

KESZTHELYI GABRIELLA

milyen színű
a valószínű?

KESZTHELYI GABRIELLA

milyen színű a valószínű?

A VÉLETLEN MATEMATIKÁJA



A könyv megjelenését a kiadói program keretében a Nemzeti Kulturális Alap, valamint a HUN-REN Rényi Alfréd Matematikai Kutatóintézet támogatta.

A szerző a könyv megírása során a HUN-REN-BME Sztochasztika Kutatócsoport támogatásában részesült.



© Keszthelyi Gabriella, Typotex, Budapest, 2025

Engedély nélkül semmilyen formában nem másolható!

keszthelyigabriella.com

Lektorálta: Ráth Balázs

Az 1., 23., 24., 30., 31., 33., 41., 43. és 55. ábrát Miklós Kelemen készítette.

ISBN 978 963 493 307 6

Kedves Olvasó!

Köszönjük, hogy kínálatunkból választott olvasnivalót!

Újabb kiadványainkról, akcióinkról a www.typotex.hu

és a facebook.com/typotexkiado oldalakon értesülhet.

Typotex Kiadó

Alapította Votisky Zsuzsa, 1989

A kiadó az 1795-ben alapított Magyar Könyvkiadók és Könyvterjesztők Egyesülésének tagja.

Felelős kiadó: Németh Kinga

Felelős szerkesztő: Juhász Emese

Szerkesztette: Mandl Orsolya

Tördelés: Szalay Éva

Borítóterv: Tillai Tamás

Nyomás, kötés: Central Dabasi Nyomda Zrt.

Felelős vezető: Szentinek Zsolt vezérigazgató

www.dabasprint.hu

TARTALOM

ELŐSZÓ	7
BEVEZETÉS	9
1. CSOMÓSODÁSOK ÉS ILLÚZIÓK	27
2. OKSÁG ÉS INDUKCIÓ	53
3. FÉLREÉRTÉSEK A FELTÉTELES VALÓSZÍNŰSÉG KÖRÜL	73
4. OKSÁG ÉS FELTÉTELES VALÓSZÍNŰSÉG	99
5. DEMOGRÁFIA ÉS ÉLETBIZTOSÍTÁSOK	131
6. FAIR JÁTÉKOK, AVAGY MIBŐL ÉL A SZERENCSEJÁTÉK ZRT.?	149
7. A NORMÁLIS MÍTOSZA	173
8. STATISZTIKA	197
9. MODERN VALÓSZÍNŰSÉGSZÁMÍTÁS	223
FÜGGELÉK	239
KÖSZÖNETNYILVÁNÍTÁS	251
JEGYZETEK	253

ELŐSZÓ

Ezt a könyvet részben a BME-n a villamosmérnök hallgatóknak tartott valószínűségszámítás-kurzus motiválta, részben pedig a saját élményeim az egyetemről: diákként és tanárként is az volt a tapasztalatom, hogy a valószínűségek világa ingoványos terep sok matematikus számára is, a laikusoknak pedig egyértelműen az. Sokat gondolkodtam, hogy mi lehet ennek a magyarázata, számos ismerőssel, kollégával és baráttal – matekossal és nem matekossal egyaránt – beszélgettem erről az évek folyamán. Arra jutottam, hogy az egyik ok a valószínűségszámítás eredete, ami a babonákhoz és a szerencsejátékokhoz kötődik. A másik ok, hogy az agyunk nem valószínűségek (matematikailag korrekt) felmérésére van huzalozva, hanem sokkal inkább a kockázatkezelésre: nagyobb valószínűségűnek ítélünk olyan eseményeket, illetve stratégiákat, amelyek jobban biztosítják a túlélésünket. De nagyban befolyásolja valószínűségi becsléseinket az is, hogy mennyi időnk van dönteni, illetve mennyire költséges (mennyi agymunkát igényel) az eseménnyel kapcsolatos adatok feldolgozása. Sokszor egyszerűen nincs elég időnk, vagy túl sok az információ, és heurisztikákhoz folyamodunk.

Így a könyvet főleg ez a két szempont strukturálja: egyrészt végighaladok a valószínűségszámítás történetén az ősidőktől napjainkig, mikor milyen szükségletek, alkalmazások segítették elő a fejlődését, másrészt pedig megpróbálom kideríteni, hogy milyen nehézségek gátolnak minket abban, hogy helyesen ítéljük meg bizonyos események bekövetkezésének valószínűségét. Ennek érdekében a könyvet sűrűn átszövik a paradoxonok (illetve olyan problémák, amelyek technikailag nem paradoxonok, de minden esetben józan ésszel nehezen megragadható,

látszólagos ellentmondások) és kognitív torzítások (gondolkodási hibák, amelyekre különösen hajlamosak vagyunk, miközben a körülötünk lévő tengernyi információt próbáljuk feldolgozni). Ezeket külön is listázom a könyv végén található *Függelék*ben. Nagyon ügyeltem arra, hogy az olvasót ne terheljem túl felsőbb matematikai fogalmakkal, és a könyv tudományos ismeretterjesztő voltaához híven érthető maradjon laikusok számára is. Ugyanakkor vannak benne könnyebb és nehezebb részek, bizonyos paradoxonok egész mély nyúluregekbe vezetnek. Amennyiben valaki nem merészkedik ezek mélyére, a könyv még akkor is konzisztens marad, de én arra biztatok mindenkit, hogy kövesse és gondolja végig a magyarázatokat.

A *Bevezetés*ben végigveszem a valószínűségek történetét az ősidőktől a reneszánszig terjedően, illetve az ebben az időben uralkodó „minden egyenlően lehetséges” hozzáállást és az ebből fakadó problémákat. A *Csomósodások és illúziók* című fejezetben azt járom körül, hogy események együttjárása, „csomósodása” lehet-e a véletlen műve. Az *Okság és indukció* című fejezetben azt vizsgálom, hogy milyen koncepciókban gondolkodhatunk valószínűségek helyett (kauzalitásban), és mi köze van ennek mégis a valószínűségekhez. Bevezetem a feltételes valószínűség fogalmát, amit tovább boncolok a *Félreértések a feltételes valószínűség körül*ben és az *Okság és feltételes valószínűség*ben. Utóbbiban külön kitérek az időre, hogy milyen szerepet játszik abban, hogy inkább ok-okozatban, mintsem valószínűségekben gondolkodunk. A *Demográfia és életbiztosítások* című fejezet főként arról szól, hogy hogyan motiválták a pestisjárványok az első népszámlálásokat, amiből az első statisztikák és az első élettartambecslések is származnak. A *Fair játékok, avagy miből él a Szerencsejáték Zrt.?* című fejezet a fogadások és különböző játékok birodalmába kalauzol el: a várható élettartam után a várható nyereség fogalmára is fény derül, illetve hogy mit hívunk racionális döntésnek a játékokban. A *normális mítosza* a leghíresebb valószínűségi eloszlás születését és mai napig tartó hatását vizsgálja. A *Statisztika* című fejezetben kifejtem annak történetét és határait, tudományban betöltött szerepét. Végül pedig a *Modern valószínűségszámítás* a terület mai állására, a matematikában és a világban elfoglalt helyére reflektál.

BEVEZETÉS

Az embereket mindig is mélyen foglalkoztatta, hogy milyen erők dolgoznak a világban, mi mozgatja a dolgokat, és milyen jövő vár rájuk. Jelen ismereteink szerint mi vagyunk az egyetlen élőlények, akik tudják magukról, hogy meg fognak halni, ezért agyi aktivitásunk nagy részét is arra fordítjuk, hogy ezt az eseményt a lehetőségekhez mérten elodázzuk. Ugyanakkor nemcsak azt szeretnénk tudni, hogy mi és hogyan lesz (mi a sorsunk, végzetünk), hanem hogy mi és hogyan volt (mi a magyarázat a már megtörtént jelenségekre), és ezt a kettőt egy narratívában akarjuk egyesíteni. Időnk végeessége rákényszerít minket arra, hogy strukturáljuk azt, jelentést adjunk neki. A narratíva egy másik fontos szükségletünket is kielégíti: a bizonytalanság és a kockázat csökkentését. Tudatosan és tudattalanul is rengeteg időt és energiát fordítunk arra, hogy mérlegeljünk, döntéseket hozzunk, aztán pedig a döntések lehetséges következményeit mérlegeljük, és újabb döntéseket hozunk. Kapcsolatunk a világgal és a benne rejlő lehetőségekkel alapvetően formálta a valószínűség fogalmát, később pedig az ellentétes folyamat is beindult: a valószínűség formálta a következtetéseinket, döntéseinket.

Pár ezer éves feljegyzett kultúránkban számtalanszor próbálkoztak ezzel az emberek a tudomány és a vallás keretein belül, de maga a folyamat nem mindig alapos átgondolás mellett zajlik. Sokszor teljesen öntudatlanul mérlegelünk lehetőségeket, vonunk le következtetéseket, vagy hozunk döntéseket. Ebben a könyvben főleg a tudomány által adott

válaszokat vesszük górcső alá. Az oksági és valószínűségi magyarázatok és ezek minden arányú keveréke átszövik mindennapjainkat: tudatosan és mindenféle átgondolás nélkül is használjuk őket.

Véletlen, bizonytalan, kockázatos, valószínű, sanszos, random, várható, esélyes, szerencsés... ezek mind hasonló, mégis egy kicsit más jelentésű szavak: vannak közöttük pozitív kicsengésűek, és arra utalók, hogy egy esemény inkább bekövetkezik, mint nem (például *valószínű, sanszos, esélyes*), illetve negatívabbak, mint a *bizonytalan, random, kockázatos*. Sokszor azonban nem világos, hogy pontosan mit értünk alattuk, nem egyértelmű a használatuk. Ez persze nem csak a magyar nyelvben van így, az 1960-as években a CIA-nál dolgozó Sherman Kent egy egész protokollt fejlesztett ki az ilyen jellegű kifejezések korrekt alkalmazására:

certain (biztos) = 100%
almost certain (majdnem biztos) = 93%
probable (valószínű) = 75%
chances about even (esélyek kb. egyenlőek) = 50%
probably not (valószínűleg nem) = 30%
almost certainly not (majdnem biztosan nem) = 7%
impossible (lehetetlen) = 0%

Konvenciója végül nem épült be a mindennapokba (még a CIA-ba se), pedig használatával ma is sok félreértést el lehetne kerülni.

De nem egyszerűen a szavak megfelelő használatáról van szó: ebben a fejezetben áttekintjük a fogalom legkorábbi megjelenéseit és kialakulásának körülményeit, amelyeket többnyire a végzettel, a bizonytalansággal és a szerencsével asszociáltak. A valószínűség jelentésének alakulása a mai napig hat arra, hogy mi hogyan érzékeljük azt.

Ahogy azt Pierre Redmont megfogalmazta:

„Az emberek [...] hajlamosak egy végzetes és vakon működő erőt társítani a jóhoz, a rosszhoz és általában minden történéshez ebben a világban. Azt hiszik, hogy Fortunát le kell kenye-rezniük, hogy kedvezzen nekik; ezért én úgy gondolom, hogy mindenkinek (nem csak a szerencsejátékosoknak) hasznos lenne tudnia, hogy a valószínűségeknek vannak szabályaik, és ezek a szabályok megismerhetők. Ismeretük hiánya hibákhoz vezet, amelyek miatt nem a sors, hanem saját maguk okolható.”¹

Már a görögök sem vágták

A valószínűség fogalma kivételesen nem az ókori görögöktől származik. A *probabilitas* (valószínűség) szó etimológiája a latin *probabilis* → *probabilitatem* (hihető, plauzibilis, jóváhagyható, megbízható), aminek a töve *probo*, ebből származik a *próba* szavunk is. Egy vélemény hihetőségét, igazolhatóságát jelentette, hogy kiállja-e a próbát. Bár a görögöknek is létezett rá szavuk: *eikos* (ΕΙΚΟΣ), ami azt jelentette, hogy „várható, bizonyos fokú bizonyossággal”, ők sokkal jobban szerették az abszolút igazságot, ami az ő definíciójuk szerint logikával, dedukcióval bizonyítható axiómákból. A platonisták szerint a matematikai fogalmak önálló létezéssel bírnak, és e fogalmak analízisével ismerhetők meg. A valószínűségi érvelést lenézték a korai filozófiában: Arisztotelész a *Phaidón*-ban panaszkodik azokra a filozófusokra, akik „[...] bár hihetően beszélnek, [...] nem azt mondják, ami igaz”.² Szókratész szerint pedig egy matematikus, aki valószínűségekkel érvel, geometriában nem ér semmit.

A „véletlen” ősidők óta a szerencsejátékokkal és a jövőmondással, illetve az istenek szeszélyeivel fonódott össze, nem éppen elősegítve

a valószínűség tudományos előmenetelét. A görögöknek volt szerencse-istennőjük is: *Tükhé* (Τύχη). Ő volt a szerencsejátékosok patrónája (római verziója: Fortuna), akivel a többi istenség sem mert vitába szállni. De Tükhé színre lépését megelőzte egy még ősibb mítosz: a három testvér, Zeusz, Poszeidón és Hadész sorsot vetett, hogy eldöntsék, ki mit fog uralni a világban – így kapta Zeusz az eget, Poszeidón a vizeket és Hadész az alvilágot.

Létezik még egy görög szó, amit a véletlen szinonimájaként használunk, ez pedig a *sztochasztikus*. A *στόχος* jelentése „célzás, találgatás”, de a *stochastic* szó először 1662-ben bukkant fel az Oxford University Press szótárában „találgatással, feltételezéssel kapcsolatos” jelentésben.³ Valójában a valószínűség-elmélethez nagyjából semmi nem állt készen az ókori görögök idejében, legalábbis Európában. Vannak, akik szerint a hinduk már rendelkeztek a megfelelő tudással: Arisztotelész idejében (Kr. e. a 4. században) íródott a *Mahábhárata* című eposz, amely többek között taglalja Nala király kalandjait.⁴

A leírás szerint Nala megszállott szerencsejátékos volt, aki a királyságát is elveszítette kockán. Ezután elszegődött szekérhajtónak Rituparna királyhoz, aki „birtokolta a kocka tudományát”, és egy fa egyetlen ága alapján meg tudta becsülni, hány levél és hány gyümölcs terem rajta. Nalát is megtanította erre, aki az újdonsült tudással felvértezve hazatért, és visszanyerte kockán a birodalmát.

A mítosz arra enged következtetni, hogy a hinduknak már megvolt az elégséges aritmetikai és kombinatorikai háttérük ahhoz, hogy matematikai fogalmi keretek között gondolkodhassanak valószínűségekről. Tőlük vették át az arabok, Európába pedig olasz matematikusokon keresztül érkezett meg.

Közülük Luca Pacioli szolgáltatta az első feljegyzett problémát:

A probléma egy olyan szerencsejátékra vonatkozik, amelyben két játékosnak minden fordulóban egyenlő esélye van a győzelemre. A játékosok egyenlő mértékben járulnak hozzá a nyereményalaphoz, és előre megállapodnak abban, hogy az a játékos, aki először nyer bizonyos számú fordulót, a teljes nyereményt megkapja. Tegyük fel, hogy a játék külső körülmények miatt megszakad, mielőtt bármelyik játékos győzelmet aratna. Hogyan osztjuk fel ekkor igazságosan a nyereményalapot? Hallgatólagosan elfogadott, hogy a felosztásnak valahogyan az egyes játékosok által megnyert fordulók számától kell függenie, úgy, hogy a győzelemhez közel álló játékos nagyobb részt kap a kasszából. A probléma azonban nem pusztán számítási probléma, hanem annak eldöntése is, hogy mi is a „fair” elosztás.⁵

Pacioli jó matematikusnak számított a maga idejében, ráadásul jóbarátja volt Leonardo da Vincinek, aki illusztrálta is Pacioli egyik geometriakönyvét. Da Vinci másik jóbarátja Fazio Cardano ügyvéd és matematikus volt, aki a Paviai Egyetemen és a milánói Piatti Alapítványnál is oktatott geometriát. Fia, Girolamo Cardano a reneszánsz egyik legismertebb matematikusa lett. Cardanót apja szintén ügyvédnek szánta, de ő orvosnak tanult Paviában, ahol nagyon hamar szerencsejáték-függő lett. Állítólag e szenvedélye – és híresen nehéz természete – miatt nem vették fel a Milánói Orvosi Kamarába, ezért hivatalosan nem praktizálhatott. Egy kis faluba, Saccóba költözött, ahol engedély nélkül is sikerült egy mérsékeltén sikeres orvosi praxist kialakítania. Minden nap játszott, és minden létező segítségre szüksége volt a játékasztaloknál (matekos és nem matekos segítségre: amikor valaki meggyanúsította azzal, hogy csal, akkor előkapta a kését, és megvágta vele az illető arcát). Az 1530-as évek elején átvette apjától a Piatti Alapítványnál betöltött posztot mint geometriaprofesszor, de szerencsejáték-függőségét az orvosi

munkájából finanszírozta. (Amilyen nehezen kapott végül engedélyt a praktizálásra, olyannyira keresett orvos lett később, Olaszországon túl Párizsba és Skóciába is elért a híre.)⁶ Kellőképp tanulmányozva a játékot, hamarosan meg is írta könyvét *Liber de ludo aleae* (A szerencsejátékról) címmel – egyesek szerint már 1525-ben befejezte, és aztán 1565-ben átirta az egészet –, amely az első munka volt a valószínűség-számításról. A könyv nemcsak azt taglalja, hogyan nyerhet többet az ember fogadásokon a matematika segítségével, de hathatós csalási módszerekről is beszámol.⁷

Ő volt az első, aki bevezette a klasszikus kedvező/összes képletet:

„Tételezzük fel, hogy valamely véletlen folyamatnak sok, egyformán valószínű végkimenetele van, ezek közül egyesek kedvezőek (vagyis nyerők), mások kedvezőtlenek (vagyis veszteséget jelentenek) számunkra. Ekkor a nyerő végeredmény elérésének valószínűsége egyenlő a számunkra kedvező eredmények részarányával.”⁸

Cardano nagy vívmánya az volt, hogy a valószínűségből következtetett a gyakoriságra.

Ugyan sok mindenre rájött, de Pacioli problémáját ő sem közelítette meg jól, ám kikövezte az utat a soron következő matematikusoknak. Jó száz évvel később, 1654-ben De Mére lovag felkereste Blaise Pascalt az osztozkodási problémával. Pascal nem volt biztos a dolgában, ezért megvitatta a kérdést a kor másik nagy matematikusával, Pierre de Fermat-val – sokan ettől a levelezéstől datálják a valószínűség-számítás megszületését. Nézzük a feladványt és a megfejtését:

Tekintsünk egy olyan verziót, ahol négy fordulót kell összesen megnyerni, és minden fordulót azonos eséllyel nyeri **A** és **B** játékos is (tehát például egy szabályos érmét dobunk fel minden

körben). Tegyük fel, hogy **A** játékos már nyert kettőt, míg **B** játékos egyet. Ezen a ponton szakítják meg a játékot. Hogyan kellene szétosztani a nyereményt?

Ekkor **A** játékosnak kettő, míg **B** játékosnak három fordulót kell megnyernie a végső győzelemhez, tehát maximum négy fordulót kell még lejátszaniuk. A lehetséges kimenetek:

AAAA AAAB AABA AABB ABAA ABAB ABBA **ABBB**
BAAA BAAB BABA **BABB** BBAA **BBAB** **BBBA** **BBBB**

Ezek mind egyformán valószínűek, és ebből öt esetben (félkövér betűs) nyer **B**, és 11 esetben **A** játékos, tehát 5:11 arányban kell elosztani a nyereményt. Felmerülhet (jogosan), hogy a fent felsorolt kimenetek mind négyfordulósak, holott *legfeljebb* négy forduló van hátra. Hiszen ha **A** játékos nyer kapásból még kettőt, akkor ott vége van a játéknak, tehát nem lesznek olyan verziók, hogy AABB, AAAB stb. Az AA-nak a valószínűsége $1/4$, a háromfordulós befejezéseké $1/8$, míg a négyfordulósoké $1/16$. Hogyan jött ki akkor a fenti számolás, és egyáltalán jó-e? Valójában az történik, hogy azokat a négyfordulósokat, amelyek hamarabb véget érnek, össze lehet vonni, tehát az AABB, AABA, AAAB, AAAA azok az AA-nak „le nem játszott alosztásai”, amelyek együttesen kiadják az AA-t valószínűségben is ($4 \cdot 1/16 = 1/4$). Azért írtunk ki minden esetet négy fordulóra nézve, hogy egyforma valószínűségű eseteink legyenek, és így lehetett használni a kedvező/összes képletet.

Ez a fajta gondolkodás nem létezett Pascal és Fermat levelezése előtt. Minden korábbi próbálkozó úgy gondolkodott, hogy csak az számít, hogy addig hány fordulót nyertek meg a játékosok (és annak

arányában kell elosztani a nyereményalapot), az nem, hogy még hányat kellett volna megnyerniük a győzelemhez. A fair elosztásnak ez az új koncepciója előrevetítette a **várható érték** fogalmát.

De Pascalnak nem egyszerűen csak a probléma megoldását – és egyéb matematikai hozzájárulásait – köszönhetjük, hanem azt is, hogy új jelentést adott a *probabilitae* (valószínűség, *probability*) szónak. Annak idején a reneszánszban a természeti jeleket bizonyítékként kezdték kezelni, és a kombinatorikai problémák sem a szerencsejátékokból, hanem a jelek alkímista értelmezéséből származnak. A jelek tehát alátámasztották a véleményeket, így azok valószínűek lettek. Ezekben az időkben a kemény tudományok – mint a mechanika vagy az optika – már elváltak az alacsonyabb rendű és tekintélyelvű orvoslástantól, illetve alkímiától, a valószínűség pedig egyértelműen az utóbbiakhoz tartozott. Így a szó maga többféle konnotációval bírt: elsősorban azt az autoritást jelölte, aki jóváhagyta a véleményt, másodsorban magát a véleményt jelentette. Harmadrészt pedig egy némileg pejoratív jelentése volt, mégpedig hogy a kérdéses állítás csak „valószínű”, és nem bizonyított (azaz nem tudományos).⁹ Pascal viszont beemelte a matematikába azáltal, hogy megalapozta a valószínűségi számításokat. Nem mintha akkortájt bármilyen tudomány különösebben elismerésre méltó lett volna: az anekdota szerint Pascalt, aki társaságban egyszer egy matematikai gondolatmenetet mutatott be, megkérdezték: „Uram, ön matematikus?” (Értsd: nem közénk való?) Mire a válasz: „Nem, kérem, én úriember vagyok.” (Pascal apja gazdag ügyvéd volt.)¹⁰

A lottót vagy megnyerem, vagy nem, az 50%, ugye?

A Cardano által bevezetett klasszikus képlet magában foglalta az **egyenlően lehetséges** fogalmát (implicite azt állította, hogy minden egyes kimenetel egyenlően lehetséges, és ezek közül a kedvezők aránya az összeshez képest a valószínűség). A megközelítés egyik legnagyobb híve Pierre-Simon Laplace volt:

„A véletlen elmélete abból áll, hogy az összes eseményt bizonyos számú egyenlően lehetséges esetre redukáljuk, azaz olyanokra, amilyenek létezésével kapcsolatban egyformán bizonytalanok lehetünk, és meghatározzuk azon esetek számát, amelyek kedvezőek az esemény számára, amelynek a valószínűségét keressük. Ennek a számnak az aránya az összes lehetséges esethez képest a valószínűség mértéke.”¹¹

Az *egyenlően lehetséges* kifejezés sok sebből vérzik, de a legnagyobb probléma vele az, hogy senki nem tudja, mit is jelent valójában. Egyrésztől felfogható úgy, hogy minden egyes kimenetel egyformán könnyen következik be, másrésztől pedig, hogy nincs okunk bármely kimenetelt előnyben részesíteni.¹² Ebből ered az **indifferencia** (vagy elégtelen ok) **elve**, ami annyit mond ki, hogy amennyiben nincs nyomós okunk mást feltételezni, akkor egy kísérlet összes kimenetelének ugyanannyi a valószínűsége. Ez viszont maga után vonta, hogy a „véletlenszerű” szinonimája lett az egyenlően lehetségesnek. Az elv a 19. századra teljesen tarthatatlanná vált, és tudatlanságra épülő megközelítésnek tartották. Ennek ellenére a mai napig megmaradt a köztudatban, ráadásul egy különösen kevésbé szofisztikált verzióban: tetszőleges esemény vagy megtörténik, vagy nem, tehát mindenre 50% esély van. Azaz ha kilépek az útra, 50% az esélye, hogy elütnek; ha veszek egy lottót, 50% az esélye, hogy megnyerem. Az emberek azzal együtt elhiszik ezt, hogy csak Magyarországon milliós nagyságrendű embertömegek lottóznak, mégsem hallunk olyan híreket, hogy félmillió ember nyerte meg az ötös lottót a 43. héten.¹³

A fentiekén túl van még néhány szembeszökő probléma az indifferencia elvével: az egyik például az, hogy ha végtelen sok kimenetelünk van, akkor mi lesz az egyes kimenetek valószínűsége? Itt persze fel lehet hördülni, hogy mégis milyen forgatókönyv adódik, amikor végtelen sok kimenetel van? A legegyszerűbb példa a darts: a táblának mint kör lapnak

végtelen sok pontja van, így a dobás lehetséges kimeneteleinek száma is végtelen. Mekkora valószínűséggel fogjuk eltalálni a céltábla egy konkrét pontját? A klasszikus kedvező/összes képlethez folyamodva a nevezőben végtelen lesz – a számlálóban meg az az egy kimenetel, hogy eltaláltuk a pontot –, ez pedig nullát fog adni. A kézenfekvő ellenérv a fenti nulla valószínűségre, hogy a dartstábla (vagy tetszőleges fizikai objektum) csak véges sok eltalálható pontból áll, és a dartsnyíl hegye sem nulla kiterjedésű. Ekkor tehát a valószínűség nem nulla, csak nagyon kicsi. Ezt sokáig egyszerűen azzal intézték el, hogy az ilyen kis valószínűségű eseményeket fizikailag lehetetlennek tekintették.*

Ha egy ennél realizistikusabb kérdést akarunk feltenni, azaz hogy nem egy konkrét pontot találunk el, hanem mondjuk a bull's eye-t, akkor a középpont körüli egy centi sugarú körlapot vesszük, az ebben lévő pontok száma lesz a kedvező eseteink száma. Viszont ez is végtelen, és ebben az esetben végtelen/végtelen-hez vezet a képlet, ami nem értelmezhető aritmetikailag.

Mit jelent akkor az egyenlően lehetséges a végtelen sok kimenetel esetén? Ehhez nézzük a következő paradoxont, amely Joseph Bertrand-tól származik, és 1889-ben jelent meg:

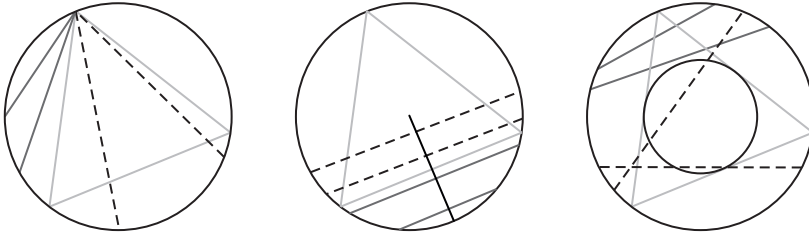
Adott egy kör és benne egy szabályos háromszög. Ha véletlenszerűen választunk egy húrt, akkor mi a valószínűsége, hogy az hosszabb lesz, mint a háromszög oldala? (Lásd 1. ábra.)

Ekkor a fenti elv alapján a „véletlenszerű” annyit tesz, hogy egyik húrt sem tüntettük ki, azaz egyenletesen választottunk. De mit is jelent ilyen esetben az egyenletes választás?

Az 1. megoldásnál lerögzítjük a véletlen húr kezdőpontját a háromszög egyik csúcánál (hiszen az egész ábra forgásszimmetrikus, tehát nem élünk semmilyen megszorítással), így pontosan akkor lesz

* Ennek az Antoine Augustin Cournot-ról elnevezett elvnek a népszerűsége egészen az 1900-as évek közepéig, a valószínűség-elmélet mértékelméleti megalapozásáig töretlen volt.

hosszabb a húr, ha a háromszöget metszi. Ez annyit tesz, hogy a körívről választottunk egyenletesen (a húr másik végpontját). A háromszög három részre osztja a körívet, és a középső ív lesz a kedvező, tehát a valószínűség $1/3$.



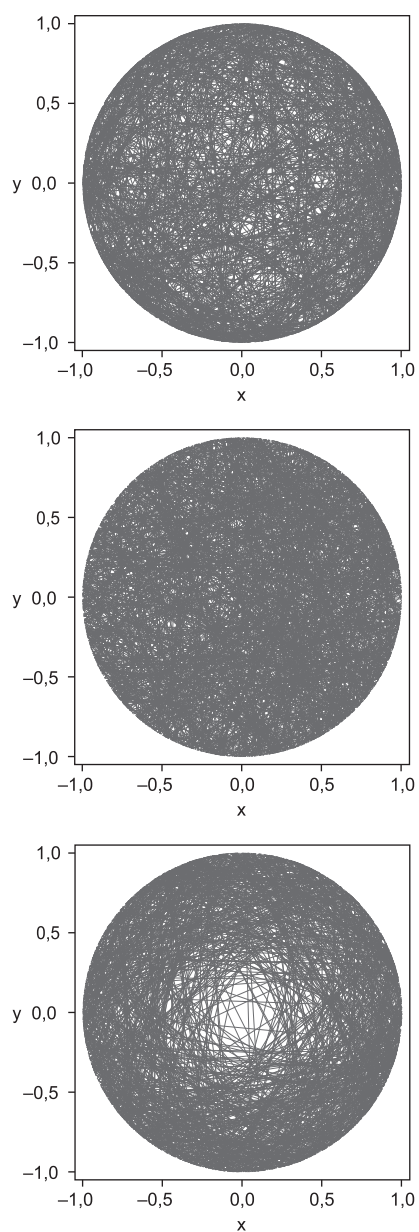
1. ábra A Bertrand-paradoxon megoldásai

A 2. megoldásnál berajzolunk egy olyan sugarat, amely merőleges a háromszög egyik oldalára, ez az oldal pont felezi a sugarat. Ekkor nézzük a sugárra merőleges húrokat (itt megint nem éltünk megszorítással, mert forgatva a háromszöget, bármely húrra teljesül): pontosan akkor lesz hosszabb egy húr az oldalnál, ha a sugár belső, háromszögbe eső felén van. Ekkor a sugárról választottuk egyenletesen a húr középpontját, a valószínűség így $1/2$.

A 3. megoldásnál a véletlen húrt megint a középpontjával azonosítjuk. Ebben az esetben a körlapról választunk egyenletesen: akkor lesz nagyobb a húr, ha a középpontja a háromszögbe írható körbe esik. Ennek a körnek a területe pedig egynegyede az eredeti nagy körnek, ez lesz a valószínűség is: $1/4$.

Az egyenlően lehetséges választás tehát nem definiálja a folyamatot, ahogy választunk: az első esetben a körvonalról, a második esetben a sugárról, a harmadik esetben pedig a körlapról veszünk egy pontot egyenletesen – ezek mind véletlenszerűek, de ahogy láhattuk, különböző eredményre vezetnek.* A Bertrand-paradoxon körül a mai napig nem

* Valójában végtelen sok legitim módszer létezik véletlen húr választására. (O. K. Bower: „Note Concerning Two Problems in Geometrical Probability”. *The American Mathematical Monthly*, 1934. október, 41[8], 506–510).



2. ábra A Bertrand-paradoxon megoldásai alapján generált véletlen húrok

ültek el a viták: egyesek szerint a kérdés aluldefiniált – azaz nem elég konkrét a megfogalmazás, hogy „véletlenszerűen választunk egy húrt” –, mások szerint egyértelműen előnyben lehet részesíteni egyik vagy másik módszert.¹⁴ Ha ránézünk a három stratégia szerint generált véletlen húrokra, egyértelműen látható a különbség (lásd 2. ábra).

A középső a leguniverzálisabb abban az értelemben, hogy ha kiválasztunk a körlapon belül egy kisebb körlapot, akkor az azt metsző húrok eloszlása nagyjából ugyanúgy fog kinézni, mint az eredeti körlapon, vagyis ez a módszer érzéketlen a skálázásra és az eltolásra. Ha tehát választanunk kell a módszerek közül, és az egyenletesség a szempont, akkor a sugárpont választása tűnik a legjobbnak.

Az egyenletességre való törekvésnél és az indifferencia elvénél tagadhatatlan a cardanói örökség, különösen amikor minden feladat (szabályos) kockákról, érmeokről és kártyákról szól. Talán a legnagyobb probléma vele a szubjektivitása: különböző emberek szemében mást jelenthet a „nyomós ok”.

Nagy számok törvénye kis számokra?

A kor gondolkodói jellemzően megfigyelésekből és kísérletekből próbálták logika segítségével következtetéseket levonni, és a klasszikus képletet elvetették. Új megközelítésre volt szükség. Az 1603-as különösen pusztító pestisjárvány után elkezdtek heti rendszerességgel lejegyezni a halálozásokat és születéseket is Londonban, ebből pedig John Graunt 1662-ben levonta az első komoly demográfiai következtetéseket. Születőben volt a statisztika, és ezzel együtt formálódott a valószínűség egy új koncepciója: a **relatív gyakoriság**.

Ugyanez foglalkoztatta a híres matematikust, Jakob Bernoullit. Bernoulli 1655-ben született Bázelen, apja kereskedelmi bankár volt, aki református lelkésznek szánta a fiát. Így ő teológiát és bölcsészetet tanult, egészen addig, amíg bele nem botlott Eukleidész egyik könyvébe. Állítólag ez az eset és a csillagászat iránti érdeklődése irányította

végül a matematikai pályára. Rengeteget utazott, és mindenhol szert tett tudós ismerősökre: Gottfried Wilhelm Leibnizcel kezdett levelezni, aki akkortájt dolgozott a differenciálszámításon, illetve Christiaan Huygens tett rá nagy benyomást.¹⁵ Jakobot 1687-ben nevezték ki a Bázeli Egyetem matematika tanszékének vezetőjévé.*

Bernoulli leghíresebb műve az *Ars conjectandi* (A találgatás művésze), amely posztumusz jelent meg 1713-ban. A könyv első fele tisztán matematikai kérdéseket boncolgat, de az utolsó (befejezetlen) részében a lehetséges orvosi, jogi és meteorológiai alkalmazásokat veszi számba. Például hogy mi a valószínűsége, hogy egy bizonyos kórokozó lakik az ember testében, ha adva van a kórokozók okozta fertőzések gyakorisága (orvosi feljegyzések révén).¹⁶ Eredeti modelljében egy urnában különböző színű golyók vannak, húzunk közülük visszatevéssel, és feljegyezzük az eredményeket. ** Bernoulli tehát a húzásokban előforduló gyakoriságokból akart következtetni az urnában eredetileg fennálló arányokra. Minél több golyót húzunk ki, a relatív gyakoriság *morális bizonyossággal**** a tényleges valószínűséggel fog nagyjából megegyezni. Ez Bernoulli arány tétele, vagy ismertebb nevén a **nagy számok törvénye**.

Ha egy szabályos pénzérmét n -szer feldobunk, és az n feldobás során k -szor kapunk fejet, akkor a dobások számának növelésével a k/n hányados – a fej dobásának relatív gyakorisága – egyre nagyobb valószínűséggel egyre pontosabban megközelíti az $1/2$ -et.

* A Bernoulli család több híres matematikust is kinevelt: Jakob volt az első tehetség a családban, és az egyetem egyik épülete az ő tiszteletére viseli a Bernoullianum nevet a mai napig. Azoknak a Bernoulliknak a száma, akik jelentősen hozzájárultak a matematikához, körülbelül 8 és 12 közé tehető, és akire én Jakobként hivatkozom, Jacob, Jacques és James néven is ismert, ez valahogy nem könnyítette meg a megkülönböztetésüket.

** Az *urna* szó eredetileg vizeskorsót jelölt, majd hamvvedret, végül pedig sorshúzásra, illetve szavazócédulák tárolására használatos edényt jelentett. Manapság a temetkezési urna és a szavazóurna jelentések élnek.

*** Bernoulli kedvenc szavajárása, ez nála 99,9%-os valószínűséget jelentett.

Arra a kérdésre, hogy a kísérletsorozatból milyen pontosan tudjuk a háttérben meghúzódó valószínűségeket kiszámítani, nem tudott felelni. Ehelyett azt válaszolta meg, hogy ha adott egy valószínűség, akkor az milyen pontosan fog visszatükröződni a relatív gyakoriságban. Tehát ha van egy szabályos kockánk, amely a klasszikus képlet szerint $1/6$ valószínűséggel mutat 6-ost, akkor ez hogyan fog visszatükröződni a dobássorozatban.

Bár az eredeti kérdést nem tudta megválaszolni, a nagy számok törvénye így is nagy horderejű állítás, és a mai napig az egyik legismertebb és leginkább félremagyarázott tétel a valószínűség-elméletben. Az arany tétel tehát azt garantálja, hogy egy populációból vett nagy minta reprezentálni fogja a populáció összetételének arányát.

„Már ötször egymás után fekete volt, most már biztosan piros fog jönni!” – halljuk Gézától a rulettasztalnál. „Eddig négy lány jött, ennek muszáj lesz már fiúnak lennie!” – halljuk a szülőszobán. „A héten harmadszorra léptem kutyakakiba, most már valami jó fog történni velem!” – halljuk a fejünkben.

Ha megkérjük Gézát, hogy fejtse ki az elméletét, akkor valami ilyesmit fog mondani:

„Ha már sokszor volt fekete, akkor szükségképpen megnő a piros valószínűsége, különben nagyszámú pörgetés után nem lehetne kb. ugyanannyi a pirosak és a feketék száma.”

Géza szerint tehát a rulett igazságossága feljogosítja arra, hogy azt várja, hogy az egyik irányba való eltérés gyorsan korrigálja magát a másik irányba történő megfelelő eltéréssel. Ugyanakkor a rulett emlékezőtehetsége

és erkölcsi érzéke is erősen behatárolt, minden pörgetésnél ugyanannyi a piros és a fekete valószínűsége, a korábbi pörgetésektől függetlenül.¹⁷ Géza fenti érvelését, a nagy számok törvényének ezt a logikai hendikepes öccsét **kis számok (hamis) törvényének**, illetve **szerencsejátékosok tévedésének** nevezik,^{*} és megfigyelhető különböző ezoterikus verziókban, a rulettet lecserélve „rendszerre” vagy „életre”. Sokan elhiszik, hogy ha sok rossz dolog történik velük, akkor az univerzum hamarosan a segítségükre siet, hogy kárpótolja őket, azaz helyreállítja a jó és rossz dolgok számának egyensúlyát.^{**} Egy kísérlet folyamán¹⁸ 6 és 15 év közötti gyerekeket kértek meg, hogy bináris kimenetelű sorozat (például érmedobás) eredményeit jósolják meg, és minél fiatalabb volt az illető, annál jobban érvényesült a kiegyenlítésre való vágy a tippelt sorozatban. Tehát hiába volt szó nagyon rövid (például 4 hosszúságú) sorozatról, a gyerekek hajlottak rá, hogy fele-fele arányra tippeljenek.

A relatív gyakoriság intuitíve nagyon ígéretes, és a 19. században heroikus erőfeszítéseket tettek a tudósok, hogy ennek a koncepciónak a segítségével adjanak végleges választ a valószínűségi kérdésekre. A mozgalom frekventizmus néven futott: a frekventisták egy esemény valószínűségét kizárólag egy kísérletsorozatban előforduló gyakoriságnak tekintették. 1866-ban John Venn a fiú/lány születési statisztikák alapján kimondta a frekventizmus alaptételét: a valószínűség nem más, mint arány.¹⁹ Tehát szerinte a nagy számok törvénye egy félreértés, mert nincsenek valószínűségek, amelyekhez közelítenek az arányok, hanem maguk az arányok a valószínűségek.

A frekventisták két csoportra oszlottak: véges és hipotetikus. A véges frekventisták igazi hardcore empiristák voltak: csakis a megfigyelhető – tehát egyben véges – sorozatokra koncentráltak, a bekövetkező események relatív gyakoriságával azonosították a valószínűséget. Hasonlóan gondolkodtak, mint rulettes Géza: számukra az, hogy 10 dobásból 3 lett 6-os, annyit jelentett, hogy a 6-os valószínűsége $3/10$ (már amennyire

* Monte-Carlo-hiba néven is fut.

** Ez a babona persze fordítva is működhet: ha sok jó dolog történt valakivel viszonylag rövid idő alatt, akkor elkezdhet félni attól, hogy „az élet majd nemsokára megbünteti”.

hajlandóak voltak egyáltalán azzal az állásponttal azonosulni, hogy szinguláris eseményeknek – mint egy darab kockadobás – szükség-szerűen lenne valószínűsége).²⁰ A hipotetikus frekventisták végtelen sorozatokat tekintettek, és azt állították, hogy azokban a relatív gyakoriság határértékben tart a valószínűséghez. Tehát ha végtelenül sokszor ismételnénk a kockadobálást, akkor az ott megjelenő frekvenciát tekinthetjük valószínűségnek. Az egy dolog, hogy nem tudunk végtelen sokszor eldobni egy kockát, az is felmerül, hogy a kocka időközben elkopna. Vajon ezt hogyan vennénk figyelembe egy ilyen sorozatnál?

Tehát ha egy fizikai világra (és főleg emberi célokra) szeretnénk a valószínűségeket használni, akkor az aszimptotikus interpretáció nem lesz kielégítő: ez így is lett, a frekventista megközelítés sokáig tartotta magát, de nem bizonyult kielégítő valószínűségi modellnek. Szemben a nagy számok törvényével, miszerint a nagy mintáknak reprezentálniuk kell a mögöttes arányt, a kis mintáknak nem. Ha elhinnénk a kis számok törvényét (ne higgyük el), azaz hogy már kis mintáknak is reprezentálniuk kell a mögöttes relatív gyakoriságot, akkor ebből az is következik, hogy hosszú egyforma sorozatok egyáltalán nem is fordulhatnak elő. Kérjünk meg azonban valakit, hogy dobjon fel egy szabályos érmét kétszázszor, egy másik valakit (például saját magunkat) pedig, hogy állítson elő egy randomnak tűnő 200 hosszú fej-írás sorozatot. Ekkor ha összevetjük a két sorozatot, az ember által kreáltban általában rövidebbek a csupa fej, illetve csupa írás sorozatok, azaz nem merjük egyforma szimbólumok hosszú sorozatát megengedni, mert az már nem tűnik számunkra véletlenszerűnek.²¹ Ha azonban 200-szor feldobunk egy szabályos érmét, akkor számíthatunk rá, hogy lesznek benne akár 8 hosszúságú egyforma fej (vagy írás) sorozatok. Tipikusan szinte senki nem mer ilyen hosszú egyforma szekvenciákat leírni egy mesterségesen előállított sorozatban.

Érdekes módon viszont a sportban egyáltalán nem tartjuk furcsának a hosszú egyforma sorozatokat. Azaz ha egy játékosnak nyerő szériája van, akkor nagyobb valószínűséget társítunk ahhoz, hogy a következő ütése/dobása is sikeres lesz. Ezt hívjuk **forró kéz jelenségnek** (*hot hand*

phenomenon). A kifejezés az Egyesült Államokból ered, a baseball területéről (de más labdajátékokban, például a kosárlabdában is ismert). Amikor valaki egymás után sokszor üt sikeresen, azt **forró sorozatnak** (*hot streak*), az illetőt pedig **forró kezűnek** (*hot hand*) nevezik. Ebben rengeteg szurkoló, edző, sőt sokszor még maguk a játékosok is hisznek: mivel a sportolóknak nem állandó a teljesítményük, miért ne lenne lehetséges, hogy megmámorosodva a nyerő sorozattól még több nyerőt fog ütni? Ezt járta körül 1985-ben Thomas Gilovich, Amos Tversky és Robert Vallone:²² 100 kosárlabda-szurkolót kérdeztek ki, és 91%-uk állította, hogy egy játékosnak több esélye van a következő ütésre, ha az előző kettőt vagy hármat eltalálta, mint ha nem. A tanulmányban profi kosárlabdajátékokat, illetve szabaddobás-sorozatokat is elemeztek statisztikailag, amiből az derült ki, hogy a hiedelemnek nincs alapja, azaz egy nyerő szériát nem követ magasabb valószínűséggel egy újabb nyerő dobás. Bár ránézésre ez a kognitív torzítás éppen az ellentéte a szerencsejátékosok tévedésének, mégis van egy erőteljes párhuzam: az abban való hit, hogy a hosszú egyforma sorozatokat nem okozhatja a véletlen. Ebben az esetben tehát a nyerő széria nem lehet a véletlen műve, hanem a játékos valami különleges képesség birtokában van, és így a következő dobás sikeressége is nagyobb valószínűségnek örvend.

Vizsgáljuk meg közelebbről ezeket a csomósodásokat!