

Dusek Tamás

A valószínűségek ingoványos terepe

Keszthelyi Gabriella: *Milyen színű a valószínű?*
A véletlen matematikája. Typotex, Budapest, 2025

Dusek Tamás, a Széchenyi István Egyetem egyetemi tanára, a *Statistikai Szemle* főszerkesztője
E-mail: dusekt@sze.hu, Tamas.Dusek@ksh.hu

Jelen ismertetés tárgyául szolgáló könyv szerzője Keszthelyi Gabriella matematikus, a BME Sztochasztika Tanszékének adjunktusa. Az *Előszó* szerint a könyv írásának egyik motivációja az volt, hogy „a valószínűségek világa ingoványos terep sok matematikus számára is, a laikusoknak pedig egyértelműen az” (7.). Egyetértek. Hiába cáfolták sokszor, sokféleképpen a véletlennel és a valószínűséggel kapcsolatos téves elképzeléseket és gondolkodási hibákat, ezek állandóan újratermelődnek, visszatérnek, olykor új elnevezések születnek rájuk, ezért rendszeresen újra és újra el kell mondani a problémákat, régebbi és aktuális példákkal megvilágítva azokat. A könyv gazdag anyagából ismertetésemben néhány részt kiemelve tárgyalok. A felsorolásszerű, száraz beszámolón túl hasznosnak tartom egyes témák kiegészítését, nem a hiányokra való rámutatás céljával, hiszen minden tárgyalás szükségszerűen szelektív lesz, hanem azért, mert ez jobban hozzájárulhat a kérdéssel kapcsolatos ismeretek elmélyítéséhez. Ezeket részben a könyv tartalmának sorrendben történő körvonalazása közben, részben ezt követően teszem meg. A megemlített koncepciók tartalmát nem mindig részletezem, ez néha az eredeti munka terjedelmét igényelné.

A valószínűségszámítással, a véletlennel, a bizonytalansággal és az ezekhez kötődő kérdésekkel és döntési szituációkkal mindenki nap mint nap találkozhat a hétköznapi életben. Nem mindegy, hogy a döntések helyes gondolatmeneten alapulnak, vagy számolási vagy gondolkodási hiányosság által terheltek. A valószínűségi gondolkodás és a valószínűségszámítás bevonását igénylő problémák megjelennek a tudomány számos területén is, ahol szintén ügyelni kell a valószínűséggel kapcsolatos koncepciók helyes kezelésére. A könyv témája rendkívül

széles körű, a gyakorlat számára fontos általános és szakmai kérdésekkel áll kapcsolatban. Nem tankönyv, amely az elmúlt évszázadokban felhalmozott absztrakt, matematikai jellegű tudást az alkalmazásoktól elszakítva ismerteti, és nem is a valószínűségszámítás tudományfilozófiai kérdéseit boncolgató szakkönyv. Inkább az ilyen tankönyveket, szakkönyveket kiegészítő és elmélyítő, a valószínűségszámítás alkalmazásaival és egyes paradoxonjaival intenzíven foglalkozó, tudományos ismeretterjesztő jellegű munka, amely érdekes és hasznos lehet a nagyközönségnek és a szakembereknek is – a statisztika- és a matematikatanárok is bizonyára nagyon izgalmasnak fogják találni –, mindenki számára olyan újdonságokat nyújtva, amelyek az egyedi előismeretek mennyiségétől és jellegétől függően eltérők lehetnek. Tartalma és stílusa miatt könnyen bekerülhet az ajánlott olvasmányok közé is egyes statisztikai és valószínűségszámítási kurzusokon.

A kötet a magyarul korábban megjelent könyvek közül számomra leginkább Warren Weaver *Szerencse kisasszony. A valószínűség elmélete* (első magyar kiadása: 1979) című könyve tematikájához és megközelítéséhez hasonlít: kevés matematikai háttér (ami nem egyenlő a szakmai tartalom csekély voltával), a koncepcionális kérdések hangsúlyosak, a klasszikus valószínűségi példák, történetek pedig sokszor közösek a két kötetben. Ugyanakkor Keszthelyi Gabriella példái nagyrészt a jelenhez közelebbiek, egy részük az adatforradalomhoz kötődik, jóval Weaver könyvének (amelyet angolul először már több mint 60 éve, 1963-ban adtak ki) megjelenése utániak. Kisebb mértékben Hurgin *Mindennapi döntéseink és a matematika* (magyar kiadás: 1981) című könyve is megemlíthető elődként, több közös témával, valamint a viszonylag újabbak közül Leonard Mlodinow *Részig bolyongás – Hogyan irányítja a véletlen az életünket?* című munkája (magyar kiadás: 2012). Az irracionális döntések tárgyalásában Stuart Sutherland *Irracionalitás* című könyve – amely részletesen foglalkozik a valószínűségszámítási, statisztikai ismeretek szerepével is – tekinthető jelentős előzménynek a tudományos ismeretterjesztő irodalomban (magyarul 2010-ben jelent meg), az orvoslásban jelentkező bizonytalanságokról, valószínűségekről, okságról pedig Siddhartha Mukherjee-től *Az orvoslás törvényei* (magyarul 2017-ben adták ki).¹

Keszthelyi Gabriella munkája 8 számozott fejezetre, valamint *Előszóra, Bevezetésre, Függelékre, Köszönetnyilvánításra és Jegyzetekre* tagolódik. A *Bevezetés* a számozott fejezetekhez hasonló terjedelmű, és a valószínűségszámítás történe-

¹ Az említett könyvek bibliográfiai adatai: Hurgin, J. I. (1981): *Mindennapi döntéseink és a matematika*. Műszaki Könyvkiadó, Budapest; Mlodinow, L. (2012): *Részig bolyongás – Hogyan irányítja a véletlen az életünket?* Akkord Kiadó, Budapest; Mukherjee, S. (2017): *Az orvoslás törvényei. Feljegyzések egy bizonytalan tudományterületről*. HVG Könyvek, Budapest; Sutherland, S. (2010): *Irracionalitás. Gondolkodásunk tévútjai*. Noran Libro Kiadó; Weaver, W. (1979): *Szerencse kisasszony. A valószínűség elmélete*. Gondolat Könyvkiadó, Budapest.

tének egy részét érinti, néhány kombinatorikai és valószínűségszámítási példa mellett. A *Már a görögök sem vágták* című alfejezet Cardanóhoz kötődően mutatja be a klasszikus valószínűség fogalmát. Az ezt követő alfejezet címe: *A lottót vagy megnyerem, vagy nem, az 50%, ugye?* Ebben az olvasható, hogy „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” (17.). Pontosabbnak gondolom az „egyenlőtlen lehetséges” kifejezést az eredeti megfogalmazási körülményekből megérteni: ezeket olyan szerencsejátékok (kockajáték, kártyajáték, érmedobás, lottó) kapcsán fogalmazták meg, ahol egyszerű helyzetekben közvetlenül, bonyolultabbakban kombinatorikai eszközökkel megszámlálható a kedvező esetek, valamint az összes eset száma. Szabályos körülmények között (pl. szimmetrikus kocka, kocka felrázása egy pohárban) az egyes esetek esélye ugyanakkora (egy osztva az összes esettel). Az igaz, hogy ezt a valószínűségértelmezést nem lehet általánosítani, de ahol megfogalmazták a 16–17. századtól kezdve (Cardano kézírata a szerencsejátékokról posztumusz jelent meg, 87 évvel halála után), ott alkalmazható ma is, de más olyan területen is, amely analóg a szerencsejátékokkal. Ennek megfelelően nem gondolom azt, hogy „az elv a 19. századra teljesen tarthatatlanná vált, és tudatlanságra épülő megközelítésnek tartották”, bár a mondat második fele annyiból lehet érvényes, hogy lehettek az elvnek indokolatlan bírálói. A szerző szerint az elv a mai napig megmaradt a köztudatban egy „különösen kevésbé szofisztikált verzióban”, és az emberek elhiszik azt, hogy „ha veszek egy lottót, 50% az esélye, hogy megnyerem” (17.). Nem kétlem, hogy létezik torz valószínűségérzékelés, például a később (33.) bemutatott hozzáférhetőségi heurisztika miatt. Ilyen mértékű torzítás feltételezése azonban túlzónak tűnik számomra.

Az első fejezet (*Csomósodások és illúziók*) a véletlen és a nem véletlen jelenségek megkülönböztetésének kérdésével kezdődik, de nagyon sok más témával is foglalkozik, például hogy az emberek racionálisak-e, azaz a meglévő információk alapján mindig logikusan és optimálisan döntenek-e. (Ez az egyik meghatározása a racionalitásnak, amelynek vannak más felfogásai is.) A szerző számos ismert (több közülük Kahneman és Tversky által megfogalmazott) kognitív torzítást mutat be, amelyek akadályozzák a logikus és optimális döntéseket, és többnyire az információk hozzáféréssel és feldolgozásával kapcsolatosak, mint például a hétköznapi életben fellépő hozzáférhetőségi heurisztika és a gyakorisági illúzió.

A nagy adatbázisokkal, sok változóval zajló statisztikai elemzéseknél előfordulhat a többszörös összehasonlítási probléma. Erre a szerző egy svéd tanulmány példáját mutatja be, ahol a távvezetékek közelében élő emberek körében vizsgál-

ták több mint 800 betegség arányát, és azt, hogy van-e olyan betegség, amely szignifikánsan sűrűbben fordul elő az itt lakók körében, mint a nem távvezeték közelében (36–37.). A probléma lényege, hogy 800 lehetséges kapcsolatvizsgálatnál a véletlen ingadozás hatására is könnyen előfordulnak statisztikailag szignifikáns, de nem valós eredmények. Ugyanakkor a probléma kezelhető, mert sok távvezeték van, a vizsgálatot sokszor meg lehet ismételni, ami alapján eldönthető (a sok változó miatt nem feltétlenül könnyen), hogy véletlen vagy tényleges hatásról van-e szó. A többszörös összehasonlítás problémája mellett egy ettől eltérő problémaként van bemutatva az adatkitermelés, más néven *data dredging*, *data snooping*, *data butchering*, p-hekkelés (38.). Az elnevezések sokfélesége mögött szerkezetileg ugyanaz a jelenség áll: a nagyon sok változó közötti még több (számuk, ha a hatás irányultságától eltekintünk: $n \cdot (n - 1) / 2$) lehetséges kapcsolatból azok lesznek kiválogatva és a publikációkban közölve, amelyek között szignifikáns kapcsolat mutatkozott. Ez a lényege a többszörös összehasonlítás problémájának is. A többszörös összehasonlítás problémája a sokváltozós elemzésekhez és jellemzően a digitalizációhoz és az adatbőséghez kötődik, de elvileg a digitális kor előtt is létezhetett volna a jelenség. Az elnevezések sokasága mutatja, hogy számos, egymástól elkülönülő elemzési terület, amellyel a statisztika foglalkozik – a csillagászatól a hidrológiai, műszaki, genetikai, klinikai, környezeti, szociológiai kutatásokon át a pszichológiai elemzésekig –, más-más terminológiát alakíthat ki. Ennek vannak negatív hatásai, például az, hogy a ténylegesnél több és többféle elemzési hibát sugalmazhat.

Az adatkitermelés illusztrációjaként a sokat emlegetett példa, Nicholas Cage filmjeinek a száma és a medencés halálokok közötti korreláció jelenik meg, amelyet a könyv 7. ábrája közöl. További érvelési hibákról is szó esik, amelyeket általánosabbnak gondolok, és elsősorban nem a valószínűségszámításhoz kötődőnek, így a texasi mesterlövész hibájáról, a megerősítési torzulásról és az adatok mazsolázásáról. „Képzeljünk el egy texasi farmert, aki azzal dicsekszik, hogy milyen jól céloz. Bizonyításként a pajta oldalába lő egy sorozatot, aztán ahol a legjobban sűrűsödnek a találatai, a köré berajzol egy céltáblát” (40.). „A fenti Nick Cage-es példát alapul véve, ha nagyvonalúan kihagyjuk az összes többi évet, amikor nem korrelált a medencébe fulladás és Cage filmjeinek száma, akkor is a texasi mesterlövészneként viselkedünk” (40.). Az adatkihagyást inkább az adatmazsolázás problémájaként írnám le, amely kevésbé jellemző, és sokkal feltűnőbb az idősoroknál, mint a nem időbeli adatoknál. A texasi mesterlövész hibáról nem gondolnám, hogy a kapcsolatvizsgálatokhoz köthető, hiszen itt egy térbeli mintázat jön létre a lövések révén. Az a besorolása, miszerint ez egy korrelációhoz tartozó torzítás, megismétlődik a 85. oldalon is. A texasi mesterlövész azonban, amennyiben nem négyzetrácsot vagy hatszögrácsot lő a pajta oldalára, vagy pont ellenkezőleg, egyetlen pontot talál el az összes lövése, egyváltozós

torzításra utal, és akkor fordul elő, ha véletlen mintázatokba és adatszoportosulásokba utólagosan magyarázunk bele valamilyen okot. A texasi mesterlövész hibájának sokféle megfogalmazása és értelmezése létezik. Két másik leírása sokkal jobban illik egy mesterlövészhez, mint az előző: az egyik szerint egyetlen lövést ad le, és a köré rajzolja a céltáblát, a másik szerint néhány lövést ad le, és minden egyes lövéshez rajzol egy célkeresztet, amivel sokkal lenyűgözőbb hatást érhet el, mint a véletlenszerűen elterülő lövések súlypontjába rajzolt céltáblával.

Az adatmazsolázás egy általános érvelési probléma statisztikai megfelelője, amennyiben egy állítást megerősítő és cáfoló érvek közül az elfogult, manipulatív szándékú elemző a narratívájának kedves megerősítő vagy cáfoló érveket válogatja ki. Összességében a bemutatott érvelési, elemzési hibák között sok az átfedés, elkövetésükre a hétköznapi életben és a tudományos kutatásban is van lehetőség. Annak eldöntése, hogy adott esetben valóban az érintett érvelési (elemzési) hibáról van-e szó, vagy inkább a sokféle érvelési hibával kapcsolatos (meta)érvelési hibáról, az körütekintést igényel és a kontextus ismeretét.

A második fejezet címe és tartalma *Okság és indukció*. Szemléletesek és hasznosak az oksági gráfok és a Bayes-tétel bemutatása. Hiányérzetet okozhat, hogy fontos alapkérdések nincsenek tisztázva, például az elméleti és a történeti magyarázatokban szerepet játszó okok megkülönböztetése. Az okság értelmezése kapcsán több kérdéses elem vetődik fel, közülük az időbeliség szerepét emelem ki. David Hume-ra hivatkozva az olvasható, hogy „ahhoz, hogy kauzalitásról beszéljünk, szükség van térbeli és időbeli szomszédságra, az oknak az időbeli elsőbbségére és az ok és okozat állandó együttjárására” (57.). „A téves oksági feltételezések gyakran úgy születnek, hogy a három kitétel közül (térbeli és időbeli szomszédság, az oknak az időbeli elsőbbsége, valamint az ok-okozat állandó együttjárása) valamelyik sérül” (60.). Majd ezután: „Nem igényel túl sok magyarázatot, hogy az oknak időben hamarabb kell bekövetkeznie, mint az okozatnak” (60.). Az okságnak többféle meghatározása és kritériuma adható meg, de következményeit tekintve ezek indokltsága eltérő. Hume felfogása ezen a területen a mai napig sok zavar forrása. Az előbbi kritérium egyik következményének érzékeltetésére nézzünk meg két példát. Ha egy diák az államvizsgája napján alkalomhoz illő öltözetet vesz fel, akkor vajon ez lesz az oka a későbbi államvizsgának? A történeti időbeli sorrendiség mint az ok elvárt kritériuma szerint ez vagy oka az államvizsgának, vagy nincs kapcsolat a két esemény között. Ha valaki vacsoravendégeket hív, amihez aznap a boltban bizonyos ételt-italt vásárol, akkor a vásárlás lesz az oka a vendéglátásnak? És így tovább, az emberi tudatosság világában ezek olyan tipikus szituációk, amelyeknél egy előre ismert időpontú későbbi esemény, amely a várakozások szerint bekövetkezik, oka lehet számos korábbi eseménynek. A természeti világban, az élettelen anyagnál, ahol nincs előrelátás és tudatosság, akár érvényesülhet is az időbeli sorrendiség kritériuma,

de az emberi cselekvés tudatos és előrelátással rendelkező világában vagy nem érvényesül, vagy csak az okság egy nagyon leszűkítő értelmű meghatározásával lehet ilyen kritériumot előírni. A jövőben várt események befolyásolják a jelenbeli tudatos cselekvést, egy jövőbeli esemény oka lehet az időben korábbi eseménynek. Az oksági viszony nem helyettesíthető matematikai függvénykapcsolattal sem, mert ez az ok-okozati összefüggés genetikus jellegét kihagyja, arról semmit nem mond (ami a Nicholas Cage-es példában és az ehhez hasonló példákban nyilvánvaló). A történeti idő és a logikai idő megkülönböztetése mellett ugyanakkor az ok mindig megelőzheti az okozatot, de ehhez az idő szokásos definícióját kell megváltoztatni. A könyv szerint az oksági érvelés szükséges feltétele az időbeli egymásutánosság, ha ez hiányzik, akkor logikai következtetésekről van szó (63.).

A harmadik fejezet a Bayes-tételt részletezi tovább, a negyedik pedig visszatér az okság, a feltételes valószínűség és az időbeliség kérdéskörére. Az oksági gráfok itt is szemléletesek, hasznosak. Számos olyan probléma van tárgyalva, amely sokváltozós (kettőnél több változós) esetben lép fel: a két kiválasztott változó közötti kapcsolat jellegét befolyásoló zavaró változók, ütköző változók, rejtett változók, amelyekre minden többváltozós elemzésben (jellemzően ilyenek a klinikai, gazdasági, szociológiai, pszichológiai, környezeti vizsgálatok) figyelemmel kell lenni.

Az ötödik fejezet a valószínűségszámítás demográfiai eredetű, a 17. századtól kezdődő fejlődésével foglalkozik. Itt egészen más kérdések vetődnek fel, mint a szerencsejátékok kapcsán tárgyalható kombinatorikus jellegű problémák. Nagyrészt a halandósági táblákról, a továbbélési valószínűségekről és az életbiztosításokról esik szó, de a fejezet végén az élettartamra vonatkozó valószínűségek általánosabban, az emberen kívüli élőlényekre és tárgyakra kiterjesztve tárgyalta. Még a nyelvek változásának időbeli leírására használt, kétes elméleti és tapasztalati megalapozottságú, spekulatív jellegű glottokronológia is szóba kerül.

A hatodik fejezet tárgyát a fair (tisztességes) játékok képezik. Ezek tárgyalása során gyakori hiányosság, hogy a játéknak a benne önként részt vevő emberek preferenciáitól, a játék nyújtotta nem pénzügyi élménytől és a hasznosságtól eltekintenek, így irracionális viselkedésként írják le a 0-nál kisebb várható értékű nyereményt szolgáltató játékokban való részvételt. A szentpétervári paradoxon sokszor tárgyalt esetének összefoglalása (ahol az 50%-os valószínűségű esemény első bekövetkezésekor a bank a 2 pénzegységnek annyiadik hatványát fizeti ki nyereményként, ahányadik alkalomkor az esemény először bekövetkezik) így hangzik: „A szentpétervári paradoxon abban áll, hogy ha valaki ilyen jellegű játékot javasol (és tegyük fel, hogy nem próbál meg átverni), akkor elvileg bármekkora árat érdemes a játékért fizetni, mégis a legtöbb résztvevő nem hajlandó 10 forintnál többet befektetni. Akkor hogy is van ez?” (155.). Warren Weaver

annyiból részletesebben találja ezt a kérdést, hogy jobban részletezi annak következményét, hogy a valóságban nincsenek végtelen kifizetési lehetőséggel rendelkező bankok (*Weaver, 1979*). A játékos számára a játszma értékét az a maximális összeg határozza meg, amelyet a bank még ki tud fizetni. A paradoxont egy korábbi könyvismertetésem, az érintett George Szpiro választáselméleti könyve alapján szintén részletesebben tárgyalja, mert bevonja azt a szempontot, hogy míg a matematikusok a pénzt a mennyiségükkel mérik, addig a játékban részt vevő emberek a pénz értékét az általa nyújtott hasznosság révén ítélik meg (*Dusek, 2023, 76–78.*).

A hetedik fejezet a normális eloszlás történetével, jellemzőivel, alkalmazásával foglalkozik (nagy nevekhez kötődően: de Moivre, Gauss, Laplace, Galton, Poincaré, Quetelet). A történeti leírások érdekesek, de az az állítás némileg bírálható, hogy „szinte minden fontos jellemzőnk a haranggörbén helyezkedik el: az egészségügyi leleteink értékei, az IQ-nk, a ruhaméretünk, az éves vízfogyasztásunk, a gyerekeink vizsgaeredményei” (193.). Ezek között van olyan jellemző is (az IQ), amely egy önkényes pontérték, nem hagyományos mérés eredménye. A pontértékek normálisához közeli eloszlását a kérdések nagy számával, nehézségük eltérő és egyenletesen változó voltával és néhány más technikával próbálják meg elérni. Az IQ normális eloszláshoz közeli volta tehát az IQ-teszt megszerkesztésének módjától függ, nem egy külső, objektív adottság. A normális eloszlás tipikus és elterjedt voltáról sokan írnak Quetelet óta, aki valóban olyan fizikai jellemzőket vizsgált (felnőttek, újszülöttek testtömege, magassága), ahol a normális eloszlástól való eltérés kicsi. Ez azonban szűk területe a jelenségeknek, hogy máshol is érvényesül-e, az tapasztalatilag vizsgálható és eldönthető kérdés. *Vargha András (2020)* a normális eloszlás pszichológiai vizsgálatokban előforduló arányára vonatkozó több olyan kutatást bemutat, amely nagyszámú változót elemezve a normális eloszlás ritka voltáról tanúskodik. Ezek közül az egyik 440 pszichológiai változót elemez a normalitás szempontjából, amelyek egyike sem volt normális eloszlásúnak tekinthető 1%-os szignifikanciaszinten. A tapasztalat azt mutatja, hogy a gyakorlati eloszlások többsége aszimmetrikus, közülük jóval gyakoribbak azok, ahol a számtani átlag a legnagyobb, a medián kisebb nála, és a módusz a legkisebb, mint fordítva. Általában az átlag fölötti kiugró értékeknek sokkal nagyobb lehetőségük van nagyon eltérni az átlagoktól, mint az átlag alattiaknak, ahol valamilyen természetes minimum ennek gátat szab. Például az egy anyára jutó átlagos gyerekszám, a települések vagy az országok átlagos lakosság-száma, vállalkozások árbevétele és egyéb méretmutatói, személygépkocsik teljesítménye és ára, sportolók, művészek jövedelme ilyen tipikus (nem ritka és kivételes) társadalmi, gazdasági jelenségek. A könyv 133. oldalán, az életkorok eloszlásának 17. századi vizsgálata kapcsán is egy ilyen irányú eloszlásra láthatunk példát: a halálozások mediánja 11,0, számtani átlaga 18,2 év (17. századi

londoni adatok alapján, amikor a módusz még 0 év volt, vagyis a legnagyobb halálozási arány a 0 évesek között volt megfigyelhető). Itt azt olvashatjuk a várható élettartamok kapcsán, hogy „ma már nagyon jól tudjuk, az átlag nem a legvalószínűbb érték és nem is feltétlenül a középérték” (133.). A 193. oldalon a normális eloszlás elterjedtségére vonatkozóan az is olvasható, hogy az átlag, a módusz és a medián ritkán esik egybe, erre az egyik példa az átlagos kereset nagyobb volta a medián keresetnél. A lábjegyzetben viszont az áll, ezzel némiképp ellentétesen és a korábbi idézethez hasonlóan, hogy gyakori tapasztalat, hogy a három jellemző (átlag, módusz, medián) egybeesik.

Azt nem lehet meghatározni, hogy az „általános” tapasztalati eloszlások hányada normális, közel normális, kicsit aszimmetrikus, vagy olyan nagy mértékben, amilyenekkel *Taleb (2012)* foglalkozik. Egyes szűkebb és tágabb vizsgálati területek (mint a pszichológia Vargha András könyvében) tapasztalati eloszlásait azonban össze lehet mérni a haranggörbével, amelyet így lehet egyfajta referenciapontként használni. De a normális eloszlás igazi jelentősége a véletlen mérési hibák és a véletlen eredetű ingadozások eloszlásának a közelítésében rejlik, nem az alapsokasági eloszlások jó közelítésében.

A nyolcadik fejezet címe *Statisztika*. 26 oldalon a statisztikának egy szűk területe tárgyalható. Nagyrészt a Fisher-féle szignifikanciateszt, a Neyman–Pearson-féle eljárás és a kettő keverékéből álló nullhipotézis-szignifikanciateszt közötti csaknem 100 éves vitasorozat egyes elemei vannak itt bemutatva. Az említett eljárások során használt terminológiák nem teljesen egységesek, de mindegyik eljárás valamilyen módon statisztikai hipotézist tesztl.

A kilencedik fejezet címe *Modern valószínűségszámítás*, tartalmában közel áll a *Bevezetés* koncepcionális és tudománytörténeti kérdéseire. A véletlen, pszeudovéletlen, kaotikus és determinisztikus jelenségek megkülönböztetése fontos, azonban a jelenségek ennél finomabb klasszifikációja is indokolható lenne, mert kérdéses, hogy a fenti kategóriákban hol helyezhetők el például a nem megismételhető egyedi események, az ügyességet, a tudatosságot, a kockázatokat, a véletlent és determinizmust elegyítő, a társadalomban gyakori szituációk. A fejezetet a valószínűségszámításhoz hozzájáruló jelentősebb magyar matematikusok munkásságának bemutatása zárja.

A könyv átfogó tartalma a fenti ismertetésből kiderül, de természetesen sok érdemi részletet nem említettem meg. A következőkben még néhány koncepcionálisan fontos elemre térek ki. Az egész munkán négy nagyobb tématerület vonul végig, egyes fejezeteken belül is keveredve: a valószínűségszámítás (koncepcionális kérdések, elmélet, alkalmazások), a kombinatorika, a döntéselmélet, az okság problémája. Ezek közül a valószínűségszámítás a leghangsúlyosabb, amely sokszor átfedésben van a kombinatorikával és a döntéselmélettel. A felosztáshoz figyelembe kell venni, hogy a matematika különböző ágainak szétválasztása nem

mindig viták nélküli. A kombinatorika különböző elemek (amelyek konkrétan nagyon sokfélék lehetnek) elrendezési, csoportosítási lehetőségeinek megszámlálásával foglalkozik, ezek lehetnek önmagukban érdekesek, valamint lehet kapcsolatuk a matematika más részterületeivel, mint a valószínűségszámítással, a gráfelmélettel, a geometriával, a számelmélettel. Emellett fontos az olyan, a gyakorlati alkalmazásokra orientált tudományágakkal való együttműködés és átfedés, mint az operációkutatás, a logisztika, a hálózattervezés, az információelmélet, a számítástechnika, ahol számos kombinatorikus jellegű és kombinatorikus optimalizációval segíthető döntési helyzet merül fel.

A kombinatorika és a valószínűségszámítás kezdete egybeolvad, amennyiben a szerencsejátékok kombinatorikai vizsgálata egyben a valószínűség három értelmezéséhez is kötődik: a kedvező esetek és az összes eset hányadosa mint (1) klasszikus valószínűség (az esetek egyenlő valószínűségűek), mint (2) gyakorisági, frekventista értelmezés (de itt hosszú sorozatok relatív gyakoriságaként van meghatározva, nem egyedi események gyakoriságaként), és mint (3) logikai értelmezés (a racionális hit foka). Az első értelmezés kötődése közvetlen, a második és a harmadik közvetett, és nemcsak a kombinatorikailag meghatározott gyakoriságok, hanem bármilyen véletlen tömegjelenség tapasztalati gyakoriságához kötődik. A valószínűség értelmezhető hányadosként, de nem minden hányados (megoszlási viszonyszám) valószínűség.

A kilencedik fejezetben a szerző felteszi azt a fontos kérdést, hogy mit hívunk valódi véletlennek. Ennek kapcsán a véletlen jelenségre felhozott egyik leggyakoribb példát, az érmedobást vizsgálja meg részletesebben, bemutatva egyes matematikusok azon álláspontját, amely szerint ez egy fizikai jelenség, semmilyen véletlenszerűség nincs benne. Ez egy érdekes kérdés, a vele kapcsolatos félreértések tisztázása valóban segíthet megérteni, hogy a véletlen elsődlegesen nem a dolgok (események) tulajdonságához, fizikai, biológiai, szellemi vagy akármilyen jellegéhez kötődik, hanem a dolgokkal kapcsolatos információkhoz és ezeknek a megfigyelő ember számára való elérhetőségéhez. Egy ügyes és gyakorlott zsonglőr, de még inkább egy tökéletesen működő érmedobógép determinisztikussá tudja tenni az érmedobás kimenetelét, amennyiben az elvárt kimenetelnek megfelelően állítják be a kiinduló körülményeket és az érmedobás erejét. A 17. századtól kezdve azonban sokan megfogalmazták, hogy a kiinduló körülmények ismeretének hiányában az egyedi érmedobások és a kockadobások kimenetele nem előrejelezhető. Popper a kockadobással kapcsolatban így vetette fel 1934-ben azt a kérdést, hogy annak kimenetele véletlen vagy nem véletlen: „Ahhoz, hogy predikciókat vezethessünk le, törvényekre és kezdeti feltételekre van szükségünk; ha nem állnak rendelkezésünkre megfelelő törvények, vagy ha nem állapíthatók meg kezdeti feltételek, akkor a tudományos előrejelzés lehetetlenné válik. A kockadobás esetében nyilvánvalóan a kezdeti feltételeket nem

ismerjük megfelelően. Ha kielégítően pontos méréseink volnának a kezdeti feltételekről, akkor ebben az esetben is lehetséges volna az előrejelzés; de a helyes kockadobás szabályai (a kockarázó edény felrázása) úgy vannak megválasztva, hogy megakadályoznak bennünket a kezdeti feltételek mérésében” (Popper, 277–278.). Amikor tehát egyes matematikusok és fizikusok determinisztikus folyamatnak kívánják láttatni az érmedobást, akkor annyiban igazuk lehet, ami a fent leírtakból következik, tehát akkor, ha ismertek lennének az érmedobásra vonatkozó információk. A véletlen jelenségre adott példákban azonban nem erre gondolnak, hanem jogosan az információhiányos érmedobásra, ahol az érme kiinduló helyzete sem ismert, és az érme pörgetését sem egy gép végzi mindig azonos erővel.

A fenti példa általános tanulsága tehát az, amellyel a szerző is összegzi a kérdést, hogy a véletlen az információ (vagy tudás, ismeret, amely itt ugyanazt jelent) elérhetőségével függ össze. Ezt Ludwig von Mises így fogalmazta meg az általa esetvalószínűségnek nevezett helyzet leírásakor: „Egy adott esemény tekintetében ismerjük az eredményét meghatározó tényezők egy részét, de vannak más meghatározó tényezők, amelyekről semmit sem tudunk. Az esetvalószínűség csak a tudásunk hiányosságában hasonlít az osztályvalószínűséghez. Minden más tekintetben a kettő teljesen különböző” (Mises, 1998, 110.).

Az érmedobás példájának másik tanulsága még általánosabb. Azt jelzi, hogy a tudomány változására vonatkozó gyakran emlegetett elképzelés, miszerint a tudomány kumulatíván növekedik és egyre tökéletesedik, kiostálva a korábbi, hibás elméleteket, elképzeléseket, a gyakorlatban nem mindig érvényesül. Popper idézete azért érdekes, mert egy jól ismert és sokat hivatkozott könyvből származik (a Google Scholar hivatkozásai meghaladták a 44 ezret 2025. október elején), és világosan leírja és megoldja a problémát. Ha a tudományos tudás az előző generációk által elért eredményeken alapulva egyre tökéletesedne, akkor 70–80 évvel Poppert követően senki nem írná le a körülmények pontosítása nélkül, hogy az érmedobás determinisztikus, és a megelőző évszázadokban tévesen használták ezt a példát véletlen jelenségek illusztrálására. Joseph Keller (1986, 191.) mégis feltette a kérdést a tanulmány első mondatában: „Miért tekintik véletlenszerűnek az érmefeldobás eredményét, annak ellenére, hogy azt kizárólag a fizika törvényei és a kezdeti feltételek határozzák meg?” Ezt követően figyelmen kívül hagyta azt a több száz éves hagyományt, amelyben benne van a válasz, és amelynek során az érmedobás mindig idealizált példa volt, tisztességes, szimmetrikus érmevel, ahol a fej és az írás valószínűsége egyenlő, és jó eséllyel egyszer sem írták le, hogy tételezzünk fel egy érmedobógépet, amelynek érmedobásai determinisztikusak, vagy egy olyan érmedobást, amelyet nagy sebességű kamerával követünk. Arról írtak korábban, hogy a valóságos érmék szabálytalanságai miatt a fej és az írás tényleges, tapasztalati valószínűsége eltérhet az idealizált,

elképzelte egyenlőségtől. Ezeknél a torzítással rendelkező érménél az a fogadást kötő, aki ennek ismeretében van, némileg (tisztességtelen) előnyt szerez az ezt nem ismerővel szemben. Amikor az érmedobást gyakorlati esetekben (sporteseményeken, sorsolásokon, előítélettől és elfogultságtól mentes döntéshozatalban) egyszerű véletlengenerátorként használják, akkor sem érmedobógépeket alkalmaznak. Lehetnek csalások, előfordulhat mindkét oldalán azonos érme, mint Rejtő Jenő *Csontbrigád*jában a mindkét oldalán fejet tartalmazó pénzdaráb, de ezek is könnyen magyarázhatók, és nem érvénytelenítik az idealizált érmedobásra és annak a sorozatára vonatkozó eredményeket.

Megjegyzéseim jól mutatják a könyv előszavában írtak érvényességét a valószínűségek világáról mint ingoványos terepről. A kötet a valószínűséggel és a véletlennel kapcsolatos számos koncepcionális, elméleti és történeti ismeretet világít meg és ad át. Emellett azon részei is inspirálók és elgondolkodtatók, amelyek valamilyen szempontból máshogyan is értelmezhetők.

Irodalom

- Dusek T. (2023): Hogyan választanak az emberek? A választáselmélet három évszázada (Könyvismertetés). *Statistikai Szemle*, 101(1), 75–80. <https://doi.org/10.20311/stat2023.01.hu0075>
- Keller, J. (1986): The Probability of Heads. *American Mathematical Monthly*, 93(3), 191–197.
- Mises, L. von (1998): *Human Action. A Treatise on Economics*. Ludwig von Mises Institute, Auburn.
- Popper, K. R. (1997): *A tudományos kutatás logikája*. Európa Könyvkiadó, Budapest.
- Taleb, N. N. (2012): *A fekete hattyú, avagy a legváratlanabb hatás*. Gondolat Kiadó, Budapest.
- Vargha A. (2020): *Normális vagy? És ha nem? Statistikai módszerek nem normális eloszlású változókkal pszichológiai kutatásokban*. Pólya Kiadó, Budapest.
- Weaver W. (1979): *Szerencse kisasszony. A valószínűség elmélete*. Gondolat Könyvkiadó, Budapest.