rés kiválasztásakor különös figyelmet fordítottunk arra, hogy kihagyjuk azon méréseket, amelyeken a kiértékelés után ilyen mozgási műtermékek adódtak. Elméletben ez a jelenség elkerülhető, ha a polarimetriához szükséges három felvételt azonos időben készítjük el. A valóságban egy ilyen megközelítés a következő nem-triviális problémákat veti fel: (i) A legfontosabb a három kamera optikai tengelye tökéletes párhuzamosságának biztosítása. (ii) Egy másik nehézség, hogy a három kamera érzékelője kissé eltérő választ ad azonos fényintenzitásokra még akkor is, ha azonos típusúak, ami keresztkalibrációt igényel. Felhasználhatók lettek volna a hajó belső szenzorainak adatai, amelyek megadják a hajó pontos állásszögét, hogy ezzel korrigáljuk az egyes képek irányítottságát. Ugyanakkor ez nagyon pontos szinkronizációt igényelt volna a hajó belső rendszerével. Ehelyett egyszerűen kihagytuk azon méréseket, amelyeken a kiértékeléskor hiba keletkezett a hajó elmozdulása miatt.

Következtetések

Azt találtuk, hogy a neurális hálózaton alapuló felhődetekciós algoritmusok voltak a legjobbak a 13 vizsgált közül, és hogy a globális paraméterek (például a fényintenzitás átlaga és varianciája), a nem-optikai információk (például a relatív nappozíció) és a polarimetrikus információk (különösen a polarizációfok) növelhetik a felhőfelismerés pontosságát. Ugyanakkor, mint azt az NNN és PNN algoritmusok hibái közti kis különbség mutatja, a polarizáció csak kismértékben növeli a pontosságot. Fontos azonban hangsúlyozni, hogy a polarizációs információkból más mennyiségekre is következtethetünk, mint például a felhőalap magasságára vagy az aeroszol koncentrációjára, amely mennyiségek mérése fotometrikus úton nem lehetséges. A polarizáció segíthet a felhőtípusok osztályozásában is, például elkülöníteni a jégfelhőket a vízfelhőktől, vagy akár a jégfelhők mikrofizikai jellemzőinek (részecskeméret és -alak) meghatározásában.

AZ ELSŐ SZÁMJEGYEK BENFORD-TÖRVÉNYE ÉS A RADIOAKTÍV IZOTÓPOK FELEZÉSI IDEJE

Gyürky György, Farkas János
MTA Atommagkutató Intézet, Debrecen

Mindennapi életünkben körülvesznek minket a számok és e számoknak olykor érdekes és meglepő tulajdonságaik vannak. Az egyik ilyen meglepő tulajdonságot Benford-törvénynek nevezzük. Bármilyen furcsa és szinte hihetetlen is a törvény által leírt jelenség, annak igazságáról bárki könnyedén meggyőződhet. Mégis kevéssé ismert még a tudomány művelői körében is. Pedig a törvénynek már különböző gyakorlati felhasználásai is vannak. E cikk apropóját egy olyan alkalmazás adja, amely – mint megmutatjuk – teljesen hibásan használja a törvényt.

Egy kis történelem

A 19. század végén *Simon Newcomb,* egy kanadai–amerikai matematikus és csillagász érdekes felfedezést tett logaritmustáblázatok tanulmányozása közben. A *Fizikai Szemle* olvasóinak jó része talán már nem is tudja, mi az, hogy logaritmustáblázat. Az elektronikus számológépek elterjedése előtt ilyen táblázatok könnyítették meg különböző matematikai műveletek elvégzését, tehát például a tudósok gyakran forgatták a könyvekbe rendezett táblázatokat. Newcomb azt találta, hogy a könyvek első oldalai sokkal inkább kopottak voltak, mint a hátrébb található oldalak. Mivel a könyvekben a listák az egyes számtól kezdődnek és a kilences számjeggyel kezdődő számokkal végződnek, így a könyvek kopása alapján úgy tűnt, a felhasználók gyakrabban keresnek alacsony számjeggyel kezdődő számokat, mint magas kezdőszámjegyűeket.

Newcomb meglepő felfedezése hosszú időre feledésbe merült, mígnem 1938-ban a General Electric fizikusa, *Frank Benford* újra rábukkant a jelenségre. Az ebből a felfedezésből születő „törvény” ezért az ő nevét viseli: az első számjegyek Benford-törvénye, vagy egyszerűen csak Benford-törvény [1]. A táblázatok kopását úgy értelmezte, hogy a természetben vagy a mindennapi életünkben gyakrabban fordulnak elő kis számjeggyel kezdődő számok.

Ezt a meglepő lehetőséget Benford igen alapos ellenőrzésnek vetette alá. A tudomány és a mindennapok számos különböző területéről vett adatokat, számsorokat, és az első számjegyek eloszlását vizsgálta bennük. Adatbázisában megtalálhatók voltak fizikai és matematikai állandók, molekulatömegek, földrajzi adatok, mint például folyók hossza és tavak területe, amerikai települések lélekszáma, híres emberek utcaházszáma, vagy akár halálozási statisztikák is. Azt találta, hogy az adatsorok nagy többségére valóban igaz, hogy az első számjegyek eloszlása nem egyenletes, hanem a kis számok irányába torzult. Például az egyessel kezdődő számok több mint hatszor gyakoribbak, mint a kilencessel kezdődőek.

Aki hitetlenkedve fogadja ezeket az eredményeket, annak javaslunk egy próbát. Listázza ki az összes számot a *Fizikai Szemle* azon példányából, amit éppen a kezében tart és vizsgálja meg az első számjegyek eloszlását! Álljuk a fogadást, hogy az adatok hibahatáron belül követni fogják a Benford-törvény alább ismertetett matematikai alakját.

A Benford-törvény matematikai jellegzetességei és alkalmazási lehetőségei

A Benford által vizsgált adatok eloszlása jól leírható egy egyszerű képlettel. Annak a valószínűsége, hogy egy szám első számjegye *d* a következőképpen adható meg:

$$P(d) = \log_{10}\left(1 + \frac{1}{d}\right), \quad d = 1,\ 2,\ \dots\ 9. \qquad (1)$$

Ezen valószínűségeket mutatja az *1. ábra*. Például annak a valószínűsége, hogy egy szám egyessel kezdődik, 30,1% (majdnem minden harmadik szám egyessel kezdődik!), míg annak, hogy kilencessel, 4,6%. Vegyük észre, hogy a valószínűségek megegyeznek annak esélyével, hogy egy logaritmikus skálán ábrázolt számegyenesen az 1 és 10 között véletlenszerűen kiválasztott pontnak mi az első számjegye. Ennek illusztrálása szintén szerepel az *1. ábrán*.

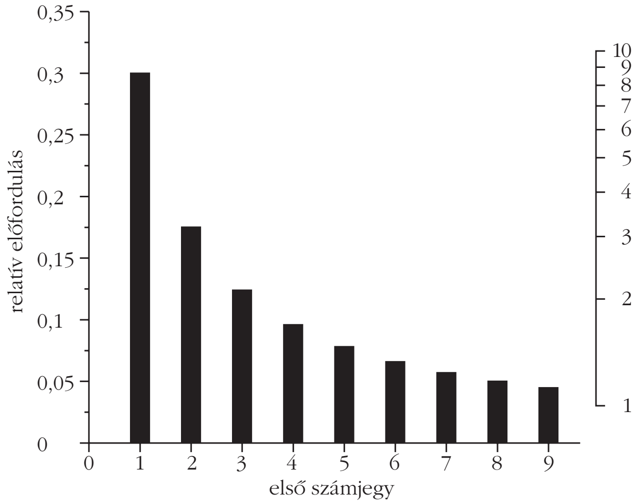
Cikkünk további mondanivalója szempontjából lényegtelen, de a teljesség kedvéért megemlítjük, hogy a nem egyenletes eloszlás nemcsak a számok első, hanem későbbi számjegyeire is érvényes. Ám ez már messze nem olyan szembeszökő, mint az első számjegyek esetén. A második számjegyben például a nulla gyakorisága 12%, a kilencesé pedig 8,5%. Ez már csak elegendően nagy adathalmaz esetén mutatható ki.

Természetesen a tízes számrendszer kiválasztásának nincs kitüntetett szerepe. Ha a Benford-törvénynek eleget tevő adatsort átszámítjuk egy tetszőleges (nem túl nagy) alapú számrendszerbe, akkor az így nyert új adatsor szintén teljesíti a törvényt az adott számrendszerben kifejezve. Az (1) képlettel adott valószínűség általános alakja *n* alapú számrendszerben adott számok esetén tehát:

$$P(d) = \log_{n}\left(1 + \frac{1}{d}\right), \quad d = 1,\ 2,\ \dots\ n-1. \qquad (2)$$

A Benford-törvény első látásra meglepő tulajdonsága a skálafüggetlenség. Nem számított, hogy Frank Benford például a vizsgált folyók hosszát kilométerben, mérföldben vagy akár lábban adta meg, az adatok minden esetben követték a törvényt. Ha tehát egy adatsorunk követi a törvényt, akkor ha minden adatot megszorzunk egy tetszőleges nemnulla állandóval, akkor az így előállt adatsor is követni fogja a törvényt. Ez a jellegzetesség jó lehetőséget teremt annak ellenőrzésére, hogy egy adatsor ténylegesen követi-e a törvényt, mint ezt a későbbiekben láthatjuk.

A Benford-törvény gyakorlati felhasználása is létezik, méghozzá meglepő területeken. Az alkalmazás



1. ábra Az első számjegyek eloszlása a Benford-törvény szerint.

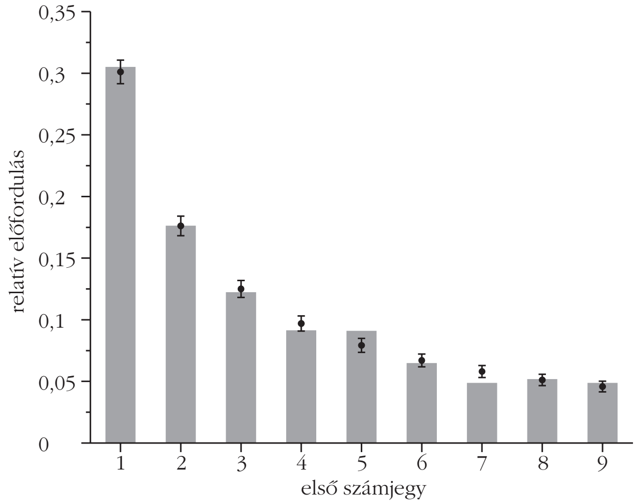
főképpen azon alapul, hogy az emberek nem ismerik a törvényt. Ha egy adatsornak valamilyen okból követnie kell a törvényt, de ezt mégsem teszi, az gyanús. Tehát könnyen leleplezhetővé válik az adatsort létrehozó olyan ember, aki nem ismeri a törvényt. Ezért leggyakrabban csalások leleplezésére használják a Benford-törvényt. Az első számjegyek eloszlását vizsgálva derítettek már fel adócsalásokat, választási visszaéléseket, lepleztek le hamisított adatsort használó egyetemi szakdolgozót. Sőt, a Benford-törvény alapján végzett ellenőrzés még azt is megmutatta, hogy Görögország „kozmetikázta” államháztartási adatait [2].

Csalások leleplezésén kívül is van példa a törvény használatára. Megvizsgálták például földrengés-érzékelő szeizmométerek adatait és más eloszlást találtak az első számjegyekben egy adott földrengést megelőző adatokban és a rengés során [3]. Szintén próbálták alkalmazni radioaktív izotópok felezési idejét szolgáltató elméletek tesztelésére. Mint megmutatjuk, hibásan. Erről szól cikkünk hátralévő része.

Radioaktív izotópok bomlásának felezési ideje

A 21. század elejére – főként a radioaktív ionnyalábokkal végzett kutatásoknak köszönhetően – több mint 3500 különböző (eltérő proton- vagy neutronszámú) izotópot ismerünk. Ezen izotópok döntő többsége radioaktív és általában ismerjük bomlásuk felezési idejét is. Mit mutat az ismert izotópok felezési időiből képzett adatsor, ha megvizsgáljuk a Benford-törvény szempontjából?

2008-ban *Dongdong Ni* és *Zhongzhou Ren* egy rangos folyóiratban publikálták cikküket, amelyben megvizsgáltak mintegy 3000 radioaktív izotópot és azt találták, hogy felezési idők jól követik a Benford-törvényt [4]. A tapasztalaton felbuzdulva a szerzők a háttérben rejlő fizikai okokat boncolgatták, és odáig jutottak, hogy a Benford-törvénnyel való összevetés alkalmas lehet olyan elméleti modellek ellenőrzésére, amelyekből bomlások felezési ideje



2. ábra. A NUBASE2003 adatbázisban található felezési idők első számjegyeinek eloszlása.

1. táblázat

A felezési idők adatbázisában az adott első számjegyek előfordulásának száma, illetve a Benford-törvény által jósolt értékek

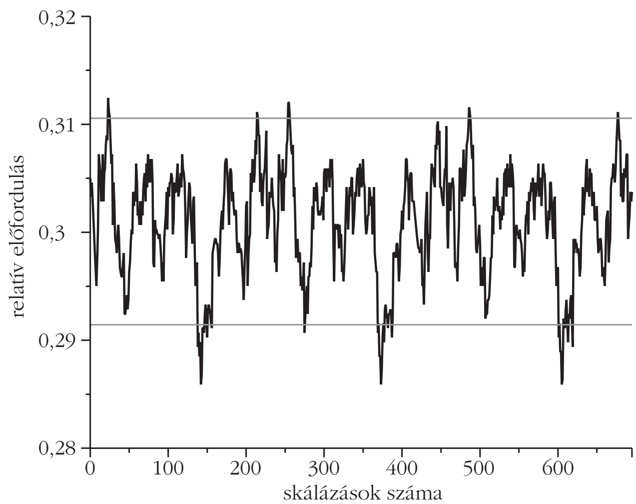
első számjegy	előfordulások tapasztalt száma az adatbázisban	Benford-törvény alapján várt előfordulásszám
1	701	692 ± 22
2	405	405 ± 18
3	281	287 ± 16
4	210	223 ± 14
5	209	182 ± 13
6	149	154 ± 12
7	112	133 ± 11
8	119	118 ± 11
9	112	105 ± 10

származtatható. Mint alább megmutatjuk, ez az elképzelés a törvény teljesen hibás értelmezéséből fakad és a leírt formában nem alkalmas elméleti modellek tesztelésére.

Előtte azonban végezzük el saját vizsgálatunkat a felezési idők adatbázisán! Ehhez a vizsgálathoz az adatokat a NUBASE2003 adatbázisból vettük [5]. Nem vettük figyelembe azokat az izotópokat, ahol nem konkrét érték, hanem csak alsó vagy felső határ volt megadva. Így összesen 2298 adatból indultunk ki. Az analízis eredménye látható a *2. ábrán,* illetve a számszerű értékeket az *1. táblázat* tartalmazza. Mint láthatjuk, a felezési időkben az első számjegyek eloszlása valóban jól követi a Benford-törvény által diktált elvárásokat az adott mintanagyságra jellemző hibahatáron belül. A Benford-törvény várható hibáját a binomiális eloszlás standard bizonytalansága alapján adtuk meg.

A Benford-törvény skálainvarianciáját kihasználva adatsorunkat még alaposabb tesztnek vethetjük alá, hogy a törvénynek való engedelmességet még erősebben bizonyíthassuk. Az adatbázisban található minden egyes értéket egy tetszőleges állandóval szorozva megvizsgálhatjuk, hogy például az adatok hány százaléka kezdődik egyes számjeggyel. Ennek a százaléknak mindig az elvárt 30,1% közelében kell lennie, akárhányszor végzünk is új skálázást. Tesztünkhöz az 1,01 értéket választottuk skálázási tényezőként. Ezzel a számmal szoroztuk meg a teljes adatbázist egymás után több száz alkalommal és minden lépésben vizsgáltuk az egyes, mint első számjegy gyakoriságát. E skálázási teszt eredménye található a *3. ábrán*. Jól látható, a gyakoriság a Benford-törvény várhatóértéke, 30,1% körül ingadozik, nagyjából a hibatartományon belül. Tehát a skálázási teszt is egyértelműen bizonyítja, hogy a felezési idők valóban követik a Benford-törvényt.

3. ábra. Skálázási teszt: az egyesek első számjegyként való relatív előfordulása a felezési idők többszöri, 1,01-dal való megszorzása után.



Ha a felezési idők ilyen kiválóan követik a törvényt, miért állítjuk azt, hogy mégsem alkalmas a törvény elméleti modellek tesztelésére? Az indokoláshoz először lássuk, hogy mi van a törvény mögött! Ez azért fontos, mert bár nem túlságosan bonyolult dologról van szó, mégis érdekes és olykor hajmeresztő magyarázatok láttak napvilágot arra, hogy a törvény miért ír le jól oly sok, mindennapi életünkben felbukkanó adatsort. Például még maga Benford is olyan magyarázattal próbált szolgálni, hogy míg mi emberek úgy számolunk, hogy 1, 2, 3, …, addig a természet úgy számol, hogy e^1, e^2, e^3, …. Ilyen misztikus magyarázatokra természetesen nincs szükség. Olyannyira nincs, hogy megmutatható, ha egy adatsorban az első számjegyek eloszlása skálainvariáns, akkor az adatoknak szükségképpen követniük kell a Benford-törvényt.

Mikor teljesül egy adatsorra a Benford-törvény?

Az a feltétel, hogy egy szám egy adott számjeggyel kezdődik, könnyebben kifejezhető a szám logaritmusának használatával. Valóban, egy x szám akkor és csak akkor kezdődik például egyes számjeggyel, ha logaritmusára igaz, hogy

$$n \leq \log(x) < n + 0{,}301,$$

ahol n egy tetszőleges egész szám. Hasonló összefüggés írható fel bármely más kezdőszámjegy esetén is. Mit jelent a skálainvariancia a logaritmus használata esetén? Az adatok konstanssal való szorzása a logaritmusnál egy konstans hozzáadását jelenti:

$$\log(a \cdot x) = \log(x) + d',$$

ahol $d' = \log(a)$. Egy a konstanssal való szorzás után tehát annak valószínűsége, hogy az első számjegy egyes újra a következő:

$$n \leq \log(a \cdot x) < n + 0{,}301,$$

tehát

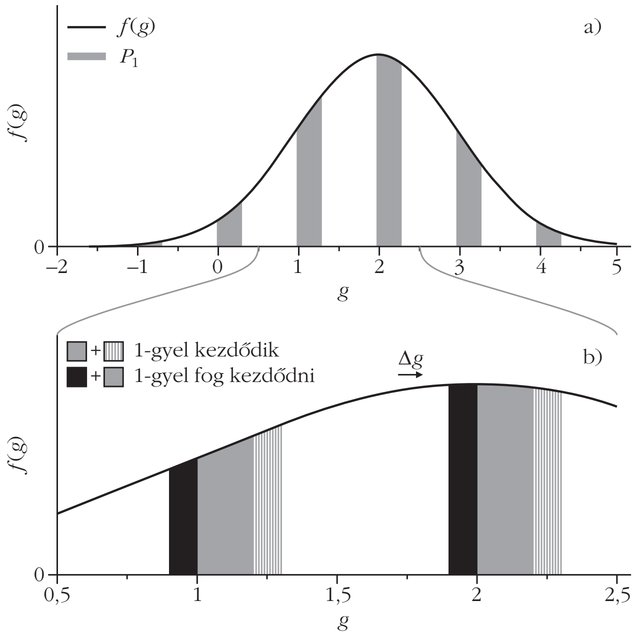
$$n - \log(a) \leq \log(x) < n - \log(a) + 0{,}301.$$

Ez mindössze egy $\log(a)$ konstanssal való eltolást jelent. A skálainvariancia megköveteli, hogy a két valószínűség megegyezzen. Ez csak akkor teljesülhet, ha a számsor számainak logaritmusa (pontosabban annak tizedesvessző utáni része) egyenletes eloszlást követ a [0,1) intervallumon!

A fenti okfejtést igyekszik szemléltetni a *4. ábra,* ahol egy vizsgálandó adatsorban lévő számok eloszlásfüggvénye látható logaritmikus skálán. A szürkével fedett terület mutatja azokat a számokat, amelyek első számjegye egyes. A fekete terület pedig azt mutatja, melyek azok a számok, amelyek a szürke területre kerülnek egy bizonyos skálázás során, tehát az egész adatsor adott konstanssal való megszorzása után. A szürke csíkos területre eső értékek a skálázás előtt egyessel kezdődnek, de a skálázás után már nem. Az adatsorunkban akkor teljesül a skálainvariancia, azaz akkor nem változik az adott számjeggyel való kezdődés valószínűsége, ha a fekete és a szürke csíkos terület egymással egyenlő, tehát a skálázás során annyit veszítünk az egyik oldalon, amennyit nyerünk a másikon.

Ezek alapján kvalitatív kritériumot adhatunk arra nézve, hogy milyen eloszlást kell követnie egy adathalmaznak, hogy a Benford-törvény igaz legyen rá. Triviálisan igaz lenne a törvény, ha az adatok például szigorúan 1 és 10 közé esnének és a logaritmusuk ezen az intervallumon egyenletes eloszlású lenne. Ilyen eloszlás a természetben vagy a mindennapi életben azonban ritkán fordul elő. Sokkal fontosabb a másik eset (*4. ábra*) amikor az eloszlás olyan, hogy sok nagyságrendet átívelően nagyjából egyenletes a logaritmikus skálán. Ekkor a szürke területek eltolása összességében nem változtatja meg az összterületüket, mert az esetleges kis eltérések kiátlagolódnak. Kimondható tehát, hogy amennyiben egy adatsorban sok nagyságrendet átfogóan találhatók értékek, mégpedig – logaritmikusan – nagyjából egyenletes eloszlásban, akkor az adatsor jól fogja követni a Benford-törvényt.

A fent elmondottak persze nem tekinthetők matematikai igényességű bizonyításnak, az túlmutatna

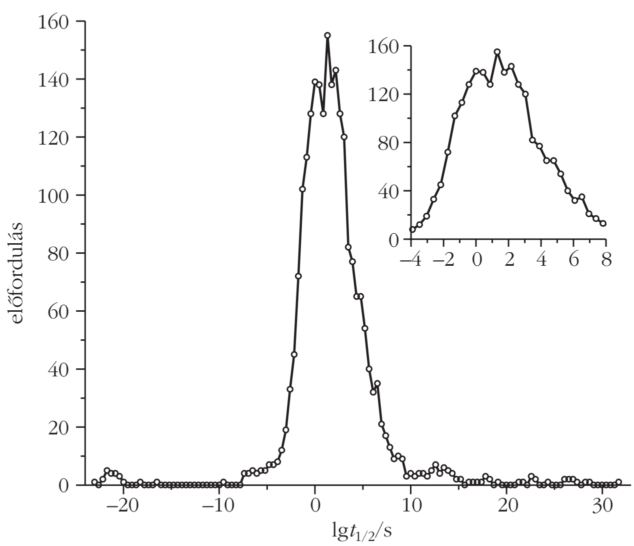


4. ábra. A skálázás hatása az 1-ek eloszlására. A szürke csíkos részre eső számok a skálázás után már nem fognak 1-gyel kezdődni, míg a fekete részek a skálázás előtt nem kezdődtek 1-gyel, de utána már igen. A szürke részre eső számok a skálázás előtt és után is 1-gyel kezdődnek.

cikkünk keretein. Bizonyítás nélkül közöljük csak az egzakt kritériumot, az érdeklődők megtalálhatják a részleteket a [6] vagy [7] dokumentumban. Tehát egy adatsor akkor és csak akkor teljesíti a Benford-törvényt, ha az adatok eloszlásfüggvényének Fourier-transzformáltja eltűnik minden egész értéknél (frekvenciánál). Gyakorlatban elegendő a törvény jó közelítéssel való teljesüléséhez, ha az említett Fourier-transzformáltak elég kis értéket vesznek fel egész frekvenciáknál.

Mit mondhatunk el a felezési időkről? Az *5. ábra* a felezési idők eloszlását mutatja logaritmikus skálán. A teljes adatbázis mintegy 50 nagyságrendet ölel fel, tehát igen széles eloszlásról van szó. Ugyan messze

5. ábra. Felezési idők előfordulása.



nem logaritmikusan egyenletes, de a középső kiemelkedő része (lásd a kis ábrán) még mindig sok nagyságrendet fog át meglehetősen sima eloszlással. Így nem meglepő, hogy az adatsorra teljesül a Benford-törvény.

Konklúzió

Miért állítjuk tehát, hogy a Benford-törvény nem alkalmas felezési időt számító elméleti modellek tesztelésére? A válasz az, hogy a Benford-törvénnyel való szembesítés csak az adathalmaz eloszlásfüggvényének alakját vizsgálja. Ha az adatok sok nagyságrendet átívelően megfelelően egyenletes eloszlásúak, akkor a Benford-törvény teljesülni fog. Tehát, ha egy elméleti modell ilyen felezési időket szolgáltat, akkor ki fogja állni a törvény próbáját. De ez a próba önmagában nem mond semmit arról, hogy a modell fizikailag mennyire helyes. A tapasztalattal teljesen összeegyeztethetetlen eredményt adó modellek is teljesíthetik a törvényt, mégsem fogadjuk el őket helyesnek. Ez fordítva is igaz: ha egy modell nagyságrendileg helyesen írja le atommagok széles körének felezési idejét a mikroszekundumtól a milliárd évig, akkor ez egy kiváló modell lehet. Ám esetleg a modell megalkotói az elméletük korlátait felismerve minden felezési időt csak nagyságrendi pontossággal, $1 \cdot 10^n$ s alakban adnak meg, akkor az adathalmaz triviálisan nem fogja teljesíteni a Benford-törvényt, pedig fizikailag igen értékes az elmélet.

A Benford-törvény a világunkban előforduló számok egy – első pillantásra meglepő, igen érdekes – tulajdonságát írja le. A körülötte kialakult kultusznak köszönhetően az érdeklődő olvasók bőséges (főként angol nyelvű) irodalmat találhatnak a témával foglalkozó gyűjtő-weboldalon [8]. Már ma is szép számmal akadnak a törvénynek gyakorlati alkalmazásai, és várhatóan ez a jövőben még inkább így lesz. Körültekintően kell azonban bánnunk a törvény alkalmazhatóságával, nehogy olyan hibát kövessünk el, mint az írásunkban idézett szerzők a felezési idők esetén.

Irodalom, hasznos weboldalak

1. http://en.wikipedia.org/wiki/Benford's_law
2. http://index.hu/tudomany/brittudosok/2011/10/25/matematikusok_jottek_ra_a_gorog_csalasra/, http://www.cesruc.org/uploads/soft/130301/1-1303011Z221.pdf
3. http://onlinelibrary.wiley.com/doi/10.1029/2010GL044830/pdf
4. D. Ni, Z. Ren, *Eur. Phys. J. A 38* (2008) 251.
5. G. Audi et al., *Nucl. Phys. A 729* (2003) 3.
6. http://goo.gl/Pv509w
7. S. W. Smith: *The Scientist and Engineer's Guide to Digital Signal Processing.* California Technical Publishing, 1997, 2008, Ch. 34, pp. 701–722.
8. http://www.benfordonline.net

MILLIKAN ÉS AZ ELEMI TÖLTÉS MEGHATÁROZÁSÁNAK TÖRTÉNETE – 2. RÉSZ

Buzády Andrea, Szegő Dóra[1]
Pécsi Tudományegyetem TTK Fizikai Intézet

Cikkünk – a korábbi számban megjelenő – első részében a híres olajcseppes kísérlet előzményeiről volt szó. Írtunk a mérések közös elvéről; megismertük, hogy miként törekedett Millikan a kortársak ötletein elindulva a kísérlet gondos kivitelezésére; a felhőszerű közeg, illetve később a csepp mozgását befolyásoló viszkozitás és hőmérséklet minél pontosabb meghatározására.

Az elemi töltés értékének meghatározása olajcseppek porlasztásával

Millikan tanítványától, *Harvey Fletchertől* származó ötlet [5] vezetett a – korábban sehogy sem kiküszöbölhető – párolgásból fakadó problémák megoldásához. A további kísérletekben a víz, illetve alkohol helyett nagy finomságú óraolajat használtak [4].

Millikan ezen felül számos módosítást vezetett be, ezek közül néhány:

- újra meghatározta a levegő viszkozitását;
- új, minden addiginál jobb optikai megfigyelő berendezést alkalmazott;
- a kísérleti elrendezést alkalmassá tette arra, hogy tetszőleges nyomáson vizsgálhassa a cseppek sebességét;
- a légáramlásból származó hibaforrásokat még precízebben védte ki;
- kísérletileg igazolta az alábbi három előfeltevést:
  – a töltés nagysága nem befolyásolja a cseppekre ható közegellenállást,
  – az olajcseppek szilárd gömbökként viselkednek,
  – a cseppek sűrűsége megegyezik magával az olaj sűrűségével.

Az olajcsepp-porlasztós kísérleti elrendezést a *4. ábra* mutatja.

A légköri nyomás akár tizenötszörösét is elbíró *D* rézkádat egy 40 l gázolajat tartalmazó *G* kádba merítették. Az *e* csapon és az *A* porlasztón keresztül beáramló, gondosan szárított levegőt használva porlasztással állították elő a rézkádban összegyűlő olajcseppeket. Az ebben uralkodó nyomást az *m* higanyma-

Jelen tudományos közleményt a szerzők a Pécsi Tudományegyetem alapításának 650. évfordulója emlékének szentelik.

[1] Egyetemi hallgató.