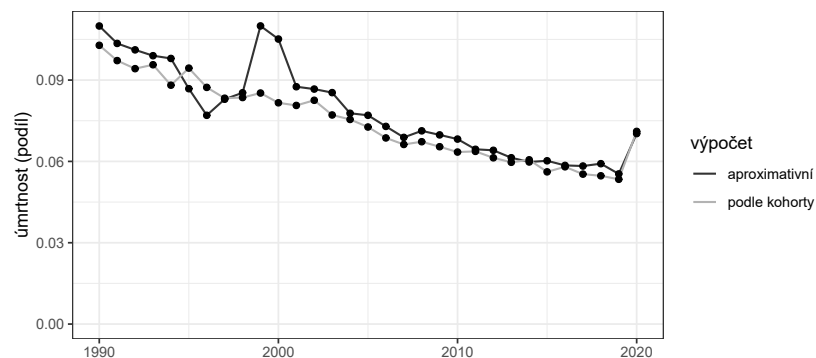
- Byla vydána 2 řádná čísla Informačního bulletinu ČStS.
- Byla pravidelně aktualizována webová stránka společnosti – bylo zveřejněno celkem 23 příspěvků v sekci novinky a byly provedeny úpravy na dalších stránkách.
- Předseda společnosti reprezentoval ČStS na slavnostním semináři *Dílo Otakara Borůvky v kontextu současné teorie grafů*, který se uskutečnil 30. května 2022 v Brně.

Rok ($z+x$)	Ročník (z)	Zemřelí			Populace		Úmrtnost	
		/1/	/2/	/3/	/4/	/5/	/6/	/7/
1994	1914	2162	4218	4565	45705	46593	0,0881	0,0980
1995	1915	1912	3291	3968	32946	45705	0,0944	0,0868
1996	1916	1158	2235	2537	24442	32946	0,0873	0,0770
1997	1917	950	1916	2027	22047	24442	0,0833	0,0829
1998	1918	915	1881	1881	21596	22047	0,0836	0,0853
1999	1919	1409	3291	2375	37209	21596	0,0852	0,1100
2000	1920	2030	3961	3912	46501	37209	0,0816	0,1051
2001	1921	2140	4320	4071	51429	46501	0,0806	0,0875
2002	1922	2278	4540	4458	52721	51429	0,0825	0,0867
		$\begin{matrix} /1/ & z+x & D_x^z \\ /2/ & z+x & D_x^z + z+x+1 & D_x^z \\ /3/ & z+x & D_x^z + z+x & D_x^{z+1} \\ /4/ & z+x & P_x^z \\ /5/ & z+x-1 & P_x^{z-1} \\ /6/ & \text{Dle vztahu (2)} \\ /7/ & \text{Aproximativní} \end{matrix}$						

Tabulka 1: Data o vývoji úmrtnosti ve věkové skupině 80 let v letech 1994 až 2002.

zejména mezi ročníky 1918 a 1919, a také 1919 a 1920. Analýza dat o úmrtnosti publikovaných v demografických ročenkách České republiky⁴ odhalila, že domnělé zvýšení úmrtnosti ve skupině 85+ kolem roku 2004 bylo jen artefaktem způsobeným nesprávným (aproximativním) výpočtem úmrtnosti. Selhání výpočtu, který jindy dává poměrně dobrý odhad pravděpodobnosti úmrtí v určitém věku, bylo způsobeno tím, že do věkové skupiny 85+ v daném čase vstupovaly populačně výrazně silnější poválečné ročníky (1919, 1920). Více než sto let stará mimořádná událost (první světová válka) tak zřejmě zmátla autora onoho grafu z jara 2021, o němž byla řeč na začátku.

Obrázek 5 znázorňuje vývoj (odhadnuté) míry úmrtnosti ve věkové skupině 80 let v období let 1990 až 2020. Poznamenejme, že odhad pomocí vztahu (2) pro rok t využívá data za roky t a $t + 1$. V současnosti jsou k dispozici data do roku 2021⁵, proto výpočet je proveden do roku 2020. Z grafu je patrný velký rozdíl odhadu úmrtnosti při použití vztahu (2) a aproximativní metody v letech 1999 a 2000. Jsou to roky, kdy své osmdesáté narozeniny slavily ročníky 1919 a 1920, tedy poválečné ročníky. Ty byly populačně výrazně silnější než ročníky válečné, jak je možné se přesvědčit z tabulky 1. Aproximativní výpočet v tomto případě vede k nesprávnému nadhodnocení míry úmrtnosti v daných letech.



Obrázek 5: Úmrtnost (vyjádřená jako podíl) ve věku 80 let. Srovnání odhadu vypočteného podle vztahu (2) a odhadu aproximativního.

⁴https://www.czso.cz/csu/czso/casova_rada_demografie a <https://www.czso.cz/csu/czso/demograficka-rocenka-ceske-republiky-2021>

⁵Demografická ročenka ČR – 2021 byla publikována 3. 10. 2022.

- Zástupci výboru ČStS Tomáš Hlavsa a Ondřej Vencálek se zúčastnili 21. slovenské statistické konference konané ve dnech 13. – 14. září 2022 v Žilině.
- Předseda společnosti se zúčastnil XIII. kolokvia *Národního metodického kabinetu Matematika a její aplikace* konaného 4. října 2022 online, kde mj. představil činnost ČStS.

8. Akce chystané ČStS v roce 2023

- Členská schůze ČStS – 12. 4. 2023.
- Statistické dny 2023 na břehu Tiché Orlice, 19. – 21. 5. 2023.
- Prastan 2023 – spolu se Slovenskou statistickou a demografickou společností – září 2023.
- Mikulášský den ČStS – kolem 6. 12. 2023.

V Olomouci dne 23. 1. 2023, upraveno 27. 3. 2023

Ondřej Vencálek
předseda ČStS

TŘI OŘÍŠKY PRO STATISTIKY

THREE GIFTS FOR STATISTICIANS

Ondřej Vencálek

Adresa: Katedra matematické analýzy a aplikací matematiky, PřF, Univerzita Palackého v Olomouci, 17. listopadu 12, 771 46 Olomouc

E-mail: ondrej.vencalek@upol.cz

Abstrakt: Článek pojednává o třech drobných problémech spojených s analýzou časových řad. Všechny tři oříšky nějakým způsobem souvisí s dobou bohatou na mimořádné události. První problém se týká tvrzení o snížení nejistoty hodnoty podílového fondu pomocí pravidelného investování ilustrovaného krajně manipulativním grafickým znázorněním historie „výkonnosti fondu“. Druhý problém upozorňuje na obtíže meziročního srovnání v současné (v mnoha ohledech mimořádné) době. Týká se nedávného vývoje obžemu „nových hypoték“. Třetí problém se týká vývoje míry úmrtnosti. Ukazujeme, jak mimořádná doba (tentokrát první světová válka) o sto let později „zneplatnila“ použití aproximativního výpočtu míry úmrtnosti.

Klíčová slova: oříšek, fond, hypotéka, úmrtnost.

Abstract: The article discusses three minor (but surprising) problems associated with time series analysis. All three problems are somehow related to an emergency-rich era. The first problem concerns the claim of reducing the uncertainty of the fund's value through regular investing which is illustrated by an extremely manipulative graphical representation of the history of “fund performance”. The second problem draws attention to the difficulties of year-on-year comparisons in the current (in many respects extraordinary) period. It refers to recent developments in the volume of “new mortgages”. The third problem concerns the development of the mortality rates. We show how an extraordinary period (the World War I in this case) a hundred years later “invalidated” the use of the approximate calculation of the death rate.

Keywords: problem, fund, mortgage, mortality.

Když se na počátku roku 2020 začaly objevovat zprávy o šíření koronavirové infekce v čínské provincii Wu-chan, málokdo v České republice si dovedl představit, že již v březnu dojde u nás k uzavření škol a vyhlášení nouzového stavu. Současná doba (doba od začátku roku 2020) je v kontextu vývoje nejen České republiky v posledních desetiletích jistě v mnoha ohledech mimořádná.

roku už se nedožili (zemřeli v roce 2021 po 80. narozeninách), a na ty, kteří se dožili roku 2022, ale zemřeli v něm před svými 81. narozeninami.

Zavedme následující označení:

- ${}_tD_x^z \dots$ počet lidí z kohorty z zemřelých ve věku x v roce t ,
- ${}_tP_x^z \dots$ počet lidí z kohorty z , kteří mají na konci roku t věk x .

Celkový počet lidí z kohorty z zemřelých ve věku x vypočteme jako součet ${}_{z+x}D_x^z + {}_{z+x+1}D_x^z$. Všechny lidi z kohorty z , kteří se dožili věku x dožili, můžeme opět rozdělit do dvou skupin: na ty, kteří se dožili konce roku $z+x$ (jejich počet je ${}_{z+x}P_x^z$), a na ty, kteří se sice věku x v průběhu roku $z+x$ dožili, ale už se nedožili konce roku $z+x$ (jejich počet je ${}_{z+x}D_x^z$). Docházíme tedy k odhadu míry úmrtnosti pro kohortu z a věk x v podobě podílu:

$$\frac{{}_{z+x}D_x^z + {}_{z+x+1}D_x^z}{{}_{z+x}P_x^z + {}_{z+x}D_x^z}. \quad (2)$$

Tento výpočet je uveden např. v publikaci [6], na str. 45, kterou doporučujeme zájemcům o problematiku konstrukce kohortních úmrtnostních tabulek. V této publikaci se čtenář rovněž poučí o názorné pomůcce známé jako demografická síť či Lexisův diagram, viz [6], str. 17–19.

Dodejme, že údaje o počtu zemřelých podle věku a ročníku narození jsou k dispozici v demografických ročenkách ČR zveřejňovaných pravidelně Českým statistickým úřadem.

3.3. Aproximativní výpočet a jeho selhání

Z výše uvedených úvah je patrné, jaké nedostatky má dříve zmíněný aproximativní výpočet. Především ne všichni, kteří zemřeli v daném roce v daném věku, měli tento věk již na počátku roku. Tito jsou započtení v čitateli, nikoliv však ve jmenovateli aproximativního vzorce. Tím se ovšem vytrácí interpretace, kdy ve jmenovateli by měli být všichni „v riziku“ (exponovaní).

V běžné situaci můžeme docela dobře předpokládat ${}_{z+x}D_x^z \approx {}_{z+1+x}D_x^{z+1}$. Při použití této aproximace ve vzorci (2) dotaneme v čitateli počet všech zemřelých v roce $z+1+x$ ve věku x (tedy čísel aproximativního vzorce). S výjimkou nejvyšších věkových skupin můžeme pak jmenovatel aproximovat členem ${}_{z+x}P_x^z$, který je v čase poměrně stabilní, tj. ${}_{z+x}P_x^z \approx {}_{z+x+1}P_x^{z+1}$. Tím je ospravedlněn výše uvedený aproximativní výpočet.

Z výše uvedených úvah je patrné, že velký rozdíl ve velikosti populačních ročníků povede k velké odchylce aproximativní hodnoty od hodnoty vypočtené podle vztahu (2). Velké rozdíly ve velikosti populačních ročníků byly

V tomto článku budeme odhadovat *pravděpodobnost úmrtí ve věku x* , tedy pravděpodobnost toho, že jedinec, který je na živu ve věku x , zemře před dosažením věku $x + 1$. Takto chápe pravděpodobnost úmrtí ve věku x např. [5]. Všimněme si, že jde o podmíněnou pravděpodobnost úmrtí v daném věku x za podmínky, že se člověk věku x dožil.

Při odhadu pravděpodobnosti úmrtí ve věku x si musíme uvědomit, že se tato pravděpodobnost v čase mění (lidé se dožívají vyššího věku). Její odhad proto vztáhneme vždy k určitému roku (označme rok t).

3.2. Odhad pravděpodobnosti úmrtí ve věku x

První nápad, jak odhadnout pravděpodobnost úmrtí ve věku x , je vydělit počet zemřelých v daném roce ve věku x počtem všech obyvatel, kteří na začátku tohoto roku byli ve věku x . Označme:

- ${}_tD_x \dots$ počet lidí zemřelých ve věku x v roce t ,
- ${}_{t-1}P_x \dots$ počet lidí, kteří mají na konci roku $t - 1$ (resp. na začátku roku t) věk x .

Odhad pravděpodobnosti úmrtí ve věku x v roce t můžeme vypočítat jako podíl

$$\frac{{}_tD_x}{{}_{t-1}P_x}. \quad (1)$$

Výhoda tohoto postupu je, že oba údaje jsou pro ČR poměrně snadno dostupné. Tento postup, který byl použit při konstrukci výše zmíněného grafu, budu označovat za aproximativní. Poznamenejme, že např. Český statistický úřad uvádí specifické míry úmrtnosti, které představují počet zemřelých v dané věkové skupině, připadající na 1000 obyvatel středního stavu (tedy stavu nikoliv na začátku roku, ale k 1. červenci) ve stejné věkové skupině. Hlubší zamyšlení nás povede k úpravě (vylepšení) tohoto odhadu pravděpodobnosti úmrtí ve věku x .

Lidé, kteří na konci určitého roku t mají věk x , se všichni narodili ve stejném roce (např. všichni lidé, kteří na konci roku 2021 měli 80 let, se narodili v roce 1941). Vezměme tedy všechny lidi narozené ve stejném roce, tj. tvořící stejnou kohortu označenou rokem narození z . V této kohortě se někteří lidé věku x nedožili, někteří zemřeli ve věku x a někteří se dožijí (či dožili) vyššího věku než x . Zajímavé by nás tedy podíl lidí narozených v daném roce z (1941), kteří zemřeli ve věku x (80 let), ve skupině všech lidí stejného ročníku z , kteří se věku x dožili.

Lidi z kohorty z , kteří zemřeli ve věku 80 let, přitom můžeme rozdělit do dvou skupin: na ty, kteří v roce 2021 oslavili osmdesáté narozeniny, ale konce

To s sebou nese řadu problémů při analýze nejrozličnějších časových řad. Nabízím čtenářům zamyšlení nad několika takovými drobnými problémy (oříšky), se kterými jsem se v nedávné době setkal.

1. První oříšek – volatilita výnosu fondu

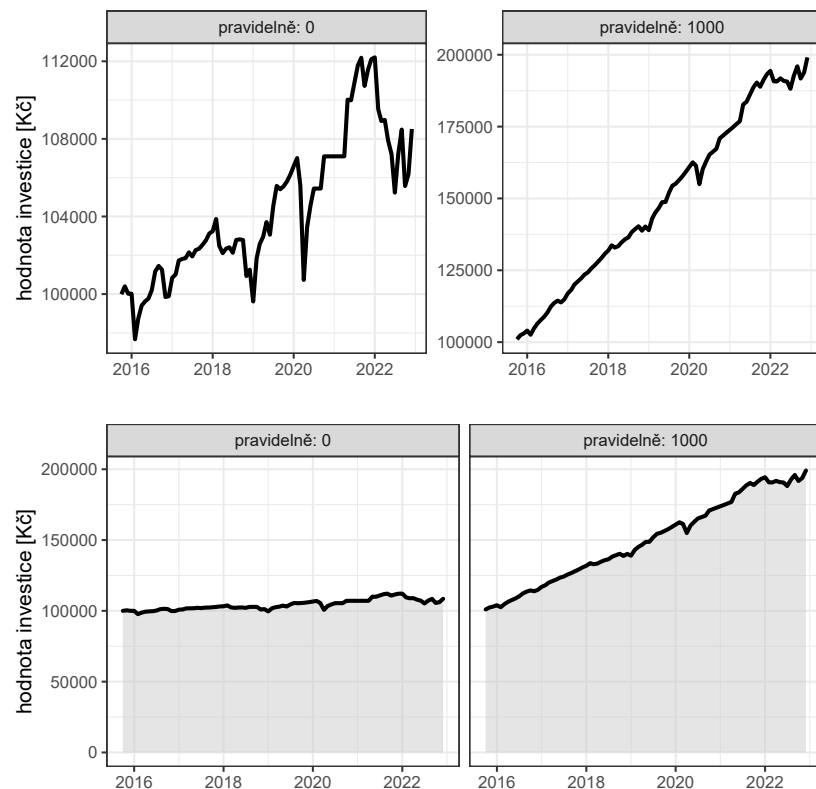
Při jedné návštěvě banky jsem byl poučen o volatilitě výnosnosti podílových fondů a dostal jsem dobrou radu, jak tuto volatilitu potlačit. Nabízím zde tento „zaručený recept“ k uvážení všem čtenářům. Nejmenovaná banka pro snadné pochopení tohoto receptu nabízí „kalkulátor výkonnosti podílových fondů“¹. V něm se klient může snadno přesvědčit, jak výhodnou strategií k potlačení volatility je pravidelné investování.

Představme si, že jsme 1.10.2015 do nejmenovaného fondu jednorázově investovali 100 tisíc korun. Zapomeňme pro tuto chvíli (i všechny chvíle další) na poplatek s tímto úkonem spjatý. Vývoj (nominální) hodnoty investované částky v následujících letech (až do 1.12.2022) vidíme na obrázku 1 vlevo nahoře. Je z něj patrné, jak velmi hodnota investice v čase kolísá. Je například patrné, že nominální hodnota investice byla po třech letech stejná jako na začátku. Také vidíme, že po 7 letech se nominální hodnota zvýšila přibližně o 8 procent, což odpovídá výnosu jen něco málo přes 1 % p.a. Pokud bychom ovšem do tohoto fondu kromě počáteční investice po celou dobu měsíčně investovali 1000 Kč, bude (nominální) hodnota naší investice celkem stabilně růst, viz obrázek 1 nahoře vpravo.

Na obrázku 1 dole jsou vykresleny stejné časové řady, ovšem se stejným měřítkem na svislé ose. Zatímco při srovnání obrázků v horní části má člověk v prvním okamžiku dojem, že pravidelným investováním dochází k významnému snížení volatility, při pohledu na spodní část obrázku by ho nic takového zřejmě ani nenapadlo.

Trič (zdánlivého) snížení volatility výnosu fondu můžeme přidat do dlouhého seznamu „optických klamů“ založených na různém měřítku porovnávání grafů. Čtenáři přenecháváme na zvážení, zda by přece jen nenašel nějaký argument obhajující tvrzení o nižší nejistotě při pravidelném investování.

¹Zvědavý čtenář odhalí, o jakou banku se jedná, když do svého oblíbeného vyhledávače zadá právě tuto čtveřici slov (v uvozovkách).



Obrázek 1: Vývoj hodnoty investice do nejmenovaného podílového fondu od 1.10.2015 do 1.12.2022 při jednorázové investici ve výši 100 tisíc Kč na začátku (vlevo) a při stejné počáteční investici a pravidelné měsíční investici 1000 Kč (vpravo). Grafy nahoře používané bankou ukazují, jak pravidelné investování snižuje volatilitu. Grafy dole používající stejné měřítko na svislé ose nejsou pro přesvědčení klienta o tomto fenoménu vhodné.

Peak v kategorii *ostatní nová ujednání ke stávajícím smlouvám* a následně i v celkovém součtu přes všechny tři kategorie ve skutečnosti vznikl v souvislosti s přijetím moratoria na splátky úvěrů a hypoték, viz dokument ČNB *Zpráva o finanční stabilitě 2020/2021* [3], str. 84–85. Toto moratorium schválila vláda ČR 1.4.2020 na návrh Ministerstva financí ČR. Dlužníci z řad fyzických osob i firem mohli podle své volby přerušit splácení na tři nebo šest měsíců. Splátky mohly být odloženy poté, co dlužník oznámil tento záměr svému věřiteli, a prohlásil, že k tomuto kroku přistupuje z důvodů negativního ekonomického dopadu pandemie koronaviru. Tyto důvody však nemusel nijak prokazovat. Více se o tomto tématu lze dozvědět na stránkách České bankovní asociace, která průběžně data o vývoji zájmu o odklad splátek v daném čase sledovala a komentovala, viz např. [4].

Při zkoumání determinantů objemu poskytnutých hypoték v České republice lze snadno sklouznout s jednoduchým, ale nesprávným vysvětlením peaku v časové řadě na obrázku 2. Mohlo by se například zdát, že prudký nárůst proměnné Úvěry v dubnu a květnu roku 2020 byl způsoben zrušením limitu DTI (Debt to Income) a uvolněním limitu LTV (Loan to Value) a DSTI (Debt Service to Income), ke kterým přistoupila ČNB 1. dubna 2020. Jak jsme vysvětlili, peak neodrážel zvýšený zájem o poskytnutí nové hypotéky, ale vznikl přijetím nového ujednání (týkajícího se odkladu splátek) u řady již existujících smluv.

Tento příklad nás vede k obecnému varování před unáhlenými závěry bez detailního pochopení původu dat, včetně způsobu sběru dat a definic jednotlivých proměnných obsažených v datech.

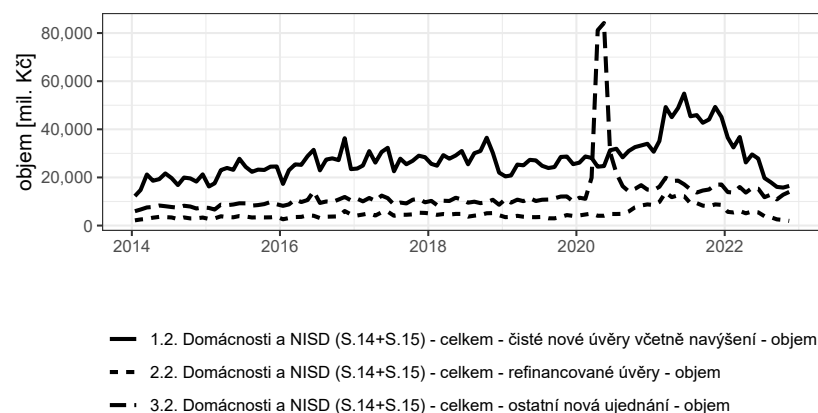
3. Třetí oříšek – úmrtnost

3.1. Podivný graf z jara 2021

Na jaře roku 2021 jsem zahlédl graf vývoje úmrtnosti v posledních třech desetiletích v šesti věkových skupinách obyvatelstva (0–14, 15–39, 40–64, 65–74, 75–84, 85+ let). Tento graf ukazoval jednak poměrně setrvalý pokles úmrtnosti ve všech věkových skupinách (který implikuje skutečnost, že „lidé se dožívají vyššího věku“), ovšem ve třech nejstarších věkových skupinách byl také patrný nezanedbatelný nárůst úmrtnosti v roce 2020. Ve věkové skupině 85+ byl však podobně dramatický nárůst také kolem roku 2004. Že by v tom období došlo k nějaké pozapomenuté či nepovšimnuté epidemii v této věkové skupině? Než budeme pátrat, co se v daném období odehrálo, připomeňme, jak vůbec „měřit“ úmrtnost.

- čisté nové úvěry včetně navýšení (skutečně nové úvěry),
- refinancované úvěry,
- ostatní nová ujednání ke stávajícím smlouvám.

Z grafu je přitom patrné, že k extrémnímu nárůstu došlo v dubnu a květnu 2020 jen v kategorii *ostatní nová ujednání ke stávajícím smlouvám*. Nešlo tedy o extrémní objem skutečně nových úvěrů.



Obrázek 4: Objem čistých nových a nově ujednaných úvěrů poskytnutých bankami domácnostem a neziskovým institucím sloužícím domácnostem (NISD) v ČR v mil. Kč dle ČNB.

Co se skrývá za nárůstem v kategorii *ostatní nová ujednání ke stávajícím smlouvám*? Metodický list k příslušnému okruhu dat (Úrokové sazby měnových finančních institucí) dostupný na stránkách ČNB³ nabízí jen velmi obecné vymezení této kategorie. Ta „obsahuje úvěry, které byly již dříve poskytnuty vykazující bankou a došlo u nich pouze k novému ujednání (např. refixace úrokové sazby u úvěrů na bydlení). Dále se v této položce vykazují konsolidované úvěry, jejichž původním věřitelem je jak jiná/jiné instituce tak i vykazující banka nebo případně pouze vykazující banka. Případné navýšení úvěru je analogicky jako u refinancovaných úvěrů součástí kategorie čistých nových úvěrů.“

³https://www.cnb.cz/cnb/STAT.ARADY_PKG.PARAMETRY_SESTAVY?p_strid=AAABAC&p_sestuid=60323&p_tab=2&p_lang=CS

2. Druhý oříšek – hypotéky

2.1. Meziroční srovnávání v době mimořádných událostí

Vývoj nejrůznějších ukazatelů často bývá prezentován v podobě procentuálního vyjádření meziročního či meziměsíčního růstu, případně poklesu. Výhody a úskalí meziročního, resp. meziměsíčního, srovnání shrnula v knize *O složitém jednoduše* Tereza Košťáková [1]: „Statistici i ekonomové obvykle přednostně uvádějí meziroční srovnání. Důvod je ten, že většina ukazatelů vykazuje v průběhu roku sezónnost, která znemožňuje smysluplné porovnání různých období roku navzájem. [...] Na druhé straně nevýhodou meziročního srovnání je skutečnost, že dané období srovnáváme s něčím, co už se stalo velmi dávno (před 12 měsíci). Tato srovnávací základna (hodnota ukazatele před rokem) mohla být například mimořádně vychýlena nahoru nebo dolů.“

O tom, jak opatrně je třeba vnímat informaci o meziročním poklesu či růstu, svědčí následující příklad. Z dat České národní banky (ČNB) vyplývá, že ve druhém čtvrtletí roku 2021 zaznamenal objem nově poskytnutých hypotečních úvěrů meziroční pokles o 43 %. O jak závažný pokles šlo? Mohlo by se zdát, že situace ve druhém čtvrtletí nebyla vůbec dobrá. Bližší pohled na data však tuto interpretaci jistě upraví.

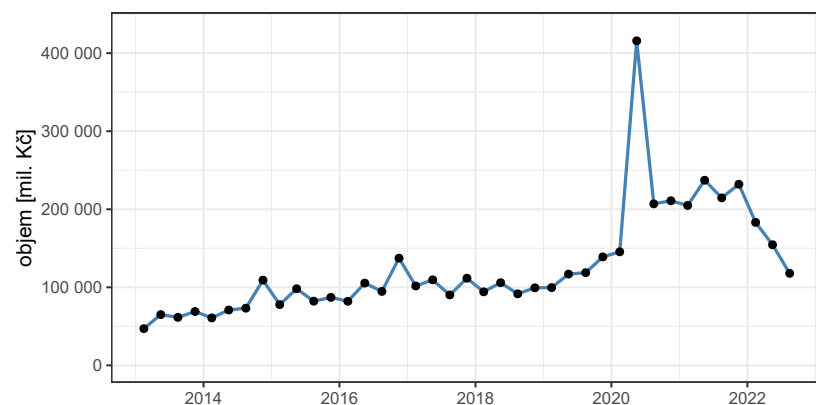
Z grafu na obrázku 2 je patrné, že ve skutečnosti byla situace zcela mimořádná nikoliv v roce 2021, ale ve druhém čtvrtletí roku 2020, kdy došlo k extrémnímu krátkodobému zvýšení objemu nově poskytnutých hypotečních úvěrů (kdy objem představoval téměř 416 mld. Kč). Hodnota ze druhého čtvrtletí roku 2021 (237 mld. Kč) je ve skutečnosti druhou nejvyšší hodnotou za sledované období (roky 2013 až 2022). Příčinu výkyvu z jara 2020 objasníme v kapitole 2.3.

Tento příklad názorně ilustruje varování, které Tereza Košťáková v [1] vyjádřila následujícími slovy: „U meziročního srovnání musíme stále po očku sledovat i to, co se dělo před rokem, abychom správně zhodnotili, do jaké míry meziroční index vypovídá o současných trendech.“ Dodejme, že toto varování nabývá na síle v době mimořádných událostí.

2.2. Záhada zmizelého peaku

Data o objemu nově poskytnutých hypotečních úvěrů jsou k dispozici ve veřejné databázi ARAD², která je součástí informačního servisu České národní banky. Některé údaje mohou být z této databáze získány a zobrazeny i na úrovni měsíčních hodnot. Nese to však s sebou neočekávaný problém. Po-

²https://www.cnb.cz/cnb/STAT.ARADY_PKG.STROM_KOREN

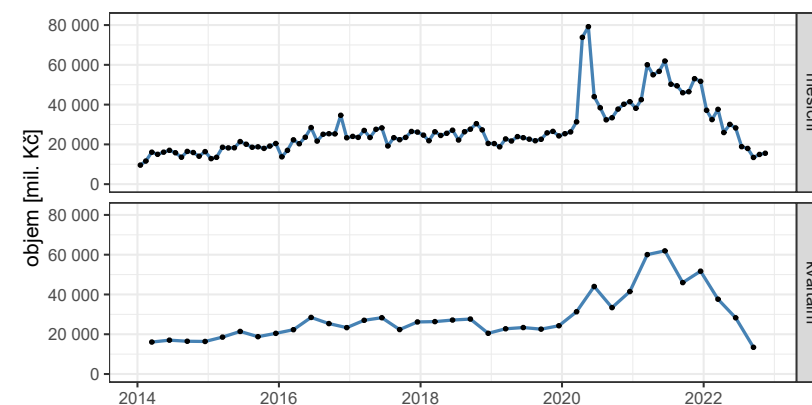


Obrázek 2: Objem nově poskytnutých hypotečních úvěrů v mil. Kč dle ČNB (čtvrtletní data).

kud při výběru dat změním volbu parametru *periodicity* z možnosti *měsíční* na možnost *kvartální*, nezískáme součty za jednotlivá čtvrtletí (kvartály), ale pouze vyfiltrujeme měsíční údaje vždy za poslední měsíc v kvartálu. *Získáme tak poměrně kuriózní časovou řadu měsíčních údajů, ovšem jen některých.* To je velmi matoucí zejména v grafu, kde jsou za rok vždy čtyři pozorování, ale nikde není upozornění, že na svislé ose jsou hodnoty odpovídající údajům vždy jen za jeden (poslední) měsíc čtvrtletí.

Měsíční data jsou k dispozici např. pro veličinu *Nové úvěry na bydlení poskytnuté bankami a stavebními spořitelny obyvatelstvu*, viz obrázek 3 nahoře. Z tohoto grafu je patrné, že k extrémnímu výkyvu došlo v dubnu a květnu roku 2020. Peak odpovídající těmto dvěma měsícům však „zmizí“, pokud zvolíme kvartální periodicitu (obrázek 3 dole). Jak jsme vysvětlili výše, je to dáno tím, že v dolním obrázku jsou zobrazena data vždy za poslední měsíc kvartálu, což v případě druhého čtvrtletí znamená, že jde o data za červen.

Nedorozumění vzniká z nepochopení povahy časové řady. Zde jde o typickou intervalovou časovou řadu, v níž hodnoty ukazatele závisí na délce časového intervalu sledování. Hodnoty takové časové řady lze agregovat sčítáním přes několik časových intervalů (např. měsíční pozorování sčítáme přes tři měsíce kvartálu a vznikají tak čtvrtletní hodnoty). Rozdíl mezi intervalo-



Obrázek 3: Objem nových úvěrů na bydlení poskytnutých bankami a stavebními spořitelny obyvatelstvu v mil. Kč dle ČNB. Srovnání grafů při volbě měsíční (nahoře) a kvartální (dole) periodicity.

vou a tzv. okamžikovou časovou řadou je přitom typicky uváděn hned v úvodu základních kurzů analýzy časových řad, viz např. kapitola 1.1 monografie [2].

2.3. Pátrání po původu peaku

Položíme-li si otázku, jak predikovat vývoj časové řady zobrazené v grafu 2, bylo by dobré nejprve pochopit, co se stalo ve druhém čtvrtletí roku 2020, v němž byla zaznamenána neobvykle vysoká hodnota. Připomeňme, že v březnu roku 2020 byly v České republice potvrzeny první případy nákazy virem SARS-CoV-2 (1. 3.), byly uzavřeny všechny školy (11. 3.) a byl vyhlášen nouzový stav (14. 3.). Konec prvního a začátek druhého čtvrtletí roku 2020 byla tedy bezesporu doba mimořádných událostí, přinejmenším v kontextu několika uplynulých desetiletí. O to překvapivější je, že v takto nejisté době dramaticky stoupl *objem nově poskytnutých hypotečních úvěrů*.

K vysvětlení tohoto překvapivého jevu nás může přivést bližší prozkoumání datové sady zveřejněné v databázi ARAD. Na obrázku 4 jsou zobrazeny *čisté nové a nově ujednané úvěry poskytnuté bankami domácnostem a neziskovým institucím sloužícím domácnostem (NISD) v ČR* rozdělené do tří kategorií: