Väčšinu vedomého života vedieme konfliktný zápas so svojimi myšlienkami. *Ak určíme nový spôsob myslenia, miznú staré otázky; tieto otázky spočívajú v spôsobe vyjadrovania; ak sa človek oblieka do nových šiat, odkladá so starými aj staré otázky*. Wittgenstein uznal: *dobré podobenstvo osviežuje rozum*.

Dušan Jedinák

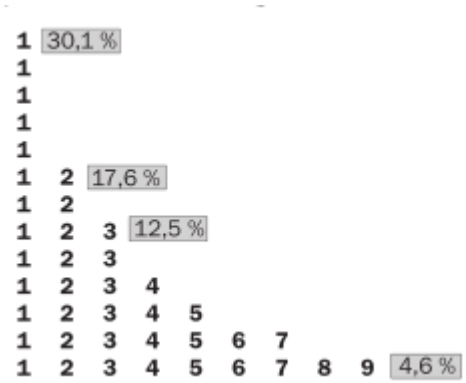
Lož má krátke nohy

Reč bude o prvej nenulovej číslici v číslach, s ktorými sa stretávame v našom každodennom živote. Povedzme v cenách tovarov v nákupnom centre. Aká je pravdepodobnosť, že to bude jednotka?

O tejto otázke by sme mohli uvažovať takto: Keďže prichádzajú do úvahy číslice 1, 2, ..., 9, tak pravdepodobnosť výskytu akejkoľvek číslice je 1/9, teda približne 11,1 %. To znamená, že z 1000 cien by zhruba 111 malo začínať číslicou 1. Táto úvaha však nie je správna. Pre väčšinu veľkých skupín čísiel reprezentujúcich čokoľvek skutočné z tohto sveta platí tzv. Benfordov zákon. Tento zákon hovorí, že pravdepodobnosť výskytu číslice *d* na prvom mieste je

$$F(d) = \log(d+1) - \log d.$$

To znamená, že vo veľkom počte čísiel je približne 30,1 % čísiel s prvou nenulovou číslicou 1, približne 12,5 % čísiel s prvou nenulovou číslicou 3 a len približne 4,6 % čísiel s prvou nenulovou číslicou 9. Teda z 1000 cien má zhruba 301 začínať jednotkou, 125 trojkou a len 46 deviatkou.



Pozreli sme sa na internetové ceny tovarov ponúkaných jednou konkrétnou spoločnosťou. Začiatkom januára 2019 sme si zaznamenali prvé číslice z 1000 cien, a to zo všetkých tovarov kategórie kreatívne hračky a stavebnice. Aby sme dosiahli spolu 1000 položiek, tak sme zobrali 182 tovarov kategórie vozidlá. Vyšlo nám, že 274 cien začína jednotkou, 131 trojkou a 128 deviatkou. Tento výsledok hovorí, že Benfordov zákon pre tento súbor dát neplatí, lebo deviatok je príliš veľa.



Mnoho cien sa však dá vyjadriť ako $x + 0{,}99$, kde x je prirodzené číslo. Urobili sme preto nový súbor dát z pôvodného tak, že sme ku každej cene pripočítali 1 cent. Vyšlo nám, že v takto upravenom súbore začína 345 cien jednotkami, 137 trojkami a 50 deviatkami. To je k Benfordovmu zákonu oveľa bližšie.

Benfordov zákon platí pre dáta z prírody aj pre spoločenské údaje. Súbory čísiel môžu predstavovať finančné čiastky z účtov, čísla domov, veľkosti populácií, dĺžky riek, hodnoty fyzikálnych veličín, veľkosti súborov uložených v počítači atď. Prvý raz o tomto zákone písal kanadský astronóm a matematik Simon Newcomb v roku 1881. V roku 1935 ho znovu objavil americký fyzik Frank Benford. Obaja objavitelia si pri čítaní logaritmických tabuliek v knižnici všimli, že prvé strany sú oveľa špinavšie ako ostatné. To znamenalo, že ľudia z rôznych príčin hľadali častejšie hodnoty logaritmov pre čísla začínajúce jednotkou. Preto by sa tento štatistický zákon mohol pokojne volať aj zákonom špinavých strán. Benford zistil, že zákon platí pre rôzne skupiny dát – počty obyvateľov, hladiny úmrtnosti, chemické a fyzikálne konštanty, bejzbalové štatistiky, polčasy rozpadu rádioaktívnych izotopov, prvočísla, prvky Fibonacciho postupnosti a mnohé iné.

Súbory dát, pre ktoré tento zákon neplatí, sú napríklad čísla v lotérii, telefónne čísla, ceny benzínu, dátumy, výšky a hmotnosti skupiny ľudí. Niektoré prípady, keď zákon platí, môžu byť vysvetlené. Napríklad adresy domov takmer vždy začínajú 1 a nižšie čísla sa musia vyskytnúť pred vyššími. Vysvetlenie väčšej časti prípadov sa však ukázalo byť tvrdým orieškom. Vedci napriek tomu, že majú veľa výsledkov o zákone, zatiaľ nenašli univerzálne platné pravidlo, podľa ktorého by mohli o skupine čísiel vyhlásiť, či pre ňu zákon platí alebo neplatí. Pri hľadaní takéhoto pravidla sa naráža na to, že treba spájať matematické dôvody pre zákon s dostatočne dobrým formálnym popisom údajov.

Pre praktickú skúsenosť s Benfordovým zákonom vám navrhujem urobiť si nasledujúci pokus s názvom **Novinová štatistika.** Tento pokus by vám mal potvrdiť veľmi prekvapujúci fakt, že rôznorodé čísla, ktoré nie je možné nijakým spôsobom dopredu predpovedať, majú súvislosť, ktorou je to, že podliehajú Benfordovmu zákonu.

KRÍŽOVKA

Minerál, ktorého meno sa ukrýva v tajničke, patrí podľa chemického zloženia do skupiny kremičitanov. Má chemický vzorec $(Mg, Al)_2((OH)Si_4O_{10}) \cdot 4\,H_2O$. Jeho rozličné odrody majú rozmanité názvy – ľudovo sa nazývajú „skalná koža“, „skalný papier“, „banská koža“ alebo „skalný korok“. Tvorí hrubé plstnaté kožovité útvary alebo spletené vlákna bielej, žltej, sivožltej až hnedastej farby, ktoré vypĺňajú pukliny magnezitu a dolomitu, alebo sa vyskytuje vo forme povlakov. Vzniká ako produkt zvetrávania bázických hornín bohatých na horčík v puklinách i v usadených horninách. Jeho vryp je biely, lesk matný. Dosahuje tvrdosť 1 – 2 a hustotu 2,1 – 2,3 g/cm^3. Je jemný, ohybný. Na Slovensku sa vyskytuje v Divíne, v Ratkovskej Suchej, na Burde, v Lubeníku, v Dúbrave a pri Jelšave.

Názov tohto minerálu je ukrytý v nasledujúcej krížovke. Do jednotlivých riadkov doplňte názvy nerastov, ktorých opisy a fotografie sa nachádzajú na nasledujúcich stránkach. Tajnička sa nachádza vo vyznačenom stĺpci.

2. Nerast má chemický vzorec Fe_3O_4. Patrí medzi oxidy. Jeho zvláštnou vlastnosťou je prirodzená magnetickosť, ktorá súvisí i s jeho názvom. Obsahuje 60 – 70 % železa, preto je významnou železnou rudou. Na Slovensku sa vyskytuje v malom množstve pri Tisovci a vo Vyhniach. Bohaté ložiská sú v Rusku (Magnitogorsk, Kursk), Švédsku (Kiruna) i Austrálii.



Zoberte si 2 – 3 kusy dennej tlače a pozrite sa na časti, ktoré obsahujú čísla. Hľadajte miery rôzneho druhu (výšky, vzdialenosti, hmotnosti atď.), údaje o peniazoch (straty, zisky, ceny) a údaje o počtoch (autá na parkovisku, obrazy na výstave, ľudia na festivale, kusy v jednom balení, strelené góly atď.). Vynechajte roky, dátumy, hodiny, percentá, vek ľudí, telefónne čísla. Nájdené čísla si zaznamenajte. Vyhľadávanie skončite, keď budete mať 600 a viac čísiel (čím viac, tým lepšie). Potom spočítajte počty jednotlivých prvých číslic, vyjadrite ich v percentách z celkového počtu čísiel, ktoré ste zozbierali, a nakoniec výsledok porovnajte s Benfordovym zákonom. To môžete urobiť sami alebo len zadať všetky vaše čísla programu nainštalovanému na stránke

https://di.ics.upjs.sk/vyucba/pomocne_materialy/bz.php.

Za nadpis článku sme zvolili známe slovenské príslovie. Možno sa pýtate, ako súvisí s našou témou. Nuž, keď si ľudia vymýšľajú falošné čísla, napr. pri voľbách alebo rôznych ekonomických ukazovateľoch, tak majú tendenciu používať číslice na začiatku čísiel tak, že to odporuje Benfordovmu zákonu. Napríklad sa zistilo, že pri falšovaní účtov sa najčastejšie ako prvé vyskytujú číslice 5 a 6. Teda Benfordov zákon sa môže používať (a aj sa používa) ako jednoduchý test pravdivosti údajov. Je stále bežnejšou súčasťou vyšetrovania v prípadoch, kde sa narába s väčším množstvom dát. Vzhľadom na to, čo sme uviedli vyššie, nám kladný ani záporný výsledok nedáva stopercentnú istotu pravdivosti či nepravdivosti uvádzaných údajov. Avšak ukazuje sa ako vynikajúci indikátor. Tento test pomohol odhaliť napríklad volebné podvody na Floride a vo Venezuele v roku 2004 a v Mexiku v roku 2006.

Emília Halušková