**Komentovaná rešerše**

Téma: Benfordův zákon v informační vědě

Tvůrce: David Jiří Šlosar

Vypracováno: 15.4. – 18.4. 2021

Popis rešeršního záměru: Smyslem této rešerše je vytvoření představy o stavu poznání v oblasti Benfordova zákona s důrazem na informační vědu. Obsahově relevantní výsledky budou použity pro dizertační práci autora. Obzvláště hodnotné jsou výsledky, které pojednávají o aplikaci Benfordova zákona ve scientometrii.

Zdroje: Web of Science, Scopus, Dimensions, Microsoft Academics, Google a přehledový web <http://www.benfordonline.net/>

Nejprve bylo provedeno vyhledávání na webu <http://www.benfordonline.net/>. Tento web je zájmovým a přehledovým webem několika vědců, kteří se Benfordovým zákonem zabývají a provádí vlastní rešerše, aby kumulovali literaturu o tomto jevu.

První iterace prohledávání tohoto webu byla čistě mechanická. Autor zvolil chronologické řazení a pročítal názvy jednotlivých článků od nejstaršího k nejnovějšímu. Hrubou a subjektivní selekcí tak vybral 30 článků níže citovaných. Ve druhé iteraci autor přes funkci prohlížeče vyhledávání hledal klíčová slova „Biblio“, „Sciento“, „Science“, „metric“ a „Citation“. Všechny pro autora relevantní články získané druhou iterací již byly obsaženy v první iteraci.

Seznam článků pro <http://www.benfordonline.net/> a jejich slovní vyhodnocení:

1. Zhao, C and Wang, CA. *A cross‐site comparison of online review manipulation using Benford’s law*. Electronic Commerce Research. DOI:10.1007/s10660-020-09455-8.

Tento článek se od tématu disertační práce odchyluje. Nepracuje s počty citací, ale transformuje výskyty slov pro Benfordův zákon a pokouší se tak o forenzní analýzu podvádění při tvorbě peer-review. Při ponechní stranou všeho ostatního, tento článek poskytuje zajímavý statistický test Kuiper’s modifies KS statistic, který je slibný co se použití v disertační práci týče. Autoři uvádí, že je méně náchylný na velikost datasetu (na rozdíl od klasického KS či Chí-kvadrátu).

1. Idrovo, AJ, Manrique-Hernández, EF and Niño, JAF. *Report From Bolsonaro’s Brazil: The Consequences of Ignoring Science*. International Journal of Health Services 51(1), pp. 31-36. DOI:10.1177/0020731420968446.

Využití Benfordova zákona při kontrole dat ohledně COVID-19. Bohužel se jedná o klasický Chí-kvadrát test použitý na vzorku. Ve studii zcela v pořádku, ale pro účely disertační práce nepoužitelné.

1. Grammatikos, T and Papanikolaou, NI. *Applying Benford’s law to detect accounting data manipulation in the banking industry.* Journal of Financial Services Research 59, pp. 115-142. DOI:10.1007/s10693-020-00334-9.

Tento článek je zajímavý jako přípracová studie z oboru bankovnictví, ale jinak nemá žádný metodologický přínos a vlastně i oborově je dost nízký.

Komentář autora: Atd. atd. Pokračoval bych hledáním v dalších zvolených databázích a přesně popsal v jakých databázích hledám, jaké jsem zadal vyhledávají řetězce a proč. Kolik výsledků jsem obdržel, jak jsem upravoval vyhledávací řetězec v dalších iteracích hledání či jak jsem měnil podmínky hledání. Nejprve je vždy logická (formální) shoda nalezených výsledků dle vyhledávacího řetězce, následně je zjišťována obsahová shoda s tím, co potřebuji. Komentovaná rešerše takovou rešerši obohacuje o pročtení článku a jeho krátké shrnutí (o čem je/k čemu mi je) v souladu s rešeršním záměrem. Citace zde uvedené nejsou dle ISO 690. Norma je to podobná, ale není stejná. Vše zde uvedené bylo původně určeno jen pro osobní potřebu a už se mi nechtělo všechny výsledky převádět do správné normy. Neinspirujte se prosím formou citování.

Cerqueti, R and Maggi, M. *Data validity and statistical conformity with Benford’s Law*. Chaos, Solitons & Fractals 144, p. 110740 . DOI:10.1016/j.chaos.2021.110740.

Barabesi, L, Cerasa, A, Cerioli, A and Perrotta, D. *On characterizations and tests of Benford’s law*. To appear in: Journal of the American Statistical Association. DOI:10.1080/01621459.2021.1891927.

Azevedo, CdS, Gonçalves, RF, Gava, VL and Spinola, MdM. *A Benford’s Law based methodology for fraud detection in social welfare programs: Bolsa Familia analysis*. Physica A 567, p. 125626. DOI:10.1016/j.physa.2020.125626.

Wei, A and Vellwock, AE. Is COVID-19 data reliable? A statistical analysis with Benford's Law. Preprint, posted September. DOI:10.13140/RG.2.2.31321.75365/1.

Vellwock, AE and Wei, A. On the Benfordness of academic citations. Preprint posted on research gate.net; last accessed March 16, 2021. DOI:10.13140/RG.2.2.22108.82562.

Silva, L and Filho, DF. Using Benford’s law to assess the quality of COVID-19 register data in Brazil. Journal of Public Health. DOI:10.1093/pubmed/fdaa193.

Sambridge, M and Jackson, A. National COVID numbers — Benford’s law looks for errors. Nature 581(7809), p. 384. DOI:10.1038/d41586-020-01565-5.

Said, T and Mohammed, K. Detection of anomaly in socio-economic databases, by Benford probability law. 2020 IEEE 6th International Conference on Optimization and Applications (ICOA), Beni Mellal, Morocco, 2020, pp. 1-4. DOI:10.1109/ICOA49421.2020.9094466.

Rosetti, CJ. Using Benford's Law to Detect Fraud. PDF Workbook Self-Study (ACFE).

Novosel, D and Alanovic, M. Analysis of Consistency of Prime-boost Covid-19 Baseline and Safety Data. Acta Scientific Medical Sciences 4(12), pp. 81-82. ISSN/ISBN:2582-0931.

Moreno-Montoya, J. Benford ́s Law with small sample sizes: A new exact test useful in health sciences during epidemics. Revista de la Universidad Industrial de Santander. Salud UIS vol. 52(2), pp. 161-163.

Miranda, AT. The Distribution of COVID-19 Cases in the Philippines and the Benford’s Law. Philippine e-Journal for Applied Research and Development 10, pp. 29-34.

Miranda, AT. World’s Distribution of Covid-19 Cases and The Benford’s Law. Cape Comorin II(V), Special Issue, pp. 16-20 .

Horton, J, Kumar, DK and Wood, A. Detecting academic fraud using Benford law: The case of Professor James Hunton. Research Policy 49(8), 104084 . DOI:10.1016/j.respol.2020.104084.

Gonsalves, RA. Benford’s Law — A Simple Explanation. Posted on towardsdatascience.com blog October 1; last accessed October 10, 2020.

Goh, C. Applying visual analytics to fraud detection using Benford's law. Journal of Corporate Accounting & Finance, pp. 1-7. DOI:10.1002/jcaf.22440.

Ducharme, RG, Kaci, S and Vovor-Dassu ,C. Smooths Tests of Goodness-of-fit for the Newcomb-Benford distribution. Preprint: arXiv:2003.00520v1 [math.ST]. FRE

Berger, A and Hill, TP. The mathematics of Benford’s law: a primer. Statistical Methods & Applications. DOI:10.1007/s10260-020-00532-8.

Li, F, Han, S, Zhang, H, Ding, J, Zhang, J and Wu, J. Application of Benford’s law in Data Analysis. Journal of Physics: Conference Series 1168, pp. 032133. DOI:10.1088/1742-6596/1168/3/032133.

Kuruppu, N. The Application of Benford’s Law in Fraud Detection: A Systematic Methodology. International Business Research, Canadian Center of Science and Education 12(10), pp. 1-10. DOI:10.5539/ibr.v12n10p1.

Kazemitabar, Jal and Kazemitabar, Jav. Measuring the conformity of distributions to Benford's law. Communications in Statistics - Theory and Methods, pp. 1-7. DOI:10.1080/03610926.2019.1590599.

Baryła, M and Pociecha, J. Euclidean distance as a measure of conformity to Benford's law in digital analysis for fraud detection. Book of Short Papers, Proceedings of the 12th Scientific Meeting of the Classification and Data Analysis Group (CLADAG) of the Italian Statistical Society (SIS), pp. 75-78.

Sarkar, T. What is Benford’s Law and why is it important for data science?. Posted on: towardsdatascience.com, last accessed September 3, 2019.

Klimov, Y and Klimov, O. Benford's Law in Information Science. Intellectual Archive 7. DOI:10.32370/2018\_09\_3. RUS

Tseng, H-C, Huang, W-N and Huang, D-W. Modified Benford’s law for two-exponent distributions. Scientometrics 110(3), pp. 1403–1413. DOI:DOI 10.1007/s11192-016-2217-6.

Fairthorne, RA. Progress in Documentation - Empirical Hyperbolic Distributions (Bradford-Zipf-Mandelbrot) for Bibliometric Description and Prediction. Journal of Documentation 25(4), pp. 319-343; reprinted 2005 in Journal of Documentation 61(2), pp. 171-193. ISSN/ISBN:0022-0418. DOI:10.1108/00220410510585179.

Egghe, L and Guns, R. Applications of the generalized law of Benford to informetric data. Journal of the American Society for Information Science and Technology 63(8), pp. 1662-1665. ISSN/ISBN:1532-2882. DOI:10.1002/asi.22690.