A VALÓSZÍNŰSÉG INTERPRETÁCIÓI

Szabó Gábor

Tartalomjegyzék

Előszó

Mit jelent az a kijelentés, hogy egy szabályos dobókockával a hatos dobás valószínűsége egy hatod?

A kérdésre az alábbi paradigmatikus válaszok adhatók:

1. *Klasszikus válasz:* Mivel szabályos kocka esetén mindegyik oldal előfordulása egyenlően lehetséges, és az esetek közül nekünk csak az egyik kedvez, ezért a kedvező esetek és az egyenlően lehetséges esetek számának aránya egy hatod lesz, és ez a hatos dobás valószínűsége.

2. *Logikai válasz:* A hatos dobás valószínűsége azért egy hatod, mert az a kijelentés, hogy az eredmény hatos lesz, egy hatod mértékben következik abból a kijelentésből, hogy a kockát eldobtuk, egy mindkét kijelentést tartalmazó nyelvben.

3. *Szubjektivista válasz:* Az, hogy a hatos dobás valószínűsége egy hatod, azt jelenti, hogy egy hatod mértékben hiszünk a kijelentés igazságában.

4. *Frekventista válasz:* A hatos dobás egy hatod valószínűsége semmi mást nem jelent, mint hogy a hatos relatív gyakorisága a kockadobások egy elegendően hosszú sorozatában közel egy hatod lesz.

5. *Propensity válasz:* A hatos dobásnak azért egy hatod a valószínűsége, mert a kocka fizikai környezetével együtt rendelkezik azzal az egy hatod mértékű kauzális hajlammal (*propensity*-vel), hogy eldobva hatos legyen.

A fenti válaszok a valószínűség öt interpretációs iskolájának szellemében születtek. Az interpretációk központi gondolata és története dióhéjban a következő.

A történetileg első interpretáció, a *klasszikus interpretáció* szerint a valószínűség *a kedvező esetek és az egyenlően lehetséges esetek számának aránya.* A gondolat Leibniztől ered, aki a definíció kényes pontját, az „egyenlően lehetséges" jelentését gazdag metafizikai eszmefuttatással igyekezett megalapozni. A 18. századra megszilárdult determinista világkép következtében azonban az „egyenlően lehetséges" kifejezést szubjektivista szempontból kezdték olvasni: két esemény akkor egyenlően lehetséges, ha egyik bekövetkezése mellett sem szól több érv. Az *indifferencia elvére* épülő valószínűség fogalma így Laplace korára szubjektív értelmezést nyert, és ez az értelmezés maradt uralkodó egészen a 19. század végéig.

A statisztikus tudományok megjelenése a 19. században azonban magával hozta az igényt egy objektív valószínűségértelmezés iránt. A *relatív gyakoriság-* vagy más néven *frekvenciainterpretáció* előfutárai cambridge-i logikusok voltak, közülük is a legjelentősebb John Venn. Venn fogalmazta meg először a férfi és női születések kapcsán a frekventizmus alaptételét: „a valószínűség semmi más, mint arány". A frekventisták szerint tehát egy esemény valószínűsége *bekövetkezésének relatív gyakorisága egy elegendően hosszú eseménysorozatban*. Ennek az értelmezésnek egyenes következménye lett azután az a meggyőződés, hogy a valószínűség fogalma csak eseménytípusokra vonatkozik, egyedi eseményekre pedig legfeljebb áttételesen alkalmazható.

Az indifferencia tarthatatlan elve miatt elvetett klasszikus interpretáció helyébe a 20. század platonista Cambridge-ében egy másik interpretáció lépett, a *logikai interpretáció*. A logikai interpretáció mögött a század elején a dedukció formalizálásában elért sikerek álltak. A deduktív következtetés mintájára a logikai interpretáció hívei egy *részleges* implikáció megadására törekedtek, amely két kijelentés, egy bizonyíték és egy hipotézis, ún. konfirmációs viszonyát fejezné ki. A valószínűség így a konfirmáció fogalmának korrelátuma lett, egy *induktív súly*, amelyet a konfirmációs viszony helyez a nyelv mondataira.

A *szubjektív interpretáció* a logikai interpretáció kritikájából indult ki, amennyiben vitatta, hogy létezne egy platóni értelemben objektív valószínűségi viszony a nyelv kijelentései között, amelyet minden racionális egyén azonos módon érzékel. A szubjektivisták ehelyett azt hangoztatták, hogy a valószínűség egy kijelentésbe vetett hit mértéke, amely egyénről egyénre változhat. Feltették, hogy e hit mértéke jól operacionalizálható alkalmas fogadási szituációkban. Másképp szólva egy kijelentésbe vetett hit mértéke akkora, amilyen mértékben az ágens hajlandó fogadni a kijelentés igazságára. A parciális hitek ilyetén operacionalizálása után a szubjektivisták feladata abban állt, hogy megmutassák, a fogadási kvóciensek formális értelemben valószínűségek.

A valószínűségi interpretációk sorát Karl Popper *propensity-interpretációja* zárta. A *propensity*-interpretáció a kvantumelmélet értelmezési vitái nyomán született. Popper meglátása szerint ugyanis amíg a klasszikus természettudományok valószínűségi kijelentéseinek értelmezéséhez a relatív gyakoriság-interpretáció tökéletesen megfelel, addig a kvantumesemények leírásához a valószínűség új interpretációja után kell néznünk. Ezért a valószínűség fogalmát Popper nem egy ismétlődő eseménysorozat relatív gyakoriságával azonosította, hanem a sorozatot generáló kísérleti elrendezés sajátosságának tekintette: olyan *kauzális hajlamnak*, amely *hosszútávú stabil frekvenciákat* képes létrehozni. Ennek a hosszútávú *propensity*-interpretációnak lett azután radikális továbbfejlesztése az ún. szinguláris *propensity*-interpretáció, amely a kauzális hajlamot, már nem hosszútávú frenvenciák, hanem *egyedi kimenetek* létrehozására vonatkoztatta.

Könyvünkben a fenti öt interpretációt igyekszünk részletesen bemutatni és értékelni. A könyv szerkezete a következő. Az első fejezetben a valószínűség fogalmának kialakulását és legfontosabb történeti paradoxonait mutatjuk be. A második fejezet a filozófiai elemzés számára nélkülözhetetlen matematikai fogalmakat vezeti be. Az olvasást megkönnyítendő a bonyolultabb fogalmakat és tételeket a könyv végén található függelékekbe utaltuk. A harmadik fejezet a valószínűség interpretációjának általános kérdéseivel foglalkozik, megrajzolva mintegy azt a fogalmi térképet, amelyen az egyes interpretációk elhelyezkednek. A könyv gerincét a negyediktől a nyolcadikig terjedő fejezetek képezik, amelyek részletesen áttekintik és elemzik a valószínűség öt standard interpretációját. A könyv rövid összegzéssel zárul.

A könyv megírásához nyújtott segítségükért sokaknak tartozom köszönettel. Elsősorban E. Szabó Lászlónak és Rédei Miklósnak szeretném hálámat kifejezni, akik érdeklődésemet a tudományfilozófia számos izgalmas fejezete, köztük a valószínűség interpretációjának kérdése felé kalauzolták. E. Szabó László *A nyitott jövő problémája* című könyvének a valószínűséggel foglalkozó fejezete, valamint Rédei Miklósnak a témában írott publikálatlan jegyzetei sokat segítettek az indulásnál. E. Szabó Lászlónak külön köszönettel tartozom azért a rengeteg beszélgetésért, amelyek során a tárggyal kapcsolatos nehézségeimet megoszthattam.

Könyvem elsősorban egyetemi jegyzetnek készült ahhoz az előadásomhoz, amelyet Bolyai-ösztöndíjasként *A valószínűség metafizikája* címmel tartottam az ELTE BTK Logika Tanszékén. Szeretném ezúton köszönetemet kifejezni az MTA Bolyai János Kutatási Ösztöndíjának munkám támogatásáért, valamint azoknak a hallgatóknak, akik az előadáson feltett kérdéseikkel és a jegyzetekhez fűzött megjegyzéseikkel hozzásegítettek, hogy számos kérdést tisztábban lássak.

A készülő könyvet vagy egyes részeit sokan olvasták. Itt szeretném megragadni az alkalmat, hogy köszönetet mondjak Gyenis Balázsnak, Gyenis Zalánnak és Szabó Máténak, akik hasznos észrevételeikkel elsősorban a matematikai részek pontosabb megfogalmazásást segítették elő, valamint Schmal Dánielnek a bevezető, történeti részekre vonatkozó tanácsaiért. Gömöri Márton, a teljes kézirat első olvasójaként, Máté András és Szegedi Péter pedig habilitációs opponensként gondos figyelemmel olvasták végig a szöveget, rengeteg hasznos javaslattal szolgálva: köszönet mindhármuknak. Köszönet illeti továbbá Barcza Istvánt, B. Gáspár Juditot, Fehér Mártát, Huoranszki Ferencet, Márton Miklóst, Tőzsér Jánost, akik megjegyzéseikkel gazdagították a könyvet vagy figyelmükkel támogatták létrejöttét. A könyvet anyám emlékének ajánlom.

1. fejezet

A valószínűség fogalmának kialakulása

1.1. Probabilitas

A valószínűség fogalma viszonylag későn jelent meg az európai gondolkodás horizontján. Ez annál meglepőbb, minthogy a szerencsejátékok korai létéről már az egyiptomi sírfeliratokon talált *astralagus*-ábrázolások tanúskodnak. Ez az *astralagus* vagy más néven *talus* egy bárány sarokcsontjából faragott négyoldalú játszócsontoska. Ian Hacking, a valószínűség történetének elsőrangú kutatója arról számol be (Hacking 1975), hogy a Kairói Múzeumban saját maga tesztelte ezeket a játszócsontocskákat, amelyek igen nagy fokú szimmetriát mutattak, és így kitűnő „véletlen-generátornak" számítottak. De hasonlóan kifinomult kombinatorikai gondolkodásról tesznek tanúbizonyságot a reneszánsz biztosítási és évjáradékszámításai is. Mindezek ellenére azonban a valószínűség elmélete csak későn alakul ki. Ennek elsősorban nem a használható aritmetikai háttér hiánya az oka – egészen egyszerűen a valószínűség *fogalma* hiányzott.

A valószínűség fogalmának kései kialakulását Hacking (1975) a következőképpen rekonstruálja. A valószínű (*probabilis*) szó értelmezéséhez a tudás és vélemény klasszikus majd középkori megkülönböztetéséből kell kiindulni: a tudás vagy tudomány (*scientia*) bizonyítható és így szükségszerű igazságokból áll, a vélemény (*opinio*) pedig egyedi reflexióból, érzékelésből vagy éppen vitából származik. A *probabilitas* e második körbe tartozott, valamely vélemény hihetőségét, igazolhatóságát jelentette; ott jelent meg, ahol szigorú bizonyítás nem volt lehetséges. Születési helyét nem a demonstrációra törekvő elméleti tudományokban, a mechanikában, az asztronómiában, az optikában kell keresni, hanem az „alacsony" tudományokban, az orvoslástanban, az alkímiában, a geológiában. A szavahihetőség valamely tekintélyre hivatkozást jelentette, és a reneszánsz fontos újítása éppen abban állt, hogy erre a tekintélyre a természetben lelt. Ahhoz azonban, hogy a természetet mint igazoló forrást használni lehessen, szükség volt még egy fogalomra, a *bizonyíték* fogalmára. A valószínűség kései megjelenése éppen a bizonyíték fogalmának kései megjelenésével magyarázható. Deduktív bizonyítás és tanúbizonyság természetesen létezett már, de nem létezett még dologi bizonyíték, vagyis az egyik létezőnek a másikra utalása, ami az induktív érvelés alapja. A bizonyíték fogalma az alkimisták *jel* fogalmából nőtt ki. Paracelsus és Agricola természeti jelekről beszélnek, olyan létezőkről, amelyek valami másra utalnak: a bolygók nevei fémekre, a szarvas agancsainak elágazásai a szarvas korára. A természeti jeleket a reneszánszban bizonyítékként kezdik értelmezni, vagyis olyan, a véleményt szavatoló tekintélyként, amilyen tekintélynek korábban csak a szent szövegek vagy néhány ókori szerző szövegei számítottak. És mivel egy vélemény valószínűsége éppen a bizonyítékkal való alátámasztottságot jelentette, a természeti jelekkel alátámasztott vélemények valószínűek lettek. Másfelől azonban a jó és megbízható jeleket éppen az különböztette meg a nem megbízhatóktól, hogy szabályos ismétlődésekben, regularitásokban jutottak kifejezésre. A bizonyítékként értett jelek gyakorisága növelte a tanúbizonyság erejét. Így kapcsolódott össze tehát a jel fogalmában az episztemológiai és aleatorikus jelleg, amely a valószínűség kettős, egyszerre szubjektív és objektív természetéért mindmáig felelős.

Az 1660-as évek előtti matematikusokat elsősorban kombinatorikai természetű problémák foglalkoztatták. Ezek a kombinatorikai problémák a jelek alkimista mágiájából nőttek ki, míg a 17. század meg nem szabadította a jeleket alkimista hátterüktől. Raymundus Lullus, akit a kombinatorika atyjának tekintenek, az Univerzum elemeinek jeleit kombinálva igyekszik leszármaztatni az Univerzum összetevőit. A három kockával való dobás kimeneteinek első ismert felsorolása egy olyan asztrológiai munkából származik, amely az égboltot az egyes kimeneteknek megfelelően 216 részre osztja, majd a dobások segítségével következtet az ég üzeneteire. Galileit is foglalkoztatják a három kocka kombinatorikai problémái: a *Sopra le scoperte dei dadi*ban (1623?) azt vizsgálja, hogy három kockával dobva mely összegekre érdemes fogadni. Az egyik korabeli elképzelés szerint a 9, 10, 11 és 12 összeg egyaránt megfelel, mivel mindegyik szám hatféleképpen áll elő mint három 1 és 6 közötti szám összege (pl. $9 = 1+2+6 = 1+3+5 = 1+4+4 = 2+2+5 = 2+3+4 = 3+3+3$). Ez azonban ellentmondani látszik a szerencsejátékosok tapasztalatának, amely szerint a 10 és 11 gyakrabban fordul elő a másik két számnál. Galilei a szerencsejátékosok álláspontját igazolta: az összegek gyakorisága attól függ, hogy a számok hányféleképpen fordulhatnak elő az összes, 216 eset között. A 10 és 11 huszonhét, a 9 és 12 azonban csak huszonöt különböző módon, így az előzők *könnyebben* valósulnak meg az utóbbiaknál. Ami pedig könnyebben valósul meg, az gyakoribb is. A „könnyű” és a „gyakori” Galileinél még szinonim fogalmak, viszonyuk nem szorul értelmezésre – a két fogalom metafizikai kapcsolatát majd Leibniz dolgozza ki. Az első átfogó kombinatorikai mű amúgy Girolamo Cardano *De ludo aleae*-ja, amely 1520 környékén született, de csak 1663-ban jelent meg egyfajta gyakorlati útmutatásként szerencsejátékosoknak.

1.2. Valószínűség és racionalitás

A valószínűségszámítás története két problémával indul, amelyeket – a legenda szerint – de Méré lovag, a kor hírhedt szerencsejátékosa ad fel Pascalnak Poitouba tartó háromnapos útjuk során. Az egyik az ún. osztozkodási paradoxon. Egy több fordulós játékban két játékos méri össze az erejét. A játékosok közül az nyeri a kitűzött jutalmat, aki a játékban 6 fordulót megnyer. A játék azonban 5 : 3-as játszmaállásnál félbeszakad. Kérdés: hogyan osszák el a játékosok maguk között a jutalmat, hogy a kifizetés a játszma állásának megfelelően igazságos legyen? Az osztozkodási paradoxont Luca Pacioli említi először 1494-ben, de megoldást nem ad rá. Tartaglia (1556) 2 : 1 arányban osztaná el a tétet, feltehetően mivel az egyik játékos két pont előnye a teljes hat fordulónak a harmada, így őt illeti meg a tét egyharmada, a maradékot pedig egyenletesen kell elosztani a játékosok között. A „helyes" megoldás azonban azon a belátáson alapul, hogy a probléma valószínűségi természetű: az elkövetkező három fordulóban a játék mindenképpen véget ér, és a nyolc lehetséges kombináció közül csak egy kedvez a hátrányban levő játékosnak – az amelyikben mind a három játékot ő nyeri. Így a tét igazságos felosztása 7 : 1 (amely szélsőségesebb az összes korábban javasolt elosztásnál).

A másik paradoxon az ún. két kockás paradoxon. Egy szimmetrikus kockával négyszer dobva $\frac{1}{2}$-nél nagyobb annak a valószínűsége, hogy legalább egy hatost kapunk; két kockával huszonnégyszer dobva azonban annak valószínűsége, hogy szintén legalább egy dupla hatost kapunk, kisebb, mint $\frac{1}{2}$. Hogyan lehetséges ez, hiszen a dupla hatos valószínűsége a hatoda a szimpla hatosénak, a huszonnégy pedig éppen a hatszorosa a négynek? Azzal, hogy a szimpla hatos valószínűsége négy dobás után $1-\left(\frac{5}{6}\right)^4 \approx 0,518$ valamint, hogy a dupla hatos valószínűsége huszonnégy dobás után $1-\left(\frac{35}{36}\right)^{24} \approx 0,492$ maga de Méré is tisztában volt. A paradoxont az jelentette, hogy a példában nem teljesül a *kritikus érték arányossági szabálya.* Kritikus értéknek azt a számot nevezzük, ahányszor egy kísérletet, amelyben egy esemény adott valószínűséggel fordul elő, meg kell ismételnünk ahhoz, hogy a sorozatban az esemény 50% valószínűséggel legalább egyszer előforduljon. A kritikus érték arányossági szabálya tehát az az intuitív elképzelés, hogy a kritikus érték fordítottan arányos az esemény valószínűségével: minél valószínűtlenebb egy esemény, annál többször kell elvégezni a kísérletet, hogy az esemény 50% valószínűséggel előforduljon. A *kritikus érték arányossági szabályának* közelítő jellegére csak Abraham De Moivre világít majd rá a *The Doctrine of Chances*-ben (1718). Tőle származik egyébiránt a „valószínűség" kifejezés első megfogalmazása is: „azon esély számának viszonylagos nagysága, hogy valami megtörténjen, azon esély teljes számához viszonyítva, hogy vagy megtörténik vagy sem, ami a valószínűség igazi mértéke".

Pascal tehát megoldja a kétkockás és az osztozkodási problémát, és a megoldásokat megírja Fermat-nak 1654-ben. Ennek a levélnek a dátumát szokás a valószínűségelmélet születési éveként számon tartani. Huygens Párizsba látogat, hogy találkozzon a két matematikussal, de Pascal addigra visszavonul a Port Royal-ba, Fermat pedig Toulouse-ban van. Huygens hazatérve megírja *De ratiociniis in aleae ludo* című művét 1657-ben, amely az első kézikönyv a témában. A könyvben jelenik meg először a várható érték fogalma mint egy játék megvásárlási ára. A várható érték fogalma a 17. században alapvetőbb a valószínűség fogalmánál. Ennek gyökerét azokban a jogi gyakorlatban alkalmazott ún. aleatorikus szerződésekben kell keresnünk, amely során az ügyfél egy jelenlegi biztos jót egy jövőbeli bizonytalanra cserél (biztosítási szerződések, évjáradékok, megvenni a halász soron következő hálójának tartalmát, stb.). Ezek az aleatorikus szerződések elsősorban az uzsora vádja alól tisztázták a kereskedőket.

A 17. század racionalizmusát az a törekvés jellemezte, hogy a racionálitás fogalmát kiterjesszék az eredetileg demonstratív tudás köréből a gyakorlati cselekvés birodalmára. Megfordult tehát az igazolási viszony a cselekvés és a hit között, a gyakorlati cselekvés sürgető morális kihívásként jelentkezett, amely racionális kritériumokat kívánt. A praktikus ész ezen kihívásának emblematikus példája Pascal *fogadása* Isten léte mellett. A *fogadás* csak Pascal halála után jelent meg 1670-ben a *Gondolatok* 233. pontja alatt, azonban már Pascal életében ismertté vált, minthogy a Port Royal *Logikája* főbb vonalakban összefoglalta. A *fogadás* – történetileg az első döntéselméleti feladat – három premisszán nyugszik, amelyek igazságát Pascal a gyakorlati racionalitásból eredezteti: (i) Isten létének p valószínűsége nem nulla. (ii) Ha Isten létezik, és mi rá fogadunk, jutalmunk végtelen lesz ($a = \infty$), míg a többi három esetben (Isten nem létezik, mi mégis rá fogadunk; Isten létezik, de mi nem fogadunk rá; Isten nem létezik, és mi nem fogadunk rá) ez a jutalom véges ($b, c, d < \infty$). (iii) A racionális döntés a várható jutalom maximalizálását kívánja. Mindezekből Pascal levezeti, hogy a racionális cselekvés az Istenre fogadás, mivel annak várható jutalma végtelen ($p \times a + (1-p) \times b = \infty$), míg az Isten ellen való fogadásé véges ($p \times c + (1-p) \times d < \infty$). Figyelemre méltó, hogy míg a skolasztikában az istenbizonyítékok demonstrációk, azaz a tudás körébe tartoznak, addig Pascal *fogadása* utat nyit a valószínű vélemény mentén megfogalmazott istenbizonyítékoknak, amely Isten létét és a holnapi napfelkeltét egyformán kezeli.

A valószínűség fogalmát a Port Royal logikusai használják először mérhető dologra az 1662-ben írt *La logique, ou l'art de penser*-ben. A műnek a valószínűséggel foglalkozó negyedik könyvét feltehetőleg Arnauld írta. A villámtól való félelem mértékét a villámcsapások relatív gyakoriságával veszi arányosnak. A csodák értékelésénél megkülönbözteti a belső körülményeket, az adott helyen történő eseményeket, és a külső körülményeket, a hagyományozódás mikéntjét. Mindez azért fontos, mert a valószínűség fogalma csak akkor bukkanhatott fel, amikor a jeleket belső evidenciaként kezdték érteni.

Leibniz tanult jogászként a jog felől közelít a valószínűség fogalmához. A *De conditionibus*-ban (1665) „jus purum”-nak nevez egy következtetést, ha az előtag implikálja az utótagot, és a következtetéshez az 1 számot rendeli. „Jus nullum”-nak nevez egy következtetést, ha az előtag az utótag ellenkezőjét implikálja, és ekkor a következtetéshez a 0 számot rendeli. Végül „jus contingens”-nek nevez egy következtetést, ha az előtag csak részben implikálja az utótagot. A jogi értelmezés hátteréről a következőket kell tudni (Daston, 1988, 2006). A Negyedik Lateráni Zsinat 1215-ben eltörölte az istenítélettel való bírósági döntést, ezzel a jogászok vállára rakva a döntés felelősségét az élet-halál kérdésekben. Ennek hatására a 13. századtól kezdve fokozatosan kialakul a bírósági bizonyítási eljárások egyfajta hierarchiája – a szóbeszédre vagy ellentmondó tanúkra épülő alacsony fokú bizonyságtól a két megbízható és független tanúra építő, „az ésszerű kételyt” kizáró bizonyosságig. A skála a 17. századra numerikus jegyeket ölt, és Leibniz csak tovább finomítja a jogi gyakorlatban már régóta alkalmazott elemeket. Leibniz tehát ismeretelméleti oldalról közelít a valószínűség fogalma felé, és 1672-ben Párizsba érkezve meglepődve tapasztalja, hogy Pascalék a valószínűséget játékokra alkalmazzák. Leibniz nevéhez fűződik a feltételes valószínűség fogalma, és a valószínűség redukciója kedvező és ellentmondó esetekre.

Az 1603-as londoni pestis kikényszeríti, hogy a születési és halálozási jegyzékeket heti osztásban vezessék. John Graunt (1662) és William Petty (1666) ezekre a statisztikai adatokra támaszkodva egy sereg jelentős következtetésre jut. 1. A születések számából a nők számára, abból a családok számára, abból pedig London lakosságának nagyságára következtetnek, és kimutatják, hogy az a korabeli vélekedésekkel ellentétben jóval kevesebb, mint kétmillió. 2. A koldusok magas számára hivatkozva szorgalmazzák egy nemzeti segélyalap felállítását. 3. Mivel a statisztikák nem adják meg az elhalálozott korát, csak a halál okát, ezért megkonstruálnak egy elhalálozási függvényt, amely majd az életjáradék-számításban kulcsszerepet játszik. 4. A régi eredetű és egészen a 19. századig tartó vitában, hogy ti. a pestis fertőzéssel vagy miazmákkal terjed-e, Graunt az utóbbi mellett foglal állást, mivel a heti áldozatok számának nagy oszcillációja csak az időjárás változásával látszik magyarázhatónak.

Az évjáradékok, életjáradékok és biztosítások gyakorlata Hollandiában ez időre széles körben elterjed. A járadékok pontos kiszámításához a halálozási görbét és az aktuális kamatot kellett ismerni. Graunt a halálozási görbét az alábbi félig empirikus, félig *a priori* megfontolásokkal konstruálja meg. A halálozási jegyzékek, mint említettük, nem adták meg az elhunyt korát, de megadták a halál okát. Graunt a halálnemek csoportosításával megállapította, hogy az elhunytak közel egyharmada gyermekbetegségben halt meg, vagyis hat évnél fiatalabb. Az elhunytak maradék kétharmadát egyenletesen elosztotta a hat és hetvenhat év között. Hasonlóan uniform halálozási görbét használtak a holland de John de Witt és John Hudde, később pedig Leibniz és de Moivre is, és azt egyfelől táblázatokkal, másfelől *a priori* feltevésekkel próbálták igazolni. Ezek az *a priori* meggondolások vezetnek aztán a valószínűségnek az „egyenlően lehetséges esetek” segítségével történő megfogalmazásához.

Az első teljes halálozási jegyzékeket a csillagász Edmund Halley készíti el 1693-ban, és segítségével megmutatja, hogy a kormány által kibocsátott évjáradékok – 6% visszatérülés hét éves futamidő alatt – alulárazottak. Az egyre kifinomultabb számítások ellenére a biztosító társaságok nem mutattak különösebb érdeklődést a valószínűségszámítás új eredményei iránt. Az elméletnek a szerencsejátékok köréből merítő példái túlságosan frivolnak tűntek; azt a kialakult gyakorlatot pedig, hogy a szerződéskötésnél az ügyfél korát, egészségi állapotát és foglalkozását esetileg mérlegelik, nem szívesen bízták vak átlagokra. A biztosítási üzlet ekkori igen jövedelmező szakaszában az ilyen matematikai megfontolások amúgy sem számítottak túlságosan. Az egyetlen kivételt (az azóta is működő) *Equitable Society for the*

Assurance of Lives (1762) jelentette, amely biztosításait rögzített statisztikus törvényekre építette, de amelynek sikere ellenére évtizedekig nem akadt követője.

1.3. Ars conjectandi

Jacob Bernoulli *Ars conjectandi*-ja (1713) a Port Royal *Ars cogitandi*-jára rímel, mintegy meghosszabbítva annak gondolatmenetét. Bernoulli a valószínűségi kalkulust a Port Royal logikájában felvetett kérdésekre alkalmazza: mi a valószínűsége annak, hogy egy bizonyos kórokozó lakik az ember testében, ha adva van a kórokozók okozta fertőzések gyakoriságának orvosi jegyzőkönyve. A problémát urnás modellje alapján tárgyalja, amelyben a különböző színű golyók aránya az urnában a kórokozók valószínűségének felel meg; a fertőzések statisztikája pedig az urnából való ilyen-olyan színű golyó húzási valószínűségének. Bernoulli tehát az okozatok valószínűségéből következtet az okok valószínűségére. Ezt a kutatási irányt a korban *a posteriori* módszernek nevezték. Ennek a módszernek lett centrális jelentőségű igazolása a műben bizonyított nagy számok (gyenge) törvénye, amelyben Bernoulli urnás modelljére támaszkodva belátta, hogy az *a posteriori* valószínűségek (a kihúzott golyók színeinek relatív gyakorisága) az *a priori* valószínűségekhez (az urnában levő színek arányaihoz) tartanak a húzások szaporodtával. A nagy számok törvénye azonban csak azon az áron volt képes betölteni az empirista program igazolásának feladatát, hogy egyfelől a bizonyítéknak korábban igen differenciált fogalmát az esetek puszta számlálására redukálta, másfelől az okok (az urna színarányai) és az okozatok (az egyedi húzások színe) közötti *valószínűségi oksági* viszonyt – amely sehogyan sem illett a kor mechanisztikus oksági elképzelésébe – pusztán látszatnak minősítette. Ez azonban oda vezetett, hogy a valószínűség egy determinisztikus világban pusztán ignoranciát jelentett. A valószínűségnek ez a felfogása lett uralkodó azután Bernoullitól egészen Laplace-ig. Bernoulli egyébként az utolsó az újkorban, aki nem-additív mértékkel kísérletezik.

Az *Ars conjectandi* átfogó elemzésére építve határozza meg a század végén Condorcet az esküdtszék optimális méretét és az igazságos szavazati procedúrát. Az 1785-ös *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* című nagyszabású munkájában Condorcet kimutatja, hogy az esküdtszéki szavazati eljárások könnyen úgynevezett ciklikus kollektív preferenciákhoz vezethetnek, vagyis a többségileg előnyben részesített döntések nem feltétlenül rendezhetők sorba. Ennek a szavazati paradoxonnak egy valószínűségi variánsa az alábbi paradoxon is, amely a véletlen események rendezésével kapcsolatos nehézségekre hívja fel a figyelmet. Két játékos előtt három szabályos, de számozatlan kocka hever az asztalon. A játék menete a következő lépésekből áll: (i) Az egyik játékos fogja a kockákat, és tetszés szerint 1-től 18-ig megszámozza az oldalaikat. (ii) A másik játékos megtekinti a számozást, majd kiválaszt egyet a három kocka közül. (iii) Az első játékos a maradék két kocka közül kiválaszt egyet; a harmadik kockát félre teszik. Ezek után mindenki a saját kockájával dob – a kérdés az, hogy melyik játékos van előnyben. Vagy másképp fogalmazva a kérdés az, hogy az első játékos meg tudja-e számozni a kockákat úgy, hogy ellenfele bárhogy válasszon is egy kockát a háromból, ő a maradék kettőből tudjon „jobb” kockát választani, vagyis olyat, amellyel $\frac{1}{2}$-nél nagyobb valószínűséggel nagyobb számot fog dobni ellenfelénél. Amint a paradoxon erre rámutat, ilyen számozás lehetséges. (Pl. első kocka: $1, 6, 11, 12, 13, 14$; második kocka: $2, 3, 4, 15, 16, 17$; harmadik kocka: $5, 7, 8, 9, 10, 18$.) A kockák úgymond körbeverik egymást: annak valószínűsége, hogy a második kockával nagyobbat dobunk, mint az elsővel $\frac{21}{36}$, és ugyanennyi ez a valószínűség a harmadik és a második, illetőleg az első és harmadik kocka viszonylatában is. A Condorcet-paradoxon tehát arra figyelmeztet, hogy véletlen események nem rendezhetők jól az „$\frac{1}{2}$-nél nagyobb valószínűséggel jobbat kapni” összehasonlítás mentén.

A valószínűségi kalkulus mind kiterjedtebb alkalmazása polgári és morális kérdésekre a tizennyolcadik századra ahhoz az általános meggyőződéshez vezetett, hogy a valószínűségi kalkulus magának a racionális gondolkodásnak közvetlen leírása. Ezzel a meggyőződéssel szemben jelentett kihívást a *pétervári paradoxon*. A pétervári paradoxon alapgondolata Nicolaus Bernoullitól (Jacob Bernoulli unokaöccsétől) származik, írásban pedig először Daniel Bernoulli (Nicolaus unokatestvére) tette közzé a porosz akadémia *Commentarii Academiae Scientiarum Imperialis Petropolitanae* című folyóiratában 1738-ban, ahonnan a paradoxon a nevét is kapta. Egy bankkal folytatott játékban a nevezési díj ellenében az alábbi játékban vehetünk részt. A bankos egy érmével dob addig, amíg fejet nem kap. Ha már az első dobás fej, akkor kapunk 2 aranyat, ha csak a második, akkor 4-et, és így tovább, ha csak az n-edik dobás lesz fej, akkor 2^n aranyat. Kérdés, mennyi legyen a játék nevezési díja. A válaszhoz ki kell számolni a játékos átlagos nyereségét, vagyis az f kifizetésfüggvény várható értékét. Ez az érték azonban $\sum_n \frac{1}{2^n} 2^n = \infty$, vagyis a játék *fair* nevezési díja végtelen lenne. Mégis „kevesen fizetnének akár csak 25\$-t is, hogy részt vehessenek a játékban” (Hacking, 1975). Az elmélet jóslata és a józan ész közötti ellentétet Bernoulli úgy vélte feloldhatónak, ha a feladatban nem a kifizetések várható értékét tekintjük, hanem a hasznosságét. A várható hasznosság fogalmát tehát Bernoulli vezeti be először, a hasznosságot pedig a kifizetés logaritmusaként definiálja, első példáját adva ezzel a közgazdaságban bevett csökkenő marginális hasznosságnak. Ha a kifizetést az $u = \log f$ hasznossággal helyettesítjük, akkor a játék nevezési díja $0,602$ hasznosságú, ami visszaszámolva 4 arany. A várható hasznosság fogalmának bevezetése arról tanúskodott, hogy a valószínűségi kalkulus elég rugalmas ahhoz, hogy a „feltevés művészetének” megfelelő eszköze legyen. (Bernoulli ötlete azonban nem oldja meg a problémát, hiszen a paradoxon megismételhető a hasznosságra és nem a kifizetésekre vonatkoztatva: fizessen a bank az n-edik fej dobásra 2^n *hasznosságot*! Megoldást csak a kifizetések korlátozása jelenthet. A bank kifizetését pl. egymillió aranyban maximálva, a várható érték kb. 21 arany lesz.)

Thomas Bayes írása az *An essay towards solving a problem in the doctrine of chance* a *Philosophical Transactions* című folyóirat 1763-as kötetében látott napvilágot Bayes klerikustársa, a szintén matematikus Richard Price gondozásában Bayes halála után két évvel. Bayes előtt a valószínűséggel foglalkozó munkák többsége azzal a kérdéssel volt elfoglalva, amit ma *direkt következtetésnek* nevezünk, vagyis, ha adva van egy szerencsejáték kimeneteinek valószínűsége, akkor hány ismétlés szükséges ahhoz, hogy egy bizonyos kimenet valószínűsége elérjen egy kívánt értéket: például hányszor kell dobni egy pár szimmetrikus kockával, hogy legalább 50% valószínűséggel legyen benne dupla hatos. Bayes megfordította a problémát: számoljuk ki a véletlen esemény valószínűségét, miután megfigyeltük az esemény egy ismétlődő sorozatának kimeneteit. Pontosabban: „*Adva van,* hogy egy ismeretlen esemény hányszor történt és hányszor hiúsult meg: *meghatározandó* annak esélye (chance), hogy az esemény valószínűsége két meghatározott valószínűségi érték közé esik." Úgy tűnik, hogy Bayes erre az *inverz következtetésre* is a direkt következtetéshez hasonló definit számokban megadható választ várt.

Ami a tágabb kontextust illeti, Bayes munkája minden bizonnyal Hume-nak az indukcióval szemben támasztott kételyeire kívánt válaszolni. Bayes esszéjének megjelenése után három évvel Price megírja *Four Dissertations* (1767) című munkáját, ahol a negyedik disszertációban erősen támaszkodik Bayes számításaira. Olvasva a művet, Hume így válaszol Price-nak: „El kell ismernem, hogy az a világosság, amelybe a vitát helyezted, újszerű, kézenfekvő, szellemes, és lehet, hogy alapvető is. Mindenesetre több időre van szükségem, hogy mindezt mérlegeljem, mielőtt a magam számára is megnyugtató ítéletet hozok." Hume azonban nem tért vissza többet a kérdéshez. Hume hallgatása azonban nem meglepő, mivel a 19. század közepéig – Bayes munkáját leszámítva – egyáltalán nem kapcsolódott össze az indukció és a valószínűség fogalma. Az indukció Bacontől Millig a bizonyosra törekedett, nem a valószínűre.

Az inverz következtetés módszere (amely elnevezés De Morgantól ered) azonban mégsem Bayes munkája révén vált közismertté, hanem Laplace munkássága nyomán. A kérdés Laplace-nál csillagászati kontextusban jelent meg, mivel a biztosítási ügyletek mellett a csillagászat volt a valószínűségszámítás empirikus alkalmazásának másik legfőbb területe: a tudósokat az asztronómiai megfigyelési hibák kezelése foglalkoztatta. A hibák értékelésénél Kepler az átlagra, Galilei – arra a jogi érvelésre támaszkodva, hogy a tanúbizonyság ereje a tanúk számával nő – a moduszra, a leggyakrabban előforduló elemre szavazott. Az igazi előrelépést azonban Thomas Simpsonnak az a felfedezése jelentette, hogy maguk a hibák is valószínűségi eloszlást követnek. Erre a meglátásra és Bernoulli urnás modelljére építi aztán Laplace hibaszámítási vizsgálatait, amelyek az inverz következtetés módszerének – mai terminológiával a becslésnek – népszerűsítéséhez vezetnek. A módszer végül Gauss munkája nyomán tesz szert igazi jelentőségre, aki, miután felfedezi a normális eloszlást, görbéket illeszt előzetesen adott pontokhoz, és megmutatja, hogy a legvalószínűbb görbe a legkisebb négyzetek módszerével származtatható. Laplace rögtön felismeri a módszer jelentőségét, és Gauss eredményeit centrális határeloszlás-tételének segítségével – amely tételt egyébiránt De Moivre fedezte fel először mintegy száz évvel Laplace előtt – általánosítja minden olyan szituációra, ahol nagyszámú, független ok működik. Végül a *Théorie analytique des probabilités*-ben (1812) publikálja az okok valószínűségének és a hibák analízisének teljes elméletét. Laplace nagyszabású művének hatása mind a matematikai módszer, mind a valószínűségnek a hit mértékével való összekapcsolása tekintetében felbecsülhetetlen.

Laplace Bernoulli urnás modelljének segítségével vezeti le a *rákövetkezés szabályát* (az elnevezés John Venntől (1866) származik), vagyis azt a szabályt, hogy ha egy végtelen nagy urnából az eddig kihúzott n golyó mind fekete volt, és feltételezve azt, hogy az urnában a fekete golyók összes lehetséges aránya egyformán valószínű, akkor annak a valószínűsége, hogy a következő golyó fekete lesz $\frac{n+1}{n+2}$ (lásd D. Függelék). Laplace a szabályt a felkelő Nap példájára alkalmazza: ha a Nap az 5000 éves történelem során eddig minden reggel felkelt, akkor annak valószínűsége, hogy holnap is fel fog kelni $0,9999994$. A példaválasztás nem véletlen. Hume számára a felkelő Nap a nem szükségszerű, mégis bizonyosan tudható tudás példája volt; Price – Hume empirista filozófiájának opponense – ellenben annak a valószínűségét, hogy a Nap holnap is felkel, a napfelkelték 5000 éves sorozatának valószínűségével azonosította, miután feltette, hogy a napfelkelték egymástól függetlenek és azonos valószínűségűek. Laplace megoldása mindkét megoldásnál naturalistábbnak számított, mivel a tapasztalatból való induktív tanulás lehetőségét modellezte, vagyis azt, hogy a vélekedés a tapasztalatok szaporodtával aszimptotikusan közelít a bizonyossághoz. Laplace analízise az enumeratív (vagyis egyedi esetről egyedi esetre következtető) indukció első valószínűségi tárgyalása.

A valószínűség klasszikus interpretációja szerint a valószínűség a kedvező esetek és az *egyenlően lehetséges* esetek aránya. A megfogalmazás Laplace *Essai philosophique sur les probabilités* (1814) című monumentális munkájából ered, de a gondolat mintegy száz évvel korábban felbukkan Leibniz igen kifinomult metafizikájában. Kezdjük a lehetőség fogalmával. Leibniz a lehetőség fogalmát a mai logika előhírnökeként konzisztencia értelemben használja: lehetséges az, ami ellentmondásmentes. Ez az ellentmondásmentesség azonban fokozatokban jön, így a lehetőségek konzisztenciájuk alapján sorba rendezhetők. A lehetőségek azonban nem pusztán formálisan léteznek – és itt torkollik a gondolatmenet Leibniz bonyolult metafizikájába –, hanem aktuális létezésre törekszenek, úgymond viaskodnak egymással a létükért: „A lehetséges saját természeténél fogva létezést követel lehetőségi fokával vagyis lényegi fokával arányosan." Vagyis a lehetőségek egyfajta *propensity*-vel, hajlammal rendelkeznek a létezésre, éppen lehetőségi fokukkal arányban. Ez a metafizikai elképzelés helyet kap azután Leibniz teremtéskoncepciójában is, amely szerint a lehetséges, azaz konzisztens világok Isten elméjében várják, hogy közülük a legkonzisztensebbet, azaz a lehetséges világok legjobbikát Isten megteremtse. A lehetőség foka az elmében tehát tükörképe annak, hogy az esemény milyen könnyen valósul meg a világban. A lehetőség foka azonban nem más, mint maga a valószínűség. Vagyis: „*Quod facile est in re, id probabile est in mente.*" Amilyen könnyen megvalósul a hatos dobás a világban, olyan valószínű az az elme számára. De honnan tudjuk, hogy milyen könnyen valósul meg a hatos dobás, vagy másképp szólva, a hatos dobás lehetőségének milyen hajlama van a létezésre? Egyszerűen onnan, hogy milyen gyakran kapunk hatost a kockadobások során. És éppen ebben áll a valószínűség fentebb már említett kettős természete: a valószínűség egyfelől aleatorikus jellemző, amennyiben az ismétlődő regularitásokkal áll kapcsolatban, másfelől episztemikus, amennyiben várakozásainkat tükrözi. Összegezve, a lehetőség éppen azért válhatott a valószínűség fogalmi alapjává, mert a valószínűséghez hasonlóan kétarcú fogalom volt.

Az „egyenlően lehetséges" kifejezés hasonlóan kettős jelentéssel bír: jelenti egyfelől azt, hogy a lehetőségek egyenlően könnyen következhetnek be, és jelentheti azt is, hogy nincs okunk az egyik lehetőséget előnyben részesíteni a másikkal szemben. Mindkét értelmezést az ún. *elégtelen ok elve* igyekezett metafizikailag alátámasztani, amely elv Leibniz *elégséges ok elvének* egyfajta kontrapozíciója. Az *elégtelen ok elve* tehát kezdetben a leibnizi metafizika talaján állva biztosította az átjárhatóságot az aleatorikus és episztemikus valószínűségértelmezés között. A kapcsolat a kétféle értelmezés között a korban általánosan elfogadott asszocionalista pszichológia miatt egészen Laplace koráig fennmaradt. Ez az asszocionizmus úgy tartotta, hogy a bennünk kialakuló benyomások tisztasága és bizonyossága a megfigyelések gyakoriságával és stabilitásával függ össze, így a valószínűség két értelmezése egyazon érmének két oldala. Miután azonban az asszocionalista pszichológia a 19. században átadta a helyét a torzítások és az idioszinkráziák iránt érdeklődő észleléselméleteknek, a valószínűség objektív és szubjektív értelmezései is széttartó utakra kerültek. Az *elégtelen ok elve* a 19. század végére végleg elvesztette objektív értelmezését, és pusztán az ignoranciára, a tudás hiányára építő episztemikus elvvé vált. Ezért történhetett, hogy az elvet végül Keynes (1921) átkeresztelte *az indifferencia elvének*.

A francia forradalom megrendítette a közös racionális alapokba vetett hitet. A nagy racionalitási elveknek, köztük az *elégtelen ok elvének* ésszerűségét egyre többen kétségbe vonták. Ráadásul az elv alkalmazása egyre több hibát eredményezett: az egyenlően lehetséges esetek számbavétele már a legegyszerűbb kombinatorikai esetekben sem volt egyértelmű. Így fordulhatott elő, hogy d'Alembert az *Enciklopédiában* a két érmével való dobásnál a fej-fej, az írás-írás és a fej-írás eseteket egyenlően lehetségesnek tartotta, és így egyenlően valószínűnek is. Az *elégtelen ok elve*, amelyre Laplace elméletét, többek között a *rákövetkezés szabályát* is alapozta, a reális okok világában egyáltalán nem nyújtott fogódzót. John Stuart Mill a teljes klasszikus interpretációt hamisnak tartotta – a tudás nem épülhet nem tudásra, hanem egyedül a tapasztalatra: „Ahhoz, hogy azt állíthassuk, hogy két esemény egyenlően valószínű, nem elegendő, hogy tudjuk, hogy az egyik vagy a másik megtörténik, és nem tudjuk, hogy melyik. A tapasztalatnak kell feltárnia, hogy a két esemény egyenlő gyakorisággal történik" (1843). Ezt a szkepszist az inverz következtetésnek még az olyan támogatói is osztották, mint George Boole.

1.4. A „nyomtatott számok lavinája"

A klasszikus felfogás hanyatlása magával hozta az igényt a valószínűség objektívebb interpretációja iránt. A 19. század elejétől Európát a „nyomtatott számok lavinája" (Hacking, 1975) borította el: a sporadikus egyházközségi feljegyzések átadták a helyüket a születésekről, halálozásokról, bűnesetekről és fertőzésekről készített hivatalos és kiterjedt statisztikáknak. A statisztikákból leszűrt és a nagy számok törvényei által alátámasztott stabil statisztikus viselkedés feleslegessé tette a háttérben működő különféle okok azonosítását, így a figyelem inkább a regularitásokra tevődött át. A belga csillagászt, Adolphe Quetelet-t a skót ezred katonái között végzett mellbőségmérés adatainak illeszkedése a Gauss-görbére annyira meggyőzi a normál eloszlás általános érvényességéről, hogy amikor a francia sorkatonák magasságában a görbétől eltérést észlel, akkor a sorozást elkerülni igyekvő hamisításra gyanakszik. Hasonló csodaszámba menő illeszkedést talál Bortkiewicz a Poisson által a *Recherches sur la probabilité des jugements*-ban (1837) publikált Poisson-eloszlás és a porosz lovasságban lórugásban meghalt katonák száma között. Mindezek a fejlemények az 1820-30-as évekre ahhoz az általános meggyőződéshez vezettek, hogy a valószínűség csak eseménytípusra vonatkoztatható mennyiség, és a helyes interpretáció a frekventista interpretáció.

A statisztika, amelyet a 19. században elsősorban a szociális viselkedés feletti kontroll, mintegy a „valószínűség megszelídítése” (Hacking, 1975) motivált a társadalomtudományokban, a század végére a fizikai és biológiai tudományokban is jelentős szerephez jut. Maxwell (1860) Quetelet normál szabályait emberek helyett gázmolekulákra alkalmazza, és kinetikus megfontolások alapján levezeti az ideális gáz molekuláinak sebességeloszlását, Boltzmann (1872) pedig ún. H-tétele segítségével a termodinamika második főtételét véli a mechanikára visszavezetni. Boltzmann programjának ellentmondásaira Loschmidt (1876) megfordíthatósági ellenvetése valamint Zermelo (1896) visszatérési ellenvetése világít rá. Mindkét ellenvetés a második főtétel egyirányú entrópiaváltozása és a reverzíbilis, azaz időtükrözésre szimmetrikus dinamikai folyamatok közötti ellentmondásra épít. A kritikákra mind Maxwell, mind Boltzmann elméletüknek valószínűségi átértelmezésével reagálnak. Boltzmann (1877) bevezeti a Boltzmann-entrópia fogalmát, és kombinatorikai érvet konstruál az entrópianövekedés alátámasztásra; Maxwell (1867) pedig, aki mindvégig szkeptikus a második főtétel mechanikai megalapozhatóságában, a főtétel statisztikus jellegét az ún. Maxwell-démon gondolatkísérlettel illusztrálja. A gondolatkísérletben egy képzeletbeli démon egy gáztartályt középen kettéválasztó falat olyan ügyesen nyitogat illetve csukogat, hogy a gyors gázrészecskék végül a tartály egyik felében, a lassú részecskék pedig a tartály másik felében gyűlnek össze – ellentmondásban az entrópianövekedés elvével. Az ellentmondás csak úgy feloldható, ha a második főtételt pusztán statisztikusan érvényes elvnek tekintjük.

De nemcsak a fizikában hoz áttörést a statisztikus gondolkodás. A biológiában Darwin unokaöccse, Francis Galton megalapítja a biometria tudományát, bevezeti a korreláció fogalmát, és felfedezi az átlag felé történő regressziót. Pearson a ferde eloszlások tanulmányozására pedig bevezeti a χ^2-tesztet.

A valószínűségnek azonban mind a szubjektív, mind az objektív interpretációja forgalomban marad egészen a század végéig. Jól mutatja ezt a kettős helyzetet az *Encyclopaedia Britannica* 1885-ös, kilencedik kiadásának szócikke, amely szerint „egy tényre vagy állításra vonatkozó valószínűség, vagy meggyőződés mértéke lényegét tekintve szubjektív, és azon elme tudása mértéke szerint változik, amelyik számára a tény megjelenik”. Majd pedig elmagyarázza, hogy „a valószínűség tudása minden esetben abból az elvből származik, hogy az arány, amelyet nagy számú próba során találunk, fenn fog állni a teljes számra is, legyen az akár végtelen.”

A valószínűség fogalmának matematikai tárgyalását a 19. és 20. század fordulóján sokan sürgették. Hilbert híres párizsi előadásában, amelyben a század matematikája előtt álló legfontosabb 23 nyitott problémát vette számba, hatodikként a modern matematikai fizikában kulcsszerepet játszó valószínűség axiomatizálását említi. A program keresztülviteléhez azonban várni kellett a mértékelmélet kialakulásáig, amely elmélet Borel, Lebesgue és Frechét nevéhez fűződik (ld. Doob 1996). A mértékelmélet a hossz-, terület- illetve térfogatfogalom általánosítása absztrakt halmazokra. Az általánosítás első lépéseként Borel (1989) a hosszúság fogalmát kiterjesztette a számegyenes intervallumairól olyan halmazokra is, amelyek az intervallumok megszámlálható halmazelméleti kombinációival kaphatók. Ezek a Borel-halmazok, a hozzájuk rendelt mérték a Borel-mérték. A Borel-mérték azonban nem teljes: a nullmértékű halmazok nem minden részhalmaza mérhető. A bajt orvoslandó Lebesgue (1904) bevezette a ma Lebesgue-mértéknek nevezett fogalmat, amely a hozzá kapcsolódó integrálfogalommal együtt a függvényanalízis központi fogalma lett. Ezt követően Frechét (1915) megmutatta, hogy a Lebesgue eredményeit meg lehet fogalmazni az $\mathbb{R}^n$ felhasználása nélkül is; más szóval, hogy a mértékelmélethez pusztán egy σ-algebra kell valamint egy σ-additív függvény.

A mértékelmélet kialakulása megérlelte a feltételeket a valószínűségfogalom egységes és elegáns tárgyalásához. Ezt a feladatot végül Kolmogorov végezte el a *Grundbegriffe der Wahrscheinlichkeitsrechnung* című munkájában 1933-ban, amely mű mára a modern valószínűségszámítás alapművének számít. A mindössze 130 oldalas kitűnő didaktikával megírt könyvecske a mértékelméleti fogalmak elegáns valószínűségi értelmezése, valamint a feltételes valószínűségnél és a véletlen folyamatok elméletében alkalmazott mértékelméleti megközelítés révén hamar magára vonta a szakma figyelmét, és – talán egy cseppet méltánytalanul – háttérbe szorította mindazokat az előfutárokat és alternatív útkeresőket, akik hozzájárultak a valószínűségelmélet kialakulásához. A következő fejezetben ennek a kolmogorovi elméletnek a matematikai alapjait ismertetjük.

2. fejezet

A valószínűségelmélet alapfogalmai

A valószínűség *mint matematikai fogalom* egy mérhető téren értelmezett σ-additív, normált mérték. Ezt a matematikai fogalmat azonban gondosan meg kell különböztetnünk attól az „igazi”, fizikai valószínűségtől, amelynek természetét ebben a könyvben feltárni igyekszünk. Ezért a leghelyesebb az volna, ha a matematikai „valószínűség” kifejezést egyáltalán nem használnánk, hanem helyette egyszerűen „normált mértéket” mondanánk. Ez a döntés azonban igencsak megnehezítené a beszédet, különösen ebben a fejezetben, ahol a valószínűségszámítás alapfogalmait szeretnénk röviden ismertetni. Ezért mi azt a megoldást választottuk, hogy a valószínűség fogalmát abban az interpretálatlan formában, ahogyan azt a matematikusok értik, *kizárólag* könyvünknek ebben a fejezetében valamint a Függelékekben használjuk. Az összes többi fejezetben „valószínűség” alatt fizikai valószínűséget értünk, ha pedig a matematikai valószínűségre akarunk utalni, akkor azt egy külön jelzővel jelöljük.

A valószínűség interpretációja szempontjából matematikailag elég volna csupán két alapfogalmat, az eseményteret és a valószínűségi mértéket bevezetni. Az interpretációk többsége azonban gazdagabb alapokra épít, vagyis gyakran olyan származtatott fogalmakra és tételekre is hivatkozik, mint a függetlenség fogalma vagy a nagy számok törvényei. Nem takaríthatjuk meg tehát a fáradságot, hogy a valószínűségszámítást legalább elemi szinten ismertessük. Ebben a fejezetben tehát röviden áttekintjük a valószínűségszámítás legfontosabb alapfogalmait és tételeit, úgymint a σ-algebrát, a mértéket, a valószínűségi változót vagy a Bayes-tételt. Hogy az ismertetőt ne terheljük túl, ezért az olyan bonyolultabb fogalmakat és tételeket, mint a feltételes függetlenség vagy a de Finetti-féle reprezentációs tétel a Függelékekben tárgyaljuk. (A valószínűségszámítás bővebb tanulmányozásához remek kézikönyvek például (Chow, Teicher 1988) valamint (Billingsley 1995) vagy a bevezető jellegű (Ross 2009).) Mivel a valószínűség interpretációi, amint majd látni fogjuk, nemcsak az *esemény* metafizikai fogalmára épülhetnek, hanem a *kijelentés* (mondat, propozíció) fogalmára is, ezért egy külön alfejezetet szentelünk a nulladrendű nyelvek Lindenbaum–Tarski-algebrájának, illetve a nyelven értelmezett valószínűségfüggvénynek.

2.1. A valószínűségi mérték

A valószínűségszámítás alapstruktúrája az *eseménytér*. Az eseménytér *hálóelméleti* bevezetése a következőképpen történik.[1] Legyen adva egy S halmaz, az események halmaza, és vezessük be a halmazon a metszet ($\wedge$) és az unió ($\vee$) műveleteket, amelyek az „és" illetve a „vagy" logikai műveleteket hivatottak reprezentálni. A megfelelő tulajdonságok (asszociativitás, kommutativitás, stb.) megkövetelése után az ilyen halmazt hálónak nevezzük. A halmazon bevezetett műveletek természetes módon meghatároznak egy $\leq$ részben rendezést a halmazon az

$$a \leq b \text{ akkor és csak akkor, ha } a \wedge b = a$$

kikötéssel, ahol $a, b \in S$. Tegyük fel, hogy erre a rendezésre nézve létezik egy legkisebb illetve egy legnagyobb eleme a halmaznak. Jelöljük ezt a két elemet 0-val illetve 1-gyel. A 0-ra mint a *lehetlen eseményre*, az 1-re pedig mint a *biztos eseményre* is szokás hivatkozni. Tegyük fel továbbá azt is, hogy az unió és a metszet műveletek disztibutívak. Ez a feltevés jelenti az elágazási pontot a klasszikus és a nem-klasszikus eseményterek között. Végül a „nem" logikai művelet reprezentálására vezessünk be egy harmadik műveletet a hálón, az ún. ortokomplementációt ($^\perp$) ugyancsak a megfelelő tulajdonságokkal (involúció, kontrapozíció, stb.) Az ezekkel a kikötésekkel kapott algebrai struktúrát a hálóelméletben *Boole-algebrának* nevezzük. A Boole-algebra tehát egy $\Sigma = (S, \wedge, \vee, ^\perp, 0, 1)$ ötös: egy disztributív, null- és egységelemes, ortokomplementáris háló. Ha a műveletek megszámlálhatóan végtelen elemre is zártak (vagyis ha megszámlálhatóan végtelen sok elem metszete és uniója szintén eleme S-nek), akkor a halmazt *Boole-σ-algebrának*, vagy röviden *σ-algebrának* nevezzük. A Boole-algebra egy elemét *atomnak* nevezzük, ha alatta a rendezésben már csak a 0 áll. Egy Boole-algebrát *atomosnak* mondunk, ha minden nem nulla elem alatt van atom, és *atomisztikusnak*, ha minden nem nulla elem „felépíthető" atomokból, azaz előáll mint atomok uniója. A klasszikus eseményteret tehát egy Boole-algebra reprezentálja, az eseményeket pedig a Boole-algebra elemei.

A Boole-algebra fogalmát azonban nemcsak hálóelméletileg lehet bevezetni, hanem *halmazelméletileg* is. Legyen Ω egy nem üres halmaz. Ω részhalmazainak egy Σ halmazát, *Boole-gyűrűnek* nevezzük, ha Σ tartalmazza minden $a \in \Sigma$ elem $\overline{a} \equiv \Omega \setminus a$ komplementerét, és zárt véges unióra nézve. Ha Σ tartalmazza Ω-t is, akkor Σ-t *Boole-algebrának* nevezzük, ha pedig az unióképzés megszámlálhatóan végtelen elemre is zárt, akkor *Boole-σ-algebrának*, vagy röviden *σ-algebrának*. A definícióból következik, hogy ha Σ egy σ-algebra, akkor tartalmazza az $\emptyset$ üreshalmazt, zárt megszámlálható metszetre nézve is, és teljesülnek benne a De Morgan-azonosságok. A lehetetlen eseményt és a biztos eseményt itt az $\emptyset$ üreshalmaz illetve az Ω alaphalmaz reprezentálja.

A Boole-algebra halmazelméleti és hálóelméleti fogalmai között a kapcsolatot a Stone-tétel teremti meg, amely kimondja, hogy az absztrakt hálóelméleti Boole-algebrák mindig reprezentálhatók halmazalgebrákkal. Mi a továbbiakban a Boole-algebra absztrakt és a

[1] A hálóelméleti alapokhoz, illetve az eseménytér hálóelméleti bevezetéséhez ld. az A. Függeléket.

halmazelméleti megközelítései között nem teszünk különbséget; az absztrakt jelölésrendszert fogjuk használjuk, de gyakran halmazokról fogunk beszélni.

Ha Ω egy nem üres halmaz, Σ pedig Ω részhalmazaiból képzett σ-algebra, akkor a (Ω, Σ) párt *mérhető térnek* nevezzük. A történetileg kialakult szóhasználat szerint Ω elemeit (amelyek általános esetben megfeleltethetők Σ atomjainak) *elemi eseményeknek*, Σ elemeit pedig *eseményeknek* vagy néha *összetett eseményeknek* nevezzük.[2]

Az eseménytér bevezetése után a valószínűség egy olyan az eseményekhez egy 0 és 1 közötti számot rendelő leképezés, amely diszjunkt elemeken additív. Szokásos bevezetése mértékelméleti alapon történik.

1. Definíció. Egy (Ω, Σ) mérhető téren értelmezett *(σ-additív) mértéken* egy olyan $\mu : \Sigma \to [0, \infty]$ leképezést értünk, amelyre az alábbiak teljesülnek minden megszámlálhatóan végtelen sok $a_i \in \Sigma$-ra:

(i) $\mu(\emptyset) = 0$,

(ii) $\mu(\vee_i a_i) = \sum_i \mu(a_i)$, ha $a_i \wedge a_j = \emptyset$ minden $i \neq j$-re.

Az (Ω, Σ, μ) hármast *mértéktérnek* nevezzük. Egy mértéket *valószínűségi mértéknek* nevezünk, ha a biztos eseményhez 1-et rendel, azaz $p(\Omega) \equiv \mu(\Omega) = 1$; a hozzá tartozó (Ω, Σ, p) hármast pedig *valószínűségi mértéktérnek*.

A valószínűségi mértéktér tehát egy halmazból, a halmaz részhalmazaiból képezett σ-algebrából, illetve a σ-algebra elemeihez 0 és 1 közötti számokat rendelő, a fenti előírásokat kielégítő valószínűségi mértékből áll.

A valószínűségszámítás nélkülözhetetlen fogalma a valószínűségi változó. Legyen (Ω, Σ) és (Ω', Σ') két mérhető tér.

2. Definíció. Egy $f : \Omega \to \Omega'$ függvényt *mérhetőnek* nevezünk, ha minden Σ'-beli elem ősképe Σ-ban van, azaz minden $a' \in \Sigma'$-re $f^{-1}(a') \in \Sigma$, ahol f^{-1} az f függvény teljes inverze. Ha (Ω', Σ')-t speciálisan az $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mérhető térnek választjuk, ahol $\mathcal{B}(\mathbb{R})$ a valós számok Borel-σ-algebrája (vagyis a valós számok nyílt intervallumait tartalmazó legszűkebb σ-algebra), akkor f-et *valószínűségi* vagy *véletlen változónak* nevezzük.

A valószínűségi változó értékkészletét időnként korlátozni szokás a valós számok valamely részhalmazára. Ha ez a részhalmaz megszámlálható, akkor a valószínűségi változót *diszkrétnek* mondjuk, egyébként *folytonosnak*. Diszkrét valószínűségi változókra fontos példa a *Bernoulli-változó*, ahol f képtere egy kételemű halmaz ($\{0, 1\}$). Ilyen Bernoulli-változók például minden $a \in \Sigma$-ra a $\chi^a : \Omega \to \{0, 1\}$ *karakterisztikus* vagy *indikátor függvények*:

$$\chi^a(x) \equiv \begin{cases} 1, & \text{ha } x \in a, \\ 0, & \text{ha } x \notin a. \end{cases}$$

[2] Az „összetett” jelző egy kissé félrevezető, hiszen ha Σ atomos, akkor elemei közé tartoznak az egyszerű események (pontosabban az egyszerű eseményeknek megfelelő halmazok) is.

A valószínűségi változó tehát egy olyan függvény, amely az elemi eseményekhez számokat rendel. Az a függvény, amely meghatározza, hogy egy valószínűségi változó értéke mekkora valószínűséggel esik egy bizonyos értéktartományba, az eloszlás.

3. Definíció. Legyen (Ω, Σ, p) valószínűségi mértéktér, $f : \Omega \to \mathbb{R}$ pedig egy valószínűségi változó. Ekkor a $p \circ f^{-1} : \mathcal{B}(\mathbb{R}) \to [0,1]$ függvényt az f valószínűségi változó *eloszlásának* nevezzük, ahol f^{-1} ismét csak az f függvény teljes inverze.

Egy $f : \Omega \to \mathbb{R}$ és egy $g : \Omega \to \mathbb{R}$ valószínűségi változót *azonos eloszlásúnak* nevezünk, ha $p \circ f^{-1}(b) = p \circ g^{-1}(b)$ minden $b \in \mathcal{B}(\mathbb{R})$ Borel-halmazra.

Szokás $p \circ f^{-1}$ eloszlás értékét egy adott $\{x\} \in \mathcal{B}(\mathbb{R})$ helyen az alábbi módon is jelölni: $p(f = x)$. Ezzel a jelöléssel az f Bernoulli-változó eloszlása, a *Bernoulli-eloszlás*, például a következőképpen írható:

$$\begin{aligned} p(f = 1) &= p, \\ p(f = 0) &= 1 - p, \end{aligned}$$

ahol $p \in [0,1]$; a χ^a karakterisztikus függvények eloszlása pedig így:

$$\begin{aligned} p(\chi^a = 1) &= p(a), \\ p(\chi^a = 0) &= p(a^{\perp}). \end{aligned}$$

Hasonlóan fontos eloszlás az *egyenletes (uniform) eloszlás*. Egy g diszkrét valószínűségi változót egyenletes eloszlásúnak nevezünk, ha g minden x_i $(i = 1 \ldots n)$ értéket azonos valószínűséggel vesz fel, vagyis ha a képterébe tartozó bármely x_i elemre:

$$p(g = x_i) = \frac{1}{n}.$$

Folytonos valószínűségi változó esetében az egyenletes eloszlásnak ez az egyszerű meghatározása nem működik, mert a $p(g = x)$ valószínűség minden x-re nulla. Így ahelyett a fogalom helyett, hogy g mekkora valószínűséggel vesz fel egy adott x értéket, a folytonos esetben csak azt tudjuk értelmezni, hogy g mekkora valószínűséggel vesz fel értéket az $[x, x + dx]$ infinitezimális *tartományban*. Ennek értelmezéséhez további két fontos fogalmat kell bevezetnünk, az eloszlásfüggvényt és a sűrűségfüggvényt. A g valószínűségi változó *eloszlásfüggvényének* az

$$F_g : \mathbb{R} \to \mathbb{R}; \; F_g(x) \equiv p \circ g^{-1}((-\infty, x))$$

nemcsökkenő, balról folytonos függvényt nevezzük, amelynek határértéke az $x \to -\infty$-ben 0, az $x \to \infty$-ben pedig 1. $F_g(x)$ annak a valószínűségét adja meg, hogy a g valószínűségi változó értéke kisebb x-nél. Ha az $F_g(x)$ eloszlásfüggvény előáll az

$$F_g(x) = \int_{-\infty}^{x} f_g(x')dx'$$

integrálként, ahol $f_g : \mathbb{R} \to \mathbb{R}$ egy nem negatív Lebesgue-integrálható függvény, akkor f_g-t a g valószínűségi változó *sűrűségfüggvényének* nevezzük. Intuitíve tehát $f_g(x)\,dx$ annak a valószínűsége, hogy a g valószínűségi változó értéke az $[x, x+dx]$ tartományba esik. Diszkrét valószínűségi változásnál az $f_g(x_i)$ sűrűségfüggvény a fenti $p(g = x_i)$ alakot ölti.

A sűrűségfüggvény segítségével most már könnyen megfogalmazhatjuk az egyenletes eloszlást folytonos valószínűségi változók esetében is. Egy g folytonos valószínűségi változó egyenletes eloszlású egy (a, b) tartományon, ha

$$f_g(x) \equiv \begin{cases} \frac{1}{b-a}, & \text{ha } x \in (a, b], \\ 0, & \text{ha } x \notin (a, b]. \end{cases}$$

Az eloszlást jellemző további két fontos mennyiség a várható érték és a szórás.

4. Definíció. Legyen (Ω, Σ, p) valószínűségi mértéktér, $f : \Omega \to \mathbb{R}$ pedig egy valószínűségi változó. Az f valószínűségi változó $Exp(f)$ *várható értékét* az alábbi integrállal definiáljuk (ha az létezik):

$$Exp(f) \equiv \int_\Omega f\,dp,$$

ami megszámlálható értelmezési tartományú diszkrét valószínűségi változó esetében az alábbi összegre egyszerűsödik:

$$Exp(f) \equiv \sum_i f(x_i)\,p(\{x_i\}),$$

ahol $x_i \in \Omega$ minden i-re. Az f valószínűségi változó *szórása* az alábbi módon definiált mennyiség:

$$Var(f) \equiv \sqrt{Exp(f^2) - (Exp(f))^2}.$$

A fenti fogalmak illusztrálására tekintsünk az alábbi példát.

Példa. Legyen $\Omega = \{1, 2, 3, 4, 5, 6\}$ hatelemű halmaz, és legyen $\Sigma = \mathcal{P}(\Omega)$ az Ω 64 elemű hatványhalmaza. Σ-ban nyilvánvalóan értelmezve vannak a szokások halmazelméleti műveletek. Az $a \equiv \{1, 2, 3\}$ és a $b \equiv \{2, 4, 6\}$ események metszete a $c \equiv a \wedge b = \{2\}$ esemény, uniója a $d \equiv a \vee b = \{1, 2, 3, 4, 6\}$ esemény, komplementereik pedig az $a^\perp = \{4, 5, 6\}$ illetve a $b^\perp = \{1, 3, 5\}$ események. A műveletek disztributívak, és teljesül a De Morgan-szabály. Az elemi események, vagyis Σ atomjai, azaz legkisebb nem nulla elemei az $a_i \equiv \{i\}$ események. Az eseménytér atomos: minden összetett esemény alatt van atom, továbbá atomisztikus: minden összetett esemény atomok uniója. A fentiek például így állnak elő:

$$\begin{aligned} a &\equiv a_1 \vee a_2 \vee a_3 \\ b &\equiv a_2 \vee a_4 \vee a_6 \end{aligned}$$

A háló koatomjai (ld. az A. Függeléket) az $a_i^{\perp}$ események, lefedésre pedig példa az $\{1\}$ esemény lefedése az $\{1, 2\}$ eseménnyel.

Definiáljuk ezek után az eseménytéren a $p : \Sigma \to [0, 1]$ valószínűségi mértéket a következőképpen:

$$p(\{i\}) \equiv \begin{cases} \frac{1}{3}, & \text{ha } i \text{ páros,} \\ 0, & \text{ha } i \text{ páratlan.} \end{cases} \tag{2.1}$$

Az a esemény valószínűsége például (az additivitás miatt) ekkor $\frac{1}{3}$ lesz, a b eseményé pedig 1.

Vezessünk be végül egy $f : \{1, 2, 3, 4, 5, 6\} \to \mathbb{R}$ valószínűségi változót az eseménytéren a következőképpen: $f(i) = i$. A valószínűségi változó várható értéke és szórása könnyen kiszámolható:

$$\begin{aligned} Exp(f) &= 1 \times 0 + 2 \times \frac{1}{3} + 3 \times 0 + 4 \times \frac{1}{3} + 5 \times 0 + 6 \times \frac{1}{3} = 4 \\ Var(f) &= \sqrt{(1 \times 0 + 4 \times \frac{1}{3} + 9 \times 0 + 16 \times \frac{1}{3} + 25 \times 0 + 36 \times \frac{1}{3}) - 16} = \sqrt{\frac{8}{3}} \end{aligned}$$

Valószínűségi mértékterekre további példákat találunk a B. Függelékben. Most azonban lépjünk tovább a valószínűségelméletnek egy másik megfogalmazása felé, amely alapfogalma nem az *esemény*, hanem a *kijelentés*. Amint azt azonban látni fogjuk, a Boole-algebra a valószínűségnek ebben a megközelítésében is kulcsszerepet kap.

2.2. A Lindenbaum–Tarski-algebra

Kezdjük a nulladrendű nyelv fogalmával.

5. Definíció. *Nulladrendű formális nyelven* egy olyan $\mathcal{L} = (Log, Con, Form)$ hármast értünk, ahol $Log = \{\&, v, \sim, (,)\}$ a logikai konstansok, Con a nemlogikai konstansok, $Form$ pedig a formulák osztálya, azaz a legszűkebb olyan osztály, amelyre az alábbiak fennállnak:

(i) Ha $A \in Con$, akkor $A \in Form$;

(ii) Ha $A \in Form$, akkor $\sim A \in Form$;

(iii) Ha $A, B \in Form$, akkor $A \,\&\, B \in Form$;

(iv) Ha $A, B \in Form$, akkor $A \, v \, B \in Form$.

6. Definíció. $\mathcal{L}$ egy *interpretációján* egy olyan $I : Form \to \{0, 1\}$ leképezést értünk, amelyre minden $A, B \in Form$-ra teljesül, hogy

$$\begin{aligned}
I({\sim}A) &= \begin{cases} 1, & \text{ha } I(A) = 0, \\ 0 & \text{különben;} \end{cases} \\
I(A \,\&\, B) &= \begin{cases} 1, & \text{ha } I(A) = 1 \text{ és } I(B) = 1, \\ 0 & \text{különben;} \end{cases} \\
I(A \,v\, B) &= \begin{cases} 0, & \text{ha } I(A) = 0 \text{ és } I(B) = 0, \\ 1 & \text{különben.} \end{cases}
\end{aligned}$$

7. Definíció. (i) *Logikai igazságnak* nevezünk egy A formulát, ha hozzá minden interpretáció 1-et rendel. Jelöléssel: $\Rightarrow A$.

(ii) *Ellentmondásnak* nevezünk egy A formulát, ha hozzá minden interpretáció 0-át rendel. Jelöléssel: $\not\Rightarrow A$.

(iii) Egy B formula akkor *következik* (szemantikailag) egy A formulából, ha $I(B) = 1$ minden interpretációban, ahol $I(A) = 1$. Jelöléssel: $A \Rightarrow B$.

(iv) Két formula, A és B *logikailag ekvivalens*, ha $A \Rightarrow B$ és $B \Rightarrow A$. Jelöléssel: $A \Leftrightarrow B$.

A logikai ekvivalencia két formula között azt jelenti, hogy a két formula igazságértéke minden interpretációban megegyezik, azaz a két formula ugyanazt jelenti. A logikailag ekvivalens formulák a nyelv egyfajta redundanciáját fejezik ki a világ tényállásaival szemben: egyazon tényállásra vagy eseményre több logikailag ekvivalens formulával is utalhatunk. Ahhoz, hogy megszabaduljunk ettől a redundanciától, és eljussunk a nyelv által reprezentált eseményekhez, a következőket kell tennünk.

A $\Leftrightarrow$ reláció ekvivalenciareláció a formulák halmazán. Tekintsük $Form$-nak a $\Leftrightarrow$ szerinti $Form/_{\Leftrightarrow}$ faktorizációját, vagyis az ekvivalenciaosztályok halmazát. Az alábbiakban megmutajuk, hogy a $Form/_{\Leftrightarrow}$ halmaz egy Boole-algebra.

Jelölje $[A]$ az A formulával ekvivalens formulák osztályát. Definiáljuk a $\wedge, \vee$ és $^{\perp}$ műveleteket $Form/_{\Leftrightarrow}$-n az alábbi módon:

$$\begin{aligned}
[A] \wedge [B] &\equiv [A \,\&\, B] \\
[A] \vee [B] &\equiv [A \,v\, B] \\
[A]^{\perp} &\equiv [{\sim}A]
\end{aligned}$$

Könnyen látható, hogy a műveletek jól definiáltak, vagyis nem függnek az ekvivalenciaosztályt reprezentáló elemtől. Ha az A' formula az $[A]$ ekvivalenciaosztályba tartozik, B' pedig a $[B]$ ekvivalenciaosztályba, akkor $A' \,\&\, B'$ az $[A \,\&\, B]$, $A' \,v\, B'$ az $[A \,v\, B]$, ${\sim}A'$ pedig az $[{\sim}A]$ ekvivalenciaosztályba fog tartozni az interpretáció függvény tulajdonságai miatt. Könnyen ellenőrizhető, hogy a $\wedge$ és $\vee$ műveletek disztributívak, a $^\perp$ művelet pedig a halmazelméleti komplementum (vagy az absztrakt hálóelméleti ortokomplementáció, ld. A. Függelék) tulajdonságaival rendelkezik. Ezekkel a műveletekkel tehát, valamint az ellentmondások $[\downarrow]$ ekvivalenciaosztályával mint nullelemmel, illetve a logikai igazságok $[\uparrow]$ ekvivalenciaosztályával mint egységelemmel a $\Sigma = (Form/_{\Leftrightarrow}, \wedge, \vee, ^\perp, [\downarrow], [\uparrow])$ struktúra Boole-algebrát definiál. Ezt az algebrát nevezik az $\mathcal{L}$ nyelv *Lindenbaum–Tarski-algebrájának.*[3]

Most jöjjön a valószínűség definíciója. A valószínűség a formulák halmazán értelmezett olyan nem negatív értékű leképezés, amely tekintetbe veszi a nyelv logikai struktúráját, azaz tautológiákhoz 1-et rendel, és egymást kizáró kijelentéseken pedig additív.

8. Definíció. Legyen $\mathcal{L} = (Log, Con, Form)$ egy nulladrendű nyelv. Egy $P : Form \to [0,1]$ hozzárendelést a nyelven *valószínűségfüggvénynek* nevezünk, ha minden $A \in Form$-ra és megszámlálható $A_i \in Form$-ra igazak az alábbi kikötések:

(i) $P(A) = 1$, ha $\Rightarrow A$,

(ii) $P(v_i\, A_i) = \sum_i P(A_i)$, ha $\not\Rightarrow A_i \,\&\, A_j$ minden $i \neq j$-re.

Ezekből a kikötésekből könnyen láthatóan adódik, hogy

$$P(A) = P(B), \quad \text{ha} \quad A \Leftrightarrow B,$$

vagyis P azonos értéket vesz fel a nyelv ekvivalens elemein. A formulák halmazát faktorizálva a $\Leftrightarrow$ ekvivalenciarelációval tehát P-ből a nyelv Lindenbaum–Tarski-algebráján értelmezett p valószínűségi mértéket kapunk. Mindezt könnyen szemléltethetjük az előző alfejezet (25. oldal) példáján.

Példa. Legyen $\mathcal{L}$ egy nulladrendű nyelv, ahol a nemlogikai konstansainak halmaza $Con = \{A_i \,|\, i = 1\ldots 6\}$, és szorítsuk meg a nyelv lehetséges interpretációit arra a hat I^{A_i} interpretációra, amelyek mindegyike Con pontosan egy elemén igazak:

$$I^{A_i}(A_j) \quad \equiv \quad \begin{cases} 1, & \text{ha } i = j, \\ 0, & \text{ha } i \neq j. \end{cases}$$

A nyelv formulái közé ilyen mondatok fognak tartozni:

$$\begin{aligned} A &\equiv A_1\, v\, A_2\, v\, A_3, \\ B &\equiv A_2\, v\, A_4\, v\, A_6. \end{aligned}$$

[3] Megmutatható, hogy minden klasszikus kijelentéslogika Lindenbaum–Tarski-algebrája Boole-algebrát definiál, a klasszikus elsőrendű logikák algebrizálása pedig ún. *cilindrikus algebrához* vezet, amely a Boole-algebrának természetes kiterjesztése.

A következményrelációra példa az $A_1 \Rightarrow A$ és a $A_2 \Rightarrow B$ következtetés, logikai ekvivalenciára pedig az

$$\begin{aligned} A \,\&\, B &\Leftrightarrow A_2, \\ A \,v\, B &\Leftrightarrow {\sim} A_5 \end{aligned}$$

relációk. Faktorizálva az ekvivalenciarelációval könnyen látható, hogy minden ekvivalenciahalmaz valamely $[A_i]$ ekvivalenciahalmazok uniójaként áll elő, vagyis $Form/_{\Leftrightarrow}$ a $\mathcal{P}(\{1,2,3,4,5,6\})$ halmazalgebrával lesz izomorf.

A nyelven ezek után definiálhatjuk a $P : Form \to [0,1]$ valószínűségi mértéket a következőképpen:

$$P(A_i) \equiv \begin{cases} \frac{1}{3}, & \text{ha } i \text{ páros,} \\ 0, & \text{ha } i \text{ páratlan.} \end{cases}$$

Ez a hozzárendelés a P függvény tulajdonságai miatt a nyelv összes formuláján jól definiált. Jól látható továbbá, hogy a P leképezés által a nyelv Lindenbaum–Tarski-algebráján meghatározott p valószínűségi mérték éppen (2.1) lesz.

2.3. Függetlenség és korreláció

A valószínűség filozófiája szempontjából nagyon fontos a feltételes valószínűség fogalma.

9. Definíció. Az a esemény *feltételes (kondicionális) valószínűségét* a b eseményre nézve, amennyiben $p(b) \neq 0$, a következőképpen definiáljuk:

$$p(a|b) \equiv \frac{p(a \wedge b)}{p(b)}. \tag{2.2}$$

A definíciója alapján minden nem nulla mértékű b eseményre értelmezhetjük a $p_b \equiv p(\,\cdot\,|b)$ hozzárendelést, amely egy valószínűségi mértéket definiál az algebrán, a *feltételes valószínűségi mértéket*, amely fontos szerepet játszik majd a bayesianizmusban. A definíciónak további közvetlen következménye a *Bayes-tétel*,[4] amely szerint minden a és b nem nulla valószínűségű eseményre fennáll, hogy

$$p(a|b) = \frac{p(b|a)p(a)}{p(b)}. \tag{2.3}$$

Mielőtt tovább lépnénk, szeretnénk megjegyezni, hogy a feltételes valószínűség fogalmát néhány szerző (pl. Rényi 1970) nem a (2.2) összefüggés alapján definiálja a *nem* feltételes valószínűségek segítségével, hanem egyszerűen alapfogalomnak tekinti. A feltételes valószínűség előnyben részesítését több dolog indokolja. Egyrészt a feltételes valószínűség nem

[4] A Bayes-tételre vonatkozó összefüggéseket ld. a D. Függelékben.

definiálható a $p(b) = 0$ esetben, vagyis nulla valószínűségű feltételre. Ha azonban a nulla valószínűségű eseményeket (ahogy azt szokás) nem tekintjük lehetetlen eseményeknek, akkor intuitíve legalábbis nem világos, hogy ezek az események miért ne szerepelhetnének feltételes valószínűségekben. Ráadásul ilyen események „sokan" lehetnek: nem megszámlálható mértékterekben nem megszámlálhatóan. A $[0, 1]$ intervallumon értelmezett uniform mérték esetében például egyik elemi eseményre mint feltételre sem lehet feltételes valószínűséget értelmezni. Visszafelé, a feltételes valószínűségből azonban könnyen származtatható a nem feltételes valószínűség – a feltétel helyébe egyszerűen az Ω biztos eseményt (vagy ha a valószínűség egy nyelven van értelmezve, akkor egy tetszőleges $\uparrow$ tautológiát) kell beírni. Az pedig egy további, már az interpretációkra tartozó kérdés, hogy vajon a feltételes valószínűség mint két nem feltételes valószínűség *egyszerű aránya* helyesen adja-e vissza a feltételes valószínűséggel szemben támasztott egyéb elvárásainkat (lásd (Szabó 2002, 87. o.)). Ezekkel a kérdésekkel azonban mi itt nem foglalkozunk, hanem a feltételes valószínűséget a fenti kolmogorovi szellemben értjük.

Az interpretációk szempontjából egy másik fontos fogalom a függetlenség fogalma. Legyen (Ω, Σ, p) egy valószínűségi mértéktér, és legyen $a, b \in \Sigma$. Az a és egy b eseményt (valószínűségi értelemben) *függetlennek* nevezzük, ha

$$p(a \wedge b) = p(a)p(b),$$

vagy két ekvivalens formában, ha

$$\begin{aligned} p(a|b) &= p(a), \\ p(b|a) &= p(b). \end{aligned}$$

Nagyon fontos látni, hogy események függetlenségéről csak azután lehet beszélni, miután a valószínűségi mértéktér adva van.[5] Ha két esemény nem független, akkor azt mondjuk, hogy *korrelálnak*. Ha az a és b esemény pozitívan korrelál, vagyis ha

$$p(a \wedge b) > p(a)p(b),$$

akkor az egyik eseményt a negáltjára (ortokomplementumára) cserélve a korreláció negatív lesz:

$$\begin{aligned} p(a^\perp \wedge b) &< p(a^\perp)p(b), \\ p(a \wedge b^\perp) &< p(a)p(b^\perp), \end{aligned}$$

mindkét eseményt felcserélve pedig ismét pozitív:

$$p(a^\perp \wedge b^\perp) > p(a^\perp)p(b^\perp).$$

[5] A függetlenség fogalmának jelentőségét mi sem bizonyítja jobban, mint Kolmogorov (1933, 21.o.) azon megjegyzése, hogy a függetlenség fogalma az, ami megkülönbözteti a valószínűségelméletet a mértékelmélettől.

A függetlenséget kettőnél több eseményre is általánosíthatjuk. Az $a_1, a_2 \ldots a_n \in \Sigma$ eseményeket *teljesen függetlennek* mondjuk, ha az $\{1, 2 \ldots n\}$ indexhalmaz bármely $\{i_1, i_2 \ldots i_k\}$ részhalmazára fennáll, hogy

$$p(\wedge_{i_k} a_{i_k}) = \prod_{i_k} p(a_{i_k}).$$

A függetlenséget ezek után halmazokra is értelmezhetjük. A $\Sigma_j \subseteq \Sigma$ $(j = 1 \ldots m)$ *halmazokat* akkor nevezzük *függetlennek*, ha mindegyik halmazból kiválasztva egy tetszőleges elemet, azok teljesen függetlenek lesznek.

Az alábbi, Kolmogorovtól (1933) származó példa arra világít rá, hogy események páronkénti függetlenségéből nem következik teljes függetlenségük. Legyen $\Omega = \{1, 2, 3, 4\}$, legyen $\Sigma = \mathcal{P}(\Omega)$, és legyen p uniform valószínűségi mérték, azaz legyen $p(\{i\}) = \frac{1}{4}$ minden $i \in \Omega$-ra. Tekintsük Σ alábbi három elemét:

$$\begin{aligned} a &\equiv \{1, 2\} \\ b &\equiv \{1, 3\} \\ c &\equiv \{1, 4\} \end{aligned}$$

Nyilvánvaló, hogy a, b és c páronként függetlenek, hiszen

$$\begin{aligned} \frac{1}{4} = p(a \wedge b) &= p(a)p(b) = \frac{1}{2} \cdot \frac{1}{2} \\ \frac{1}{4} = p(a \wedge c) &= p(a)p(c) = \frac{1}{2} \cdot \frac{1}{2} \\ \frac{1}{4} = p(b \wedge c) &= p(b)p(c) = \frac{1}{2} \cdot \frac{1}{2} \end{aligned}$$

ugyanakkor azonban

$$\frac{1}{4} = p(a \wedge b \wedge c) \neq p(a)p(b)p(c) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$$

így az események nem *teljesen* függetlenek. Erre a példára mint *Bernstein-paradoxonra* is szokás hivatkozni, de a példa paradox jellege természetesen nem látható mindaddig, amíg a valószínűséget nem interpretáltuk.[6]

2.4. A nagy számok törvényei

A függetlenség fogalma mellett a valószínűség interpretációiban kulcsszerepet játszanak a nagy számok törvényei. A nagy számok törvényei valószínűségi változók aritmetikai átlagára vonatkozó konvergenciatörvények. A törvényekhez a valószínűségi változók kétféle konvergenciáját kell definiálnunk:

[6] A Bernstein-paradoxon és a reichenbachi közös ok elv kapcsolatáról lásd még (Uffink, 1999).

10. Definíció. Az (Ω, Σ, p) valószínűségi mértéktéren valószínűségi változók egy $f_1, f_2 \ldots$ sorozata *valószínűségi értelemben konvergál* egy f valószínűségi változóhoz, jelöléssel $f_i \xrightarrow{p} f$, ha minden $\varepsilon > 0$ valós számra

$$\lim_{n\to\infty} p(|f_n - f| \leqslant \varepsilon) = 1,$$

vagyis ha az $a_n \equiv \{x \in \Omega \ : \ |f_n(x) - f(x)| \leqslant \varepsilon\}$ események valószínűsége n növekedtével 1-hez tart.

11. Definíció. Az (Ω, Σ, p) valószínűségi mértéktéren valószínűségi változók egy $f_1, f_2 \ldots$ sorozata *majdnem biztosan (majdnem mindenütt) konvergál* egy f valószínűségi változóhoz, jelöléssel $f_i \xrightarrow{m.b.} f$, ha

$$p(\lim_{n\to\infty} f_n = f) = 1,$$

vagyis ha a valószínűségi változók pontonkénti konvergenciájával definiált $a \equiv \{x \in \Omega \ : \lim_{n\to\infty} f_n(x) = f(x)\}$ esemény valószínűsége 1.

Nem nehéz belátni, hogy a második konvergencia erősebb, azaz a majdnem biztosan való konvergenciából következik a valószínűségi értelemben vett konvergencia.

A nagy számok törvényei valószínűségi változók átlagára vonatkozó konvergenciákat fogalmaznak meg a konvergencia fenti két definíciójának megfelelően. A valószínűségi értelemben vett konvergenciára vonatkozó tételeket közös néven a *nagy számok gyenge törvényeinek* nevezzük, a majdnem biztosan való konvergenciára vonatkozó törvényeket a *nagy számok erős törvényeinek*. Mindkét csoportba tételek tucatjai tartoznak a valószínűségi változókra kirótt különféle feltételeknek megfelelően. Mi az alábbiakban mindkét csoportból csak egy-egy klasszikusnak számító tételt mondunk ki. (A bizonyításokat lásd például (Billingsley 1995, 85-88. o.)).

1. Tétel. (A nagy számok gyenge törvénye, Bernoulli-tétel.) Legyen (Ω, Σ, p) valószínűségi mértéktér, és legyen $f_1, f_2 \ldots$ páronként független[7] és azonos eloszlású valószínűségi változók egy végtelen sorozata, ahol a valószínűségi változók várható értéke és szórása minden i-re létezik, $Exp(f_i) = m$ és $Var(f_i) = \sigma$. Ekkor az

$$s_n \equiv \frac{1}{n}\sum_{i=1}^{n} f_i \tag{2.4}$$

összeg mint valószínűségi változó valószínűségi értelemben konvergál a közös várható értékhez, azaz $s_n \xrightarrow{p} m$.

[7] Valószínűségi változók különféle függetlenségeinek definícióit lásd a D. Függelékben.

2. Tétel. (A nagy számok erős törvénye.) Legyen (Ω, Σ, p) valószínűségi mértéktér, és legyen $f_1, f_2 \dots$ páronként független és azonos eloszlású valószínűségi változók egy végtelen sorozata, ahol a valószínűségi változók várható értéke minden i-re $Exp(f_i) = m$, és minden valószínűségi változó véges $Exp(f_i^4) = m_4$ negyedik momentummal rendelkezik. Ekkor a (2.4) összeg majdnem biztosan konvergál a közös várható értékhez, azaz $s_n \xrightarrow{m.b.} m$.

A nagy számok gyenge törvényét Bernoulli (1713) bizonyította először Bernoulli-eloszlású változókra, az erős törvényre pedig az első történeti példa Borel (1909) ún. *normál szám*-tétele. (A Bernoulli-tétel alkalmazását karakterisztikus függvényekre lásd a D. Függelékben.) Megjegyezzük továbbá, hogy a nagy számok törvényei nem specifikálják a konvergencia gyorsaságát.[8]

[8] Ehhez az ún. *iterált logaritmus törvényét* kell segítségül hívni, ld. (Billingsley, 1995).

3. fejezet

A valószínűség interpretációi

Miután az előző fejezetben röviden ismertettük a matematika „valószínűség” fogalmát, rátérhetünk könyvünk központi kérdésére: mi is az a valószínűség? Ahhoz azonban, hogy erre a kérdésre választ kapjunk, vissza kell lépnünk egy lépést, és tisztáznunk kell, hogy mit is értünk a valószínűség interpretációja alatt?

Először is gondosan el kell különítenünk a valószínűség fogalmának interpretációját a fogalom explikációjától. Egy köznyelvi fogalom *explikációja* vagy magyarázata a fogalom helyettesítése valamilyen tudományos vagy filozófiai fogalommal. Hogy ez a helyettesítés mennyire sikeres, az attól függ, hogy a magyarázó fogalom egyfelől mennyire őrzi meg a hasonlóságát a magyarázandóval, másfelől mennyiben ad pontosabb alkalmazhatósági kritériumokat a kezünkbe az adott tudományos vagy filozófiai taxonómiában. A hőmérséklet fogalmának tudományos explikációja a köznapi hőmérsékletetfogalom helyettesítése egy pontosan körvonalazható jelentéssel bíró termodinamikai fogalommal; a tudás fogalmának filozófiai explikációja a tudás köznapi fogalmának helyettesítése egy pontosan körvonalazható jelentésű ismeretelméleti fogalommal.

Ezzel szemben szoros értelemben *interpretálni* egy formális, tehát jelentéssel nem rendelkező struktúrát szokás – az interpretáció vagy más néven szemantika éppen ennek a jelentés nélküli struktúrának a felruházása jelentéssel. A formális struktúra általában egy elsőrendű formális nyelv, a nyelv egy mondatának jelentése pedig – legalábbis a jelentés igazságfeltétel-elmélete szerint – „valami olyasmi, ami meghatározza azokat a feltételeket, amelyek mellett a szóban forgó mondat igaz, illetve hamis” (Lewis 1972).

Mármost explikálnunk vagy interpretálnunk kell a valószínűséget? Nyilvánvaló, hogy a „valószínű” a köznyelv sajátos használattal bíró kifejezése, és így annak a tudományos vagy filozófiai elméletnek, amely használni kívánja a valószínűség fogalmát, tekintettel kell lennie erre a köznyelvi használatra. Ez hasonlít arra, ahogy a „hőmérséklet” vagy a „tudás” terminusok csak a kifejezések *common sense* használatának figyelembevétele mellett kaphatnak helyet egy új elméletben – míg ezek a kényszerek nem kötik például az olyan definícióval bevezetett fogalmakat, mint az „entrópia” vagy a „ráépülés”. Mindez azt mutatja, hogy a valószínűség fogalmát mindenekelőtt explikálnunk kell.

Az explikáció a tudományelőtti fogalmat egy tudományos elmélet terminusával helyettesíti. Egy tudományos elmélet egy formális vagy részben formalizált nyelvből és egy interpretációból álló pár. A hőmérséklet fogalmának explikációja tehát a köznyelvi hőmérsékletfogalom helyettesítése a termodinamika mint tudományos elmélet, azaz mint részben interpretált formális nyelv hőmérsékletfogalmával. A többi tudományos fogalommal is hasonló a helyzet: a „gyorsaság" köznyelvi fogalmát a kinematika sebességfogalma explikálja, az „egyszerre" fogalmát a relativitáselmélet szimultaneitásfogalma.

Az explikáció szempontjából a valószínűség fogalma azonban a többi fogalomtól eltérő történeti utat járt be. Míg a fenti explikált fogalomak elnyerték a helyüket egy tudományos elméletben, addig a valószínűség fogalma az európai gondolkodásban betöltött kardinális szerepe ellenére *nem vált fizikai fogalommá*. E helyett viszont kialakult egy matematikai fogalom és egy matematikai elmélet: a matematikai valószínűség illetve a valószínűségelmélet, amely úgy tűnt, hogy helyettesíteni képes a valószínűség köznyelvi fogalmát. Ezt az igényét tükrözte már pusztán a szóhasználata is: eseményekről és ezek valószínűségeiről beszélt, miközben halmazokat és függvényeket használt. A helyzet a geometriával rokonítható, ahol is a „pont", „egyenes", stb. köznyelvi fogalmából megszületett egy matematikai elmélet, a geometria.

Persze mondhatnánk erre, hogy a valószínűség éppen olyan explikált fogalom mint a hőmérséklet, hiszen számos fizikai elméletben felbukkan: használja a statisztikus fizika, a kvantumelmélet, stb. Másrészt a „pont" és „egyenes" fogalmával is hasonló a helyzet: önmagukban mint geometriai terminusok még interpretálatlanok, de egy fizikai elméletbe, például a relativitáselméletbe kerülve interpretációt nyernek. Van azonban egy nagy különbség. Amíg a termodinamika a hőmérséklet fogalmát *közvetlenül* interpretálja, vagyis megmondja, hogy mit ért hőmérsékleten: azt, ti. amit a hőmérő mér (legalábbis bizonyos hőmérséklettartományban) – addig a valószínűséget tartalmazó elméletek egyike sem mondja meg, hogy mit is ért valószínűségen. A fogalom pusztán egy elméleti terminus. A kérdés éppen az, hogy mit jelent a valószínűség akár a statisztikus fizikában, akár a kvantumelméletben.

Vannak tehát tudományos elméleteink egyfelől, bennük szereplő valószínűségekkel, és van egy matematikai elméletünk, a valószínűségelmélet, benne a matematikai valószínűséggel – de egyik elmélet sem mondja meg, hogy mit is értsünk valószínűségen. Ahogy Poincaré tréfásan megjegyzi, a matematikusok azt hiszik, hogy a valószínűség törvényeit a természetből olvastuk ki, a természettudósok pedig azt, hogy matematikailag vezettük le őket. Ezen felül ott van továbbra is a köznyelv explikálatlan valószínűségfogalma. Vagyis a valószínűségnek mind az explikációja, mind az interpretációja egyelőre várat magára. Melyik a helyes kiindulópont a valószínűség értelmezéséhez?

Nyilvánvalóan a valószínűség fogalmának elemzésében nem bízhatjuk magunkat teljes mértékben a fogalom köznyelvi használatára. A köznyelvi fogalomhasználat regulatív szerepét elfogadni egyet jelentene azzal, hogy magunkévá tesszük a hétköznapi nyelv filozófiájának azt a metafizikai prekoncepcióját, amely szerint a fogalmak metafizikai elemzésének *par excellence* terepe a köznyelv, így minden filozófiai elemzés alfája a természetes nyelvi használat. Számosan érveltek már ez ellen az elképzelés ellen, amely érveket itt most nem soroljuk fel. A valószínűség elemzéséhez ennél többre van szükség – valójában nem kevesebbre, mint *a valószínűség empirikus elméletének megalkotására*, vagyis egy olyan elméletre, amely közvetlenül vagy közvetve empirikus fogalmakban *definiálja*, hogy mit is értünk valószínűségen. Egy olyan interpretált formális nyelvre tehát, amelyben szerepel egy p terminus, és amely nyelv interpretációja egyértelműen rögzíti ennek a terminusnak a fizikai jelentését. A köznyelvi elemzés itt pusztán annyiban juthat szerephez, amennyiben koordinálja az elméletalkotás folyamatát: kiemeli a fogalom köznyelvi használatának legfontosabb jegyeit, előzetesen osztályozza a használat esetlegesen eltérő köreit, stb.

Hasonló a helyzet azonban az interpretáció tekintetében is. Ha a valószínűség fogalmának analízisét úgy értelmezzük, mint a valószínűségszámítás formális nyelvének interpretációját, akkor úgymond a szekeret fogjuk a ló elé. Milyen garanciáink vannak ugyanis arra nézve, hogy ennek a formális elméletnek az interpretációja éppen a fizikai valószínűséghez vezet majd bennünket? Elvégre a normált területmérték is kielégíti a valószínűség axiómáit, így akár egy lakóház tulajdonhányadait is tekinthetnénk a valószínűség „helyes interpretációjának". Vagy megfordítva a kérdést, mi a garanciánk arra, hogy a „tényleges" valószínűség (akármi legyen is az) éppen úgy viselkedik, ahogy a Kolmogorov-féle valószínűségszámítás állítja? Erre azon túl, hogy a valószínűségszámítás háromszáz év kombinatorikai, statisztikai problémáinak gondos elemzésével jött létre, és így helyesen „kell", hogy leírja a tényleges valószínűséget – nincs semmilyen garanciánk.

Ezért a valószínűség fogalmi elemzésében a helyes sorrend a következő. Amennyiben adva van a valószínűségnek egy értelmezése, először is azt kell megvizsgálnunk, hogy a kérdéses értelmezés *empirikusan* rögzíti-e a benne szereplő „valószínűség" terminus jelentését – röviden, hogy az értelmezés valóban egy empirikus elmélet-e. Az explikáció és interpretáció kérdése csak ezek *után* jöhet szóba; vagyis csak ezek után lehet feltenni a kérdést, hogy a szóban forgó értelmezés valószínűségfogalma milyen viszonyban áll a köznyelv valószínűségfogalmával illetve a történetileg létrejött valószínűségszámítással. Ennek hangsúlyozása azért különösen fontos, mivel a valószínűség empirikus meghatározásának feladata az irodalomban nagyon sok esetben teljesen háttérbe szorul a másik két vizsgálati szempont mögött.

Ezzel a hibával összefüggésben az irodalomban az a szóhasználat honosodott meg, hogy a valószínűség empirikus elméleteinek vizsgálatát a „valószínűség interpretációi" kifejezéssel adják vissza. A valószínűség elemzésével foglalkozó munkák jelentős része ilyen címet visel, és tartalmilag is a kiindulópontot legtöbbször a valószínűségszámítás matematikai elmélete jelenti, a feladat pedig abban áll, hogy ehhez a formális elmélethez fizikai szemantikát rendeljünk. Mi ezt a megközelítést a fentebb elmondottak fényében nem fogadhatjuk el; ugyanakkor nem kívánunk elszakadni a hagyományos szóhasználattól sem. Ezért a továbbiakban mi is a „valószínűség interpretációi" kifejezést fogjuk használni (valójában a könyv címétől fogva ezt használjuk), mégpedig abban a *tág értelemben*, amely magában foglalja a valószínűség fizikai elméletének valamint a köznyelvi fogalomhoz illetve a kolmogorovi elmélethez vett viszonyának (azaz a szűk értelemben vett interpretációnak) vizsgálatát. A *szűk* *értelemben* vett interpretáció helyett innentől fogva igyekszünk a „szemantika” kifejezést használni.

Ebben a fejezetben a következőképpen fogunk haladni. Mivel a szemantika alapjául szolgáló valószínűségszámítás alapjait az előző fejezetben ismertettük, ezért ezt a fejezetet az explikáció tárgyával, azaz a *common sense* valószínűségfogalom rövid vizsgálatával kezdjük. Ezek után bemutatjuk a valószínűség interpretációinak Wesley Salmontól származó és a szakmában szinte egyöntetűen elfogadott kritériumait, és megmutatjuk, hogy e kritériumok az interpretáció természetére vonatkozó milyen előfeltevésekkel élnek. Majd rátérünk arra a kérdésre, hogy mi is az a metafizikai entitás, amely valószínűséggel rendelkezhet – vajon események, tények vagy kijelentések-e a valószínűség hordozói? Ezt követően feltesszük a könyv legfontosabb kérdését: mit is értünk egy esemény vagy kijelentés valószínűségén – a választ azonban a soron következő fejezetekre bízzuk. Itt csupán a valószínűségi interpretációk tág térképét igyekszünk megrajzolni, és a térképen megjelölni azokat a metafizikai „tartományokat”, amelyeken az interpretációk egyáltalán elhelyezkedhetnek. Mivel az irodalomban szokásban van a valószínűséggel kapcsolatos filozófiai problémákat úgy elemezni, hogy közben a valószínűség fogalmát nem interpretálják vagy csak intuitív értelemben használják, ezért a fejezetet két példával, a Simpson-paradoxon és a reichenbachi közös ok példájával zárjuk, és a példákon megmutatjuk, hogy a felvetett filozófiai kérdések nem válaszolhatók meg azelőtt, mielőtt a valószínűség fogalmát tisztáztuk volna.

3.1. A *common sense* valószínűség

„Mi a valószínűsége annak, hogy elérem a vonatot, hogy elkapom a betegséget, hogy az eső elmossa a kirándulást?” „Mi a valószínűsége, hogy az evolúcióelmélet helyes?” „Mi a valószínűsége, hogy van élet a Földön kívül is, hogy létezik Isten?” „A bizonyítékok fényében mi a valószínűsége, hogy ő volt a tettes?” „Mi a valószínűsége, hogy nyerek a lottón, hogy hatost dobok a kockával?” – ilyen és ehhez hasonló fordulatokban használja a köznyelv a valószínűség fogalmát. Mik ennek a köznyelvi használatnak az általános szabályai?

A köznyelvi felfogás szerint ami valószínű, az sem nem lehetetlen, sem nem biztos.[1] Ha az urnából valószínűleg fekete golyót fogunk húzni, akkor lehetséges, hogy feketét húzunk, de nem biztos. A lehetőség itt empirikus lehetőség, és nem pusztán logikai. Az, hogy lehetséges, hogy fekete golyót húzok, nemcsak azt implikálja, hogy az előzetes feltételek (az urna összetétele, a húzás módja) nincs logikai ellentmondásban az eredménnyel (a fekete golyó húzásával), hanem azt is, hogy empirikusan sem zárják ki azt (ahogy mondjuk kizárják, hogy egy nyulat húzzak ki az urnából). Ezek a kezdeti feltételek implicit módon benne foglaltatnak minden köznyelvi valószínűségi kijelentésben, akár explicitté képes tenni őket a beszélő, akár nem.

[1] A *common sense* valószínűségfogalom elemzésében Max Black-nek (1967) a *Macmillan Encyclopedia of Philosophy*-beli *Probability* szócikkére támaszkodom.

A „valószínűleg" kifejezés nem metanyelvi kifejezés, hanem a hétköznapi tárgynyelv része, így „Az urnából valószínűleg fekete golyót fogunk húzni" mondat a világ tényeire vonatkozik. Normál esetben az, hogy „Az urnából valószínűleg fekete golyót fogunk húzni" implikálja, hogy „Az urnából fekete golyót fogunk húzni". Ezért normál körülmények között abszurd azt állítani – mintegy Moore paradoxonjának mintájára –, hogy „Az urnából valószínűleg fekete golyót fogunk húzni, de nem fogunk feketét húzni". Vagyis „Az urnából valószínűleg fekete golyót fogunk húzni" mondat állításával a beszélő elköteleződik „Az urnából fekete golyót fogunk húzni" igazsága mellett, megengedve ugyanakkor, hogy az előzetes feltételek esetleg nem elegendően erősek az eredmény létrehozásához. Ha a beszélő fehér golyót húz, kijelentheti: „Tévedtem", ugyanakkor reflektálhat is a mondat megengedő jellegére: „Fehéret húztam, mindenesetre a fekete valószínűbb volt". Ennélfogva a „valószínű" kijelentés szerepe kettős: egyrészt valószínűséget tulajdonít egy eseménynek, másrészt meg is jósolja egy esemény bekövetkezését. Ezzel függ össze, hogy a valószínűségi kijelentések szorosan kötődnek a valószínűségi becslések gyakorlati szerepéhez a cselekvésekben. Abszurd tehát azt állítani, hogy „Ez felettébb valószínű, de nem hiszem, hogy így lesz" vagy, hogy „Ez felettébb valószínű, de ezt inkább ne vedd figyelembe". Más szóval a bizonytalan helyzetekben való döntés igazodik a valószínűségi becslésekhez: követi a csaknem bizonyost, és kerüli a nagyon valószínűtlent.

Amíg a köznyelvi használatban a „valószínű" jelzői és a „valószínűleg" határozói fordulat egyfajta elkötelezettséggel jár az állított mondat igazsága vagy a esemény bekövetkezése mellett („Valószínű, hogy feketét húzunk."), addig a „valószínű, hogy" és még inkább az „annak a valószínűsége, hogy" névszói kifejezés lazít ezen az elköteleződésen, és egyfajta episztemológiai távolságot tart az állított mondattól („A fekete golyó húzásának valószínűsége nagy.").

A valószínűség szoros fogalmi rokonságban áll a lehetőséggel. Mind a „valószínű", mind a „lehetséges" kifejezés fokozható. – „Mennyire valószínű?" – „Nagyon valószínű." – „Rendkívül valószínű." „Mennyire lehetséges?" – „Igencsak lehetséges." – „Csaknem biztos." A „biztos"-t a köznyelv nem tekinti a „valószínű" speciális esetének. A szoros fogalmi rokonság hátterében az a természetes meggyőződés áll, hogy a valószínűség a kezdeti feltételekre vonatkoztatott empirikus lehetőség, egyfajta távolság a bizonyostól. Lyuk van a táskámon, amelyben a lakáskulcsomat tartom. A *common sense* szerint „lehetséges", hogy a kulcs kiessen a lyukon, de nem biztos, mert a kulcs lehet éppen a táska másik felében is. Ha a táskámon egy másik lyuk is támad, akkor a kulcs kap egy extra lehetőséget is, hogy kiessen, és így nagyobb lesz a valószínűsége annak, hogy kiesik. A példa mögötti racionálé világos: a kezdeti feltételek (a táska) nem határozzák meg teljes mértékben a lehetőségeket a végeredmény (a kulcs kiesése) tekintetében. Amennyiben inkább blokkolják a lehetséges utakat a siker felé, annyiban a kimenet inkább valószínűtlen, amennyiben szabad utat engednek a sikernek, annyiban a kimenet valószínű. Vagyis a kezdeti feltételek egyfajta tendenciával rendelkeznek, hogy a végeredményt realizálják. Ennek a tendenciának a pontos mibenléte a köznyelvi használatból nem derül ki, annyi azonban megállapítható, hogy a tendencia a szóban forgó szituáció objektív tulajdonsága.

Az, hogy lehetséges, hogy a kulcs kiessen a lyukon, empirikusan igazolható: néha elvesztem a kulcsomat. Ha az esetek többségében kiesik a kulcs, akkor valószínűnek fogom tartani, hogy a kulcs kiessen, ha szinte sohasem esik ki, akkor csaknem lehetetlennek. A *common sense* elképzelés azonban nem azonosítja a valószínűséget a relatív gyakorisággal, hanem az utóbbit egyszerűen az előbbi jelének veszi. Az empirikus verifikáció mellett a valószínűségi kijelentéseket igazolhatjuk másfelől pusztán introspekció révén is. – „Mi a valószínűsége annak, hogy egy exobolygón életet találnak?" – „Kisebb, mint egy százalék." Ilyenkor egyszerűen reflektálhatunk arra az meggyőződésünkre, hogy az exobolygókon az életet kevésbé tartjuk valószínűnek, mint mondjuk egy dupla hatos dobását két kockával. A valószínűségi jóslatokba némi uniformitás is belejátszik. Azt, hogy két hasonló összetételű urnából fekete golyót húzunk, azonos valószínűségűnek tartunk. Ez az elv egyszerűen az „azonos okokból azonos okozatok" maxima általánosítása valószínűségi eloszlásokra.

Összefoglalva a *common sense* valószínűséget az alábbi három tulajdonság jellemzi: (i) valószínűséggel szinguláris események és általános tényállások egyaránt rendelkezhetnek, (ii) valószínűségi kijelentéseink a külvilágra vonatkoznak és nem a beszélő hitállapotára, (iii) a relatív gyakoriság nem azonos a valószínűséggel, legfeljebb igazolhatja azt.

3.2. A valószínűség interpretációjának kritériumai

Amint a bevezetőben már említettük, a valószínűség interpretációjának feladatát az irodalomban legtöbbször abban látják, hogy egy formális rendszerhez, a valószínűségszámításhoz valamilyen fizikai szemantikát adjanak meg. Adva van tehát egy (Ω, Σ, p) valószínűségi mértéktér; kérdés, hogyan interpretáljuk a benne szereplő mennyiségeket. Ehhez konkrétan a következő kérdésekre kell választ kapnunk:

1. Mi jelent a valószínűségelmélet eseménystruktúrája? Hogyan interpretáljuk az eseményeket, azaz mi az Ω és Σ halmaz jelentése?

2. Hogyan értelmezzük a $p : \Sigma \to [0, 1]$ hozzárendelést? Mit értünk azon, hogy egy a esemény $p(a)$ valószínűsége egyenlő x-szel?

Ezt a formális kiindulópontot választva valamiképpen azt is garantálni kell ezek után, hogy az így megadott szemantika „helyes" lesz: valóban a keresett valószínűséghez vezet el. Ehhez valamilyen kritériumok szükségesek. Ilyen *adekvátsági kritérumokat* fogalmazott meg Wesley Salmon (1966), és a szakma többnyire az ő kritériumait használja a valószínűségi interpretációk megítélésére. A kritériumok a következők:

> „*Megengedhetőség* [*Admissibility*]: Azt mondjuk, hogy egy formális rendszer interpretációja megengedhető, ha a primitív terminusokhoz rendelt jelentések a formális axiómákat, és következésképpen minden tételt igaz kijelentésekké transzformálnak. A valószínűségfogalommal szemben támasztott alapvető követelmény, hogy kielégítse azokat a matematikai relációkat, amelyeket a valószínűségi kalkulus határoz meg.

> *Kideríthetőség* [*Ascertainability*]: Ez a kritérium azt kívánja, hogy álljon rendelkezésre valamilyen módszer, legalább elvben, amellyel kideríthetjük a valószínűség értékeit. Mindez pusztán azt a tényt fejezi ki, hogy a valószínűség fogalma hasztalan volna, ha lehetetlen volna elvben kideríteni hogy mik is a valószínűségek ...
>
> *Alkalmazhatóság* [*Applicability*]: Ennek a kritériumnak az erejét legjobban Bishop Butler híres aforizmája fejezi ki: 'A valószínűség az élet kalauza.' Megkerülhetetlen tény tehát, hogy olyan valószínűségfogalmat keresünk, amelynek gyakorlati prediktív ereje van." (63-64. o.)

A kritériumok értelmezést kívánnak. A megengedhetőség kritériuma a fenti megfogalmazásban egy kissé homályos, ugyanis a matematikai relációkat kielégíteni csak matematikai fogalmak tudnak, nem pedig fizikai eljárások vagy jelentések. Másfelől, ha egy valószínűségi kalkulusban megadjuk az alapfogalmak fizikai jelentését, akkor azok *a fortiori* kielégítik a formális követelményeket, lévén azok interpretációi. A kideríthetőség kritériuma valójában pusztán a formális rendszerek interpretációját fogalmazza újra: interpretálni annyi, mint megadni az elmélet terminusainak fizikai jelentését. A „legalább elvben" kitétel a megfogalmazásban a szigorú verifikacionizmussal szemben tesz engedményeket. Alkalmazhatóság kritériumával Salmon a valószínűség interpretációjára terheli azt a feladatot is, hogy szolgáljon alapul a valószínűség fogalmának az indukció, a tudományos következtetés, a racionális hit témakörében való használatához vagy éppen a tudományban való paradigmatikus használathoz.

A kritériumok értelme valójában akkor válik világossá, ha elhagyjuk azt a formális kiindulópontot, amelyet a valószínűség szokásos elemzései követnek, vagyis a valószínűség interpretációjának problémáját nem az (Ω, Σ, p) valószínűségi mértéktér szemantikájának a megadásában látjuk, hanem – ahogy azt már a bevezetőben említettük – a valószínűség egy empirikus elméletének felállításában. Ha a valószínűség interpretációjának feladatára így tekintünk, akkor rögtön értelmet nyer például Salmon megengedhetőségi kritériuma. *Miután* ugyanis rendelkezésünkre áll egy elmélet, értelmes lesz feltenni azt a kérdést, hogy az elmélet formális része milyen viszonyban áll a kolmogorovi valószínűségelmélettel. Egy példát említve, ha a valószínűség relatív gyakoriság végtelen sorozatokban, akkor feltehetjük azt a kérdést, hogy ezen sorozatokban értelmezett relatív gyakoriságok vajon modellezik-e a valószínűségeket mint mértékeket. Ha az derül ki, hogy a kérdésre a válasz nemleges, abból persze még korántsem következik, hogy a valószínűség relatív gyakoriság elmélete nem megengedhető, ahogy azt Salmon első kritériuma sugallja. Ebből a tényből mindössze annyi következik, hogy a fizikai valószínűség *más*, mint amit a matematikai valószínűségelmélet alapján várnánk. Az eltérés okán természetesen el lehet töprengeni: vajon nem idealizálja-e túl a mértékelmélet a valószínűség fogalmát, vajon a fogalomnak nem más jegyeit tartja-e szem előtt a matematikai és a fizikai elmélet, stb. A lényeg azonban az, hogy amennyiben a fizika elmélet formális része nem modellje a klasszikus valószínűségszámításnak, még nem szükséges ez előbbit elvetnünk, hiszen nincsenek a priori előírások arra nézve, hogy a tényleges valószínűségeknek éppen a kolmogorovi elmélet szerint kellene viselkedniük.

Valójában az interpretációnak ezt az általunk javasolt formáját sugallja a második, kideríthetőségi kritérium is. Az, hogy az interpretáció kideríthető, egyszerűen annyit jelent, hogy egy fizikai elmélet áll rendelkezésünkre, amely fizikailag értelmezi a benne szereplő valószínűség terminust.

Ami Salmon harmadik, megengedhetőségi kritériumát illeti, ez megítélésünk szerint egyszerűen az explikáció követelményét fogalmazza meg. Egy fizikai elméletet csak akkor tarthatunk a valószínűség, nem pedig valamely más köznyelvi fogalom fizikai elméletének, ha az elmélet azokat az előzetes jegyeket explikálja, amelyekkel a valószínűség a köznapi használatban rendelkezik. Salmon ezek közé sorolja a prediktív erőt is, így egy olyan fizikai elméletet, amely alapján események jövőbeli bekövetkezésére nem lehet jóslásokat adni, nem minősülhet a valószínűség helyes elméletének. Nehéz azonban belátni, hogy a valószínűség fogalmának milyen extra prediktív erővel kellene rendelkeznie azon túl, amellyel minden egyéb fizikai fogalom rendelkezik. Ha a hőmérséklet egy valóban létező fizikai fogalom, akkor rengeteg predikcióra nyújt lehetőséget – kezdve onnan, hogy mi történik, ha télen nem veszek kabátot, egészen addig, hogy mi történik egy gáztartályban, ha növelem a nyomást. A hőmérséklet tehát „kalauz az életben" – mind a hétköznapiban, mind a tudományosban. Ezen a fizikai fogalmakra jellemző általános prediktív erőn túlmenően a valószínűség sem rendelkezik semmiféle extra prediktív erővel. Ha a valószínűség fizika fogalom, és így beleillik egy fizikai elméletbe, akkor a fizikai elmélet és alkalmazási köre meg fogja határozni, hogy a valószínűség milyen predikciókra használható. Az természetesen egy további kérdés, hogy a valószínűség fogalma az indukció, a racionális döntés, a tudományos elméletek konfirmációjának céljait szolgáló predikcióra is alkalmas lesz-e. Mindenesetre ezek a szempontok nem tartozik bele a fogalom előzetes használatába, így nem is befolyásolhatják (bár történetileg nem ritkán befolyásolták) a valószínűség fizikai elméleteinek kialakulását.

Összefoglalva, Salmon három adekvátsági kritériuma a helyes értelmezés szerint a következő programot szorgalmazza. 1. Alkossuk meg a valószínűség fizikai elméletét (kideríthetőségi kritérium); 2. vizsgáljuk meg az elmélet formális részének viszonyát a valószínűségszámításhoz (megengedhetőségi kritérium); 3. végül ellenőrizzük az elméletben szereplő valószínűségfogalom viszonyát a köznyelvi valószínűségfogalomhoz (alkalmazhatósági kritérium). Mindez egybevág azzal, amit a bevezetőben a valószínűség tág értelemben vett interpretációjának neveztünk. Így értelmezve a kritériumok helyesek, és a továbbiakban mi is használni fogjuk őket. Most azonban térjünk át annak a kérdésnek a tisztázásához, hogy minek is van valószínűsége.

3.3. Az események és kijelentések metafizikája

A valószínűségelmélet interpretációinak alapfogalma vagy az *esemény* vagy a *kijelentés*. Az események és a kijelentések ontológiájának elemzése valamint viszonyuknak tisztázása mély metafizikai kérdés, amelynek részleteitől itt most eltekintünk, és csak a legszükségesebbeket foglaljuk össze. (A részletesebb tárgyaláshoz ld. (Schneider 2005) vagy magyarul (Tőzsér 2009)).

Az események ontológiájának vizsgálata viszonylag új területe a metafizikának, mondjuk a fizikai tárgyakéhoz képest, ami annál is inkább meglepő, mivel a józan ész a tárgyakhoz hasonlóan realista az eseményeket illetően is: az eseményeket meg tudjuk különböztetni egymástól, referálni tudunk rájuk, számlálni tudjuk őket, stb. Az események a fizikai tárgyakhoz hasonlóan partikulárék. A partikuláris vagy *szinguláris eseményeket* (*token events*) meg kell különböztetni az általános eseményektől vagy *eseménytípusoktól* (*type events*). Az eseménytípusok természetét illetően a két szembenálló metafizikai hagyománynak megfelelően két elképzelés létezik. A realista elképzelés szerint az eseménytípus egy *tulajdonság* (univerzálé), a nominalista hagyomány szerint az eseménytípus szinguláris *események osztálya*. Hogy ebbe az eseményosztályba azután csak aktuális szinguláris események tartoznak-e bele vagy lehetséges események is, az a nominalista álláspontón belül további változatokat eredményez. Szinguláris esemény tehát az a konkrét hatos dobás, amelyet ezzel a kockával ma délben végrehajtottam, eseménytípus pedig a realista értelemben a „hatos dobás” tulajdonsága, nominalista értelmezésben a szinguláris hatos dobások osztálya. A szinguláris események és az eseménytípusok közötti viszonyra a szokásos metafizikai terminus a realista tradícióban az *instanciálás*: a ma déli hatos dobás a „hatos dobás” tulajdonságának instanciálódása, a nominalista hagyományban pedig az *eleme*: a déli hatos dobás a hatos dobások osztályának eleme. Amint majd látni fogjuk, a realista-nominalista szembenállás nem jelenik meg a valószínűségi interpretációk közötti vitában, azonban a valószínűség minden interpretációjának határozottan állást kell foglalnia abban a kérdésben, hogy a valószínűséget szinguláris eseményhez vagy eseménytípushoz rendeli-e.

Az események metafizikájának több rivális elmélete létezik. Az egyik elmélet szerint a szinguláris események téridőtartományok, vagyis a téridő részhalmazai. A ma délben végrehajtott kockadobás a térnek az a régiója, amelyet a kocka (esetleg a kezem, vagy az asztal) elfoglalt, és az időnek az a szakasza, amely a kocka eldobásától a kocka landolásáig eltelt. E felfogás szerint az események primitív, ontológiai szerkezet nélküli entitások, amelyeket pusztán relációs, azaz a többi eseményhez fűződő kapcsolataik révén azonosítunk. Ez a szinguláris kockadobás abban különbözik egy másik kockadobástól, hogy a téridő másik helyén történt. De vajon nem történhet-e a téridő egyazon régiójában két esemény? Ugyanaz-e az a két esemény, hogy a kockával hatost dobtam, és az, hogy páros számot dobtam – ha történetesen délben hatost és így páros számot dobtam? Ugyanaz-e hangosan beszélni és zavarni a szomszédot? Vagy Davidson (1980, 178. o.) példájával: ha egy golyó melegszik, és közben forog is, akkor ez két esemény vagy egy? Ha meg szeretnénk engedni, hogy a téridő egyazon régiójában több esemény is helyet kaphasson, akkor az eseményeket másképpen kell individuálnunk, nem pusztán a téridőbeli *locus*-uk alapján.

Ezt teszi a szinguláris eseményeknek egy másik elmélete, amely az eseményeket az oksági folyamatokban betöltött szerepük alapján azonosítja. Az események ebben az elméletben is belső szerkezet nélküliek, azonban nem kötődnek kizárólagosan téridőtartományokhoz. A hatos és a páros dobás különböző események, mivel más oksági hálóba illenek bele: néhány társasjátékban hatos dobás esetén még egyet dobhatok, páros dobás esetén azonban nem. A téridő elmélettel szemben az oksági elmélet többletfeladata, hogy előzetesen meghatározza az okság fogalmát az eseményfogalom felhasználása nélkül. Ez azonban komoly kihívásnak bizonyul, hiszen az okok és okozatok maguk is események.

Egy harmadik, Jaegwon Kim (1998) nevéhez fűződő elmélet még finomabban individuálja az eseményeket. Kim szerint az események összetett entitások, amelyeket egy szubsztancia, egy tulajdonság és egy időpont hármas struktúrája karakterizál. A ma délben történt kockadobás tehát három komponensből áll: a kockából mint szubsztanciából, az eldobásból mint tulajdonságból (eseménytípusból), és a ma déli időpontból. Ha tehát nem a kockát dobtam volna el, vagy nem ma délben, hanem tegnap délben, vagy ha nem eldobtam volna, hanem a táskámba tettem volna – mindhárom esetben más eseményt kapnék. Az esemény azonban nem pusztán egy rendezett hármas: a kockadobás eseménye csak akkor következik be, ha a kocka ma délben *instanciálja* (exemplifikálja) az „eldobva lenni" tulajdonságát.

Kim elméletének egyebek közt azt szokás felróni, hogy elköteleződik a fizikai tárgyak léte mellett, vagyis feltételezi, hogy minden eseményhez egy szubsztancia tartozik, ahogy a hatos dobáshoz a kocka. Ezt a feltételezést küszöböli ki David Lewis elmélete. Lewis szerint a szinguláris események a fenti téridőrégió-elmélethez hasonlóan téridőrégiók. Individuációjukhoz azonban nem elegendő megadni pusztán azt a téridőrégiót, amelyben történnek, hanem ezen felül meg kell adnunk a téridőrégiók egy osztályát is, amelynek az illető esemény eleme. Ez az osztály hivatott képviselni Lewis nominalista metafizikájában azt a tulajdonságot, amelyben az osztály elemei közösek.[2] A golyó felmelegedése és forgása tehát azért különböző események Lewis szerint, mert az egyiknek megfelelő eseménytípus különbözik a másiknak megfelelő eseménytípustól: a világban (vagy valamelyik lehetséges világban) vannak melegedő, de nem forgó gömbök, és fordítva.

A fenti négy elmélet többek között abban tér el egymástól, hogy különböző finomságban tagolja az eseményeket. A skála egyik végén a téridőrégió áll, amely szerint azzal, hogy kijelöljük egy esemény helyét és idejét, és kikötjük, hogy az események az aktuális világ részei, már egyértelműen individuáljuk az eseményeket: az aktuális világban *akkor* és *ott*, nem történhet más, mint ami ténylegesen történik. A hatos dobás és jobb kézzel való dobás legfeljebb egy tulajdonsága lehet annak a szinguláris kockadobásnak, amely ma délben történt az asztalomon.

A skála másik végén Kim és Lewis elmélete áll, amely igen finoman tagolja az eseményeket. A hatos dobás, és a páros szám dobása két esemény, mert bennük két különböző tulajdonság instanciálódik a téridő adott pontján. De nem szaporodnak-e így el túlságosan a déli kockadobással kapcsolatos események? Azonos-e például az az két esemény, hogy ma délben eldobtam a kockát, valamint az, hogy ma délben *jobb kézzel* eldobtam a kockát – ha történetesen így volt? Kim igyekszik elejét venni az események ilyetén proliferációjának, és azt állítja, hogy a déli kockadobás és a déli jobb kezes kockadobás – szemben a hatos és a

[2] Hogy ebbe az osztályba Lewis azután lehetséges világok téridőtartományait is beleveszi, az itt most nem lényeges.

páros dobással – *egy* esemény, mivel jobb kezes kockadobás esetében nem a kocka instanciálja az „jobb kézzel eldobva lenni" tulajdonságot, hanem a kockadobás *eseménye* (amely maga az eldobás instanciálódása a kocka által) instanciálja a "jobb kézzel való" (határozói) tulajdonságot.

Az egyre finomabban strukturált eseményfogalom azonban átvezet bennünket egy másik metafizikai kategóriához, a *tényekhez.* A standard meghatározás szerint a tény egy tulajdonság instanciálása egy konkrét partikuláré által. Mint ilyen a tény absztrakt és atemporális entitás. Az, hogy ma délben eldobtam egy kockát, egy tény, amelyben a kocka mint konkrét partikuláré instanciálja az „eldobva lenni" tulajdonságát. Az, hogy ma délben hatost dobtam, egy másik tény, mert a kocka itt a „hatost dobni" tulajdonságát instanciálja. Vagyis ha ma délben hatos dobtam a kockával, akkor mindkét tény *fennáll*, az is, hogy eldobtam a kockát, és az is, hogy hatost dobtam. Mint látható az események oksági elmélete és még inkább Kim tulajdonság-exemplifikációs elmélete szerint ennek a két ténynek két esemény felel meg: a hatos dobás és a páros dobás eseménye,[3], míg az események téridőrégió elmélete szerint csak egy: az az esemény, amelyet az adott téridőrégió azonosít.

Az eseményfogalom metafizikai skálája tehát a durván tagolt téridőeseményektől a finoman tagolt tényekig terjed. A továbbiakban ezért szem előtt kell majd tartanunk, hogy esemény alatt milyen metafizikai entitást is értünk.

Folytatva a szokásos metafizikai *parlance*-t: a tények *igazságalkotók* (*truth-makers*). Mit jelent ez? Vagyis melyek azok a metafizikai entitások, amelyek képesek igazak és hamisak lenni attól függően, hogy bizonyos tények fennállnak-e vagy sem? A standard válasz a kérdésre: az igazság hordozói (*truth-bearers*) *propozíciók.* A tények metafizika bevezetését éppen az teszi indokoltá, hogy velük megmagyarázzuk, miért tartunk bizonyos propozíciókat igaznak. A propozíciók bevezetése azonban csak további metafizikai feladatokat róna ránk: el kellene számolnunk a propozíciók mint absztrakt partikulárék metafizikai természetéről. Éppen ezért mi a továbbiakban igazsághordozóknak nem a metafizikailag terhelt propozíciókat fogjuk tekinteni, hanem a neutrálisabb *kijelentéseket.* Kijelentés alatt itt nem egy konkrét téridőbeli megnyilatkozást értünk, hanem egyszerűen egy formális nyelv elemét. A tények absztrakt struktúráját tehát egy formális nyelv jól definiált kijelentései (formulái) reprezentálják, örökölve ezáltal a tények finoman tagoltságát.

A kijelentések és a tények közötti viszonyt a *szemantika* teremti meg: a nyelv egy bizonyos formulája akkor és csak akkor igaz, ha egy bizonyos tény fennáll – vagy visszatérve a finoman tagolt események nyelvére: ha egy esemény bekövetkezik. A kijelentése és a tények közötti szemantikai viszony természete igen nehéz metafizikai kérdés. (Ehhez ld. Szabó 2003) Hasonlóan nehéz metafizikai kérdés, hogy vajon a tények szerkezete „hasonlít-e" a formális nyelv szerkezetére. Vannak, akik szerint az összetett formulákat összetett tények teszik igazzá; vannak, akik szerint minden tény atomi; és van olyan is, Davidson, aki szerint csak egy tény létezik, az Igaz. Ismét csak maradjunk a metafizikailag ökonomikus válto-

[3] Éppen ez az, amiért Kim elméletét azzal szokás vádolni, hogy az nem is eseményeket individuál, hanem tényeket.

zatnál, és anélkül, hogy elköteleződnénk a konjunktív, diszjunktív stb. tények létezésének kérdésében, mondjuk egyszerűen azt, hogy egy nyelv atomi kijelentéseinek a világ tényeire vonatkozó szemantikája a konnektívumok segítségével egyben az összetett formulákat is szemantikával látja el.

Két metafizikai entitásunk van tehát: az immanens, temporális és (legalábbis a téridőrégió-elmélet szerint) durván tagolt esemény illetve az absztrakt, atemporális és finoman tagolt tényeket reprezentáló kijelentés. Melyik legyen a kettő közül a valószínűség hordozója? Valószínűsége a ma déli kockadobásnak van, vagy annak a kijelentésnek, hogy „Ma délben a kockával hatost dobtam (dobok)"?[4]

Mellor (2005, 9. o.) az utóbbira szavaz, mégpedig az alábbi egyszerű érv alapján. Hogyan lehetne a ma déli hatos dobás *eseménynek* egy hatod a valószínűsége, ha történetesen ma délben ötöst dobok? A hatos dobás szinguláris eseménye ekkor egészen egyszerűen nem következik be, ezért nem lehet a valószínűség hordozója sem, hacsak nem köteleződöm el lehetséges események létezése mellett. Ezzel szemben az a kijelentés, hogy „Ma délben a kockával hatost dobok" továbbra is létezik mint egy formális rendszer része, és így valószínűséggel is rendelkezhet. A valószínűség tehát kijelentésekhez tartozik.

Az érv huszáros, és az ember igen csak csodálkozna, ha ilyen egyszerű érvvel diszkreditálni lehetne a két metafizikai kategória közül az egyiket. És nem is lehet. Mielőtt nem nyilatkoztunk arról, hogy mit értünk valószínűségen, egyáltalán nem zárható ki, hogy a valószínűség az aktuálisan megvalósuló ötös dobáshoz tartozzon. Az természetesen kiderülhet, hogy a valószínűség fogalma csak eseménytípusokra vagy tényekre illetve kijelentésekre értelmezhető, de ez a belátás semmi esetre sem születhet meg előbb, mint hogy megmondtuk, mit értünk valószínűségen. Az alábbiakban tehát nyitva hagyjuk a kérdést, hogy vajon a valószínűség eseményekhez vagy kijelentésekhez tartozik-e, és párhuzamosan megvizsgáljuk mindkét lehetőséget. Mielőtt azonban a valószínűséget eseményekhez vagy kijelentésekhez rendeljük, meg kell vizsgálnunk, hogy az eseményekhez és a kijelentések milyen struktúrával rendelkeznek. Kezdjük az események vizsgálatával.

Induljunk ki ismét csak egy konkrét hatos dobásból, és tekintsünk el az esemény individuációját érintő metafizikai kérdésektől. A hatos páros szám, és osztható hárommal, vagyis ez a szinguláris esemény a „páros dobás" és a „hárommal osztható szám dobása" eseménytípusnak egyaránt instanciája. Ha legközelebb hatost dobunk, akkor ebben a másik szinguláris eseményben is instanciálódik mindkét eseménytípus. De nemcsak ez a két eseménytípus instanciálódik a szinguláris dobásokban, hanem egyúttal a „hatos dobás" eseménytípus is. Ugyanakkor „páros dobás" és a „hárommal osztható szám dobása" eseménytípusok sehogy máshogy nem instanciálódhatnak együtt, csak úgy, ha hatost dobok. Röviden a „hatos dobás" eseménytípusa akkor és csak akkor instanciálódik, ha „páros dobás" és a „hárommal osztható szám dobása" eseménytípusai is. Hasonló viszony áll fent a többi eseménytípus kö-

[4] Valójában van még egy harmadik opció is: a valószínűség a kockához, vagyis a fizikai tárgyhoz tartozik, amely Kim szerint az eseményt konstituálja. Ezt a lehetőséget Mellor *propensity*-nézete kapcsán fogjuk megvizsgálni. (Ld. 136. oldal.)

zött is: „páros dobás” és a „négynél kisebb szám dobása” akkor és csak akkor instanciálódik, ha a „kettes szám dobása” instanciálódik.

Hasonló a helyzet a negációval és a diszjunkcióval. Egy szinguláris esemény akkor és csak akkor instanciája a „páros dobás” eseménytípusának, ha *nem* instanciája a „páratlan dobás” esememénytípusának. Egy szinguláris esemény akkor és csak akkor instanciája a „hárommal osztható szám” eseménytípusának, ha vagy a „hármas dobás” vagy a „hatos dobás” eseménytípusának instanciája. A különböző eseménytípusok instanciálódása között tehát logikai összefüggések állnak fent. Természetesen a „eseménytípus instanciálódása” realista kifejezés helyett – a metafizikai meggyőződéstől függően – mindenütt élhetnénk azzal a nominalista fordulattal is, hogy a szóban forgó szinguláris esemény egy adott eseményosztályba *tartozik*, vagy még egyszerűbben egy bizonyos típusú szinguláris esemény *bekövetkezik*. Mindezek után az események bekövetkezésének fenti „logikáját” kifejező struktúrát (immár a nominalista beszédmódban) a következő módon jellemezhetjük:

(i) Bármely két a és b eseményosztályhoz egyértelműen létezik egy $a \wedge b$ *metszet eseményosztály*, amelybe egy szinguláris esemény pontosan akkor tartozik bele, ha mindkét eseményosztályba beletartozik.

(ii) Bármely két a és b eseményosztályhoz egyértelműen létezik egy $a \vee b$ *unió eseményosztály*, amelybe egy szinguláris esemény pontosan akkor *nem* tartozik bele, ha egyik eseményosztályba sem tartozik bele.

(iii) Bármely a eseményosztályhoz egyértelműen létezik egy $a^{\perp}$ *komplementer eseményosztály*, amelybe egy szinguláris esemény pontosan akkor tartozik bele, ha az a eseményosztályba *nem* tartozik bele.

(iv) Létezik a 0 *lehetetlen eseményosztály*, amelybe egyetlen szinguláris esemény sem tarozik bele.

(v) Létezik az 1 *biztos eseményosztály*, amelybe minden szinguláris esemény beletartozik.

Ez utóbbi két „mesterséges” eseménytípus feltevése egyszerűen annak a két tapasztalatunknak a kifejezésére szolgál, hogy a „hatos dobás” és a „nem hatos szám dobása” eseménytípusai együtt sohasem instanciálódnak, a kettő közül azonban valamelyik – feltéve, hogy a kockadobás megtörtént – mindig instanciálódik. Vagyis létezik egy sohasem instanciálódó és egy mindig instanciálódó eseménytípus.

A fenti öt kikötésnek (illetve az elvárásokba implicite beleértett egyéb tulajdonságoknak: disztributivitás, De Morgan-szabályok stb.) megfelelő struktúra, vagyis a metszetre, az unióra és a komplementer műveletre zárt algebrai struktúra azonban éppen a Boole-algebra, vagy annak megszámlálható végtelen műveletre általánosított formája, a σ-algebra. A Boole-algebra matematikai struktúrája tehát az eseményektípusokat modellezi az instanciálódás/eseményosztályba tartozás „logikája” szempontjából.

A fenti kikötések alapján úgy tűnhet, mintha implicit módon állást foglaltunk volna abban a kérdésben is, hogy valószínűsége szinguláris eseményeknek lehet-e vagy eseményosztályoknak. A Boole-algebra műveletek ugyanis eseményosztályok között vannak értelmezve, és így úgy tűnhet, hogy a valószínűség is csak eseménytípusokhoz tartozhat. Szinguláris események között a fenti műveletek nem értelmezhetők, hiszen mi is volna az $a^{\perp}$ komplementer szinguláris esemény értelme, hogyan is nézne ki egy szinguláris nem-hatos dobás?

Ez azonban nincs így, ha észben tartjuk, hogy a Kim-féle és a Lewis-féle szinguláris események belső szerkezettel rendelkeznek. Kim szerint egy szinguláris eseményt az individuál, hogy milyen tulajdonságot instanciál, Lewis szerint pedig egy szinguláris eseményt mint téridőrégiót az individuál, hogy a téridőrégiók mely osztályának, tehát mely eseményosztály tagja. Vagyis ha a valószínűség az eseménytípusokon értelmezve van, akkor közvetve a szinguláris eseményeken is: ha a „hatos dobás" tulajdonságnak (a hatos dobások eseményosztályának) van valószínűsége, akkor a szinguláris hatos dobások instanciálhatják egyben ezt a (másodlagos) tulajdonságot is: a szinguláris hatos dobásnak egy hatod a valószínűsége, *mivel* a „hatos dobás" tulajdonságnak annyi. Ha pedig szinguláris események alatt nem a Kim-féle vagy a Lewis-féle belső szerkezettel rendelkező szinguláris eseményeket értjük, hanem egyszerűen a téridőrégiókat, a Boole-algebra-műveleteket az események között akkor is értelmezhetjük mint a téridőrégiók halmazelméleti metszeteit, unióit és komplementereit. Vagyis a kérdés, hogy a valószínűség szinguláris eseményekre vagy eseményosztályokra van-e értelmezve, ezen a szinten – ahogy azt el is várjuk – még nincs eldöntve.

A fentiek fényében az előző fejezet (25. oldal) példája a következőképpen interpretálható.

A kockadobások eseménytere. Az eseménytérben olyan események szerepelnek, mint a „6-os dobás", a „nem 6-os dobás", a „páros dobás", a „hárommal osztható szám dobása". Az eseménytér zárt a metszet és az unió műveletére nézve: a „páros dobás" és a négynél kisebb szám dobása" események metszete a „2-os dobás" esemény, uniója pedig a „nem 5-ös dobás" esemény. Értelmezve van az ortokomplementáció: a „páros dobás" ortokomplementere a „páratlan dobás". A műveletek disztributívak, vagyis az eseménytér klasszikus. Az eseménytér a lehetetlen és a biztos eseménnyel együtt egy 2^6 elemű Boole-algebrát alkot, a $\mathcal{P}(\{1,2,3,4,5,6\})$ halmazalgebrát. A kockadobás eseményterének atomi eseményei, az

$$a_i: \qquad \text{„az } i \text{ szám dobása"}$$

események. Az összetett események közé olyanok tartoznak, mint

$$a: \qquad \text{„4-nél kisebb szám dobása"}$$
$$b: \qquad \text{„páros dobás"}$$

Az eseménytér továbbá atomos és atomisztikus.

Ami az (Ω, Σ, p) hármas utolsó tagját, a valószínűségi mértéket illeti, felhívjuk a figyelmet, hogy éppen ez a

$$p(\{i\}) \equiv \begin{cases} \frac{1}{3}, & \text{ha } i \text{ páros,} \\ 0, & \text{ha } i \text{ páratlan} \end{cases}$$

hozzárendelés nincs még interpretálva. Vagyis hiába tudjuk már, hogy a p függvény *eseményekhez* rendel valószínűségeket, egyelőre nem tudjuk, hogy mit jelent itt a valószínűség, vagyis, hogy hogyan értelmezzük *fizikailag* azt a kijelentést, hogy a páros dobások mind $\frac{1}{3}$ valószínűséggel következnek be, a páratlan dobások pedig 0 valószínűséggel. Ennyiben az (Ω, Σ, p) valószínűségi mértéktér csak félig interpretált, még ha éppenséggel a „cinkelt" kocka mértékterének hívjuk is.

A kockadobást leíró nyelv. Most induljunk ki a fenti metafizikai opció másik ágán, a *kijelentésektől*, vagyis tegyük fel, hogy a valószínűség hordozói nem események, hanem kijelentések. Egészen pontosan induljunk ki egy nulladrendű $\mathcal{L}$ nyelvből, amely a kockadobás eredményeire vonatkozó tényeket reprezentálja. Legyen tehát $\mathcal{L}$ nemlogikai konstansainak halmaza $Con = \{A_i \,|\, i = 1 \ldots 6\}$, amely atomi kijelentések az alábbi tényeknek felelnek meg:

$$A_i : \text{„A kockadobás eredménye } i\text{"},$$

vagyis A_i akkor és csak akkor igaz, ha fennáll az a tény, hogy a kockadobás eredménye i. Az összetett formulák közé ekkor ilyen jelentésű formulák fognak tartozni:

$$\begin{aligned} A \equiv A_1 \vee A_2 \vee A_3 : &\quad \text{„A kockadobás eredménye kisebb, mint 4"} \\ B \equiv A_2 \vee A_4 \vee A_6 : &\quad \text{„A kockadobás eredménye páros"} \end{aligned}$$

amely formulák interpretációját a konnektívumok már meghatározzák: A akkor és csak akkor igaz, ha az A_1-nek, A_2-nek és A_3-nak megfelelő három tény közül legalább egy fennáll. Az interpetáció és a konnektívumok azt is meghatározzák továbbá, hogy a nyelv logikailag ekvivalens formuláit ugyanazok a tények teszik igazzá. Például mind az A_2, mind az $A \wedge B$ formulát az a tény teszi igazzá, hogy a kockadobás eredménye a kettes szám. A nyelv logikailag ekvivalens formulái tehát ugyanarra a tényre vonatkoznak, ugyanaz a jelentésük. Másképp fogalmazva a kockadobás eredményeivel kapcsolatos tények a nyelvnek a logikailag ekvivalens formuláinak osztályaival lesznek „izomorfak". Ez utóbbi viszont éppen az $\mathcal{L}$ nyelv Lindenbaum–Tarski-algebrája (ld. 28. oldal), vagyis a $\mathcal{P}(\{1,2,3,4,5,6\})$ halmazalgebra. Ezen a Boole-algebrán ezek után már megadható a valószínűség mint nomált mérték.[5]

[5]Hangsúlyozzuk, hogy nyelv formuláinak „redundanciája" a tényekhez képest nem azonos a tények „redundanciájával" az eseményekhez képest. Vagyis az, hogy a kettes dobásnak megfelelő tény a nyelvben többféleképpen is reprezentálható (A_2 és $A \wedge B$), nem ugyanaz a redundancia, mint hogy a kettes dobás és a jobb kézzel való kettes dobás ugyanarra a durván tagolt eseményre vonatkozik. Az előbbi a nyelv algebrizálásával megszüntethető, az utóbbi – mint korábban láttuk – az eseményfogalmommal kapcsolatot metafizikai döntéstől függ.

Összefoglalva tehát a következőket mondhatjuk. A Boole-algebra az események tekintetében természetes módon jelenik meg, ha az eseményeket pusztán a bekövetkezés és be nem következés „logikája” szempontjából vizsgáljuk; de hasonlóan természetes módon jelenik meg egy formális nyelv algebrizálása során is, azaz ha a nyelv azonos jelentésű mondatainak struktúráját tekintjük. A valószínűség mint matematikai struktúra ezek után egyszerűen egy a Boole-algebrához számot rendelő függvény. Hogy ezek után ezt a függvényt miként értelmezzük mint az események bekövetkezési valószínűségét illetve az kijelentésekbe vetett hit mértékét, az a következő fejezeteknek lesz a témája.

3.4. Objektív, szubjektív és logikai valószínűség

De térjünk vissza az alapkérdésünkhöz: mit jelent egy esemény (egy kijelentés) valószínűsége? Ennek a kérdésnek a megválaszolása a soron következő fejezeteknek lesz a feladata; ebben az alfejezetben pusztán azokat az „ontológiai tartományokat” kívánjuk felvázolni, amelyekből a válaszokat egyáltalán meríthetjük. Három ilyen tartomány lehetséges:

1. *Objektív valószínűség:* A valószínűség a külvilág eseményeinek tulajdonsága.
2. *Szubjektív valószínűség:* A valószínűség egy (intencionális) elmeállapot tulajdonsága.
3. *Logikai valószínűség:* A valószínűség egy a külvilág eseményei közötti logikai viszony.

Nézzük a tartományokat sorjában.

Objektív valószínűség. A valószínűség objektív interpretációi abból indulnak ki, hogy a valószínűség a külvilág objektív eseményeinek a tulajdonsága. Ez a meghatározás természetesen még nagyon sok szabadsági fokot nyitva hagy az értelmezések számára. A valószínűség lehet szinguláris események és lehet eseményosztályok tulajdonsága. Az események lehetnek a fenti értelemben durvábban és finomabban tagolva. Az események közé tartozhatnak lehetséges események, vagy korlátozhatjuk az események körét az aktuális eseményekre. De bárhogyan döntünk is, a valószínűség ezeknek az eseményeknek lesz a tulajdonsága, nem pedig egy az eseményeket észlelő elméé. Az események valószínűsége nem függ tehát attól, hogy az elme észleli-e ezeket a valószínűségeket, vagy észleli-e egyáltalán magukat az eseményeket.

Szubjektív valószínűség. A valószínűség szubjektív interpretációi a valószínűséget egy az eseményeket észlelő elme intencionális állapota tulajdonságának tekintik. Más szóval a valószínűség a külvilág eseményeinek észlelése során az elmeállapotra jellemző tulajdonság. Ez a külvilági eseményekre vonatkozó intencionális elmeállapot a hit. A hatos dobás egy hatod valószínűsége tehát a hatos dobás szinguláris eseményére vagy eseménytípusára vonatkozó hitünknek valamilye speciális tulajdonsága, például az intenzitása. A valószínűség tehát nem maguknak az objektív (aktuális vagy éppen modális) események a tulajdonsága, hanem a rájuk vonatkozó hité.

Logikai valószínűség. Van azonban egy harmadik lehetőség is: a valószínűség objektív események vagy tények között fennálló logikai vagy a priori viszony. Vegyük a hatos dobás és a páros dobás eseményét. Egy megfelelő formális nyelven belül a között a két állítás között, hogy egy szám hatos, és hogy páros, természetesen lehet logikai a viszony, de a kérdés itt az, hogy lehet-e a hatos dobás és a páros dobás *mint események* közötti viszony is logikai. A standard válasz a kérdésre az, hogy igen, lehetséges, hiszen amennyiben egy adott hatos dobásról empirikusan az derülne ki, hogy az nem páros dobás, akkor a szemantikát változtatnánk meg, vagyis továbbra is ragaszkodnánk ahhoz, hogy a kettőjük közötti viszony logikai (a priori).[6]

Ha elfogadjuk, hogy a hatos dobás és a páros dobás közötti viszony logikai viszony, innen már csak egy lépés azt állítani, hogy *bármely* két, a kockadobásokra vonatkozó esemény közötti viszony is lehet logikai. Nemcsak a hatos dobás eseményéből következik a páros dobás eseménye, hanem visszafelé, a páros dobás eseményéből is következik a hatos dobás eseménye – csak nem teljes, hanem, úgymond, egy harmad mértékben. A valószínűség éppen ennek a részleges következtetésnek a mértéke. Vagy a szokásos beszédmódban, a valószínűség a *konfirmáció* mértéke két kijelentés, a *bizonyíték* (evidencia) és a *hipotézis* között egy *ideális (racionális) ágens* számára

Meg kell azonban jegyeznünk, hogy a logikai valószínűség megengedésével letérünk arról az ösvényről, amelyen eddig haladtunk. A valószínűség interpretációja alatt ugyanis eddig mindvégig *fizikai* interpretációt értettünk, vagyis a valószínűség fogalmával szemben azt a követelményt támasztottuk, hogy azt valamilyen fizikai (operacionalista) eljárással értelmezzük. Ha viszont megengedjük azt, hogy a valószínűség események közötti logikai viszony legyen, akkor ezzel lemondunk a valószínűség ilyen *közvetlen* verifikációjáról. A *közvetett* verifikációról azonban nem kell lemondanunk. A valószínűség ekkor egy fizikai elmélet elméleti terminusa lesz, amely elméletet más pontokon leszünk kénytelenek verifikálni. A lényeg azonban az, hogy a valószínűség fogalmát ekkor is fizikailag interpretáljuk, csak ezúttal egy teljes fizikai elmélet kerülőútján. Vagyis ha valaki azt állítja, hogy a páros dobás eseménye szabályos kocka esetében egy harmad mértékben konfirmálja a hatos dobás eseményét, vagy hogy egy racionális ágens egy harmad mértékben hisz az egyik alapján a másikban, akkor evvel még nem ért az interpetáció végére – meg kell tudnia mondani, hogy *empirikusan* mit ért itt racionalitásan, szabályos kockán, konfirmáción, és hogyan jön ki az egy harmad érték.

A valószínűségnek a következő fejezetekben ismertetett interpretációi mind besorolhatóak a fenti három kategória egyikébe. A relatív gyakoriság-, valamint a *propensity*-interpretáció az „objektív", a szubjektív interpretáció illetve a logikai interpretáció, ahogy nevük is mutatja, a „szubjektív" illetve a „logikai" kategóriákba esnek. A klasszikus interpretáció asszocionista eredetét tekintve alapvetően felette állna ennek a kategorizálásnak, mégis a „szubjektív" kategóriába szokták sorolni.

[6]Hogy ez minden esetben lehetséges manőver-e, az természetesen nyitott kérdés. A kvantumlogika esetét leszámítva azonban valóban nincs történeti példánk az ellenkezőjére.

Most azonban térjünk vissza egy pillanatra az explikáció kérdéséhez. Amint a fejezet bevezetőjében említettük, egy fogalom explikációja a fogalom helyettesítése egy tudományos elmélet terminusával. Ahhoz, hogy a helyettesítés sikeres legyen, előzetesen tisztáznunk kell azonban egy sereg kérdést a helyettesíteni kívánt fogalommal kapcsolatban: melyek a releváns és melyek az irreleváns köznapi használati módjai; vajon a fogalom köznyelvi használatai egy egységes jelentést tükröznek-e vagy a különféle használatok pusztán homonim viszonyban vannak egymással, stb. Ez utóbbi kérdés tisztázásának az a tétje, hogy az előzetes fogalmat egy vagy több tudományos fogalommal helyettesítsük-e, azaz hogy *egy* vagy *több*, a különböző jelentésköröket lefedő tudományos elméletet alkossunk. A „hőmérséklet” fogalmának termodinamikai explikációja jó példa az előbbi esetre; a „tudás” fogalmának két explikációja aszerint, hogy propozícionális tudásról vagy képességtudásról van szó, pedig példa az utóbbira. Mi a helyzet azonban a valószínűséggel?

A fenti interpretációk a valószínűséget minden esetben egyfajta *homogén* fogalomként kezelik, azaz feltételezik, hogy a fogalom minden használata maradéktalanul besorolható egyetlen interpretáció alá. Ennek a megközelítésnek szép hagyománya van a valószínűség interpretációi között. Ilyen hajlíthatatlan álláspontot képviselt például de Finetti, aki valószínűség alatt kizárólag szubjektív valószínűséget értett, vagy von Mises, aki a valószínűség egyetlen értelmes fogalmának a relatív gyakoriságot tekintette. Ezzel a megközelítéssel szemben azonban lehetséges egy másik álláspont is, amely szerint a valószínűség *heterogén* fogalom, vagyis a különböző kontextusokban használt valószínűségfogalom különböző elméleteket von maga után. Ezen az állásponton volt például Carnap, aki a valószínűség két párhuzamosan létező fogalmát különböztette meg, az objektív és logikai valószínűséget. Hasonló módon érvelt az interpretációk pluralitása mellett Mellor is, aki szerint az alábbi három mondat valószínűségfogalma más-más módon interpretálandó: „A dohányosok nagyobb valószínűséggel kapnak rákot, mint a nem dohányosok.” „Azt hiszem, hogy valószínűleg esni fog holnap.” „A csillagászati adatok nagyon valószínűvé teszik, hogy az Univerzumnak van kezdete.” Az első mondat valószínűségfogalma objektív, a másodiké szubjektív, a harmadiké logikai. A legjobb tehát külön jelölést bevezetni a háromra – objektív valószínűség: ch (*chance*), szubjektív valószínűség: cr (*credence*), logikai valószínűség: lo (Mellornál: ep (*epistemic*)).

Aki többféle valószínűség létezésében hisz, az a különböző valószínűségek közötti viszonyokra vonatkozóan hipotéziseket is megfogalmazhat. Ezeket a hipotéziseket hívják valószínűségi *koordinációs elveknek*. A szubjektív és objektív valószínűség közötti viszonyra vonatkozó leghíresebb ilyen koordinációs elv az ún. *Principal Principle*. Az elv David Lewis-tól (1986) ered, és a következőket állítja. Ha egy X szubjektumnak egy a állításra vonatkozóan nincs egyéb releváns tudása, mint hogy az a esemény objektív valószínűsége y, akkor X-nek az esemény bekövetkezésére vonatkozó szubjektív valószínűsége is y. A feltételes valószínűséget használva ez így írható:

$$cr(a \mid ch(a) = y) = y$$

A *Principal Principle*-lel kapcsolatos filozófiai viták általában az alábbi kérdések körül forognak. (i) Milyen típusú elv a *Principal Principle*: empirikus, racionalitási vagy egyszerűen az objektív valószínűség definíciója? Akik az elvet empirikus elvnek tekintik, igazolniuk kell, hogy az empirikus ágensek hitfüggvényei az elvnek megfelelően alakulnak. A feladat elől úgy lehet kitérni, hogy az elvet normatív elvként értjük, még pedig a *racionális* ágensre vonatkozó normatív elvként. Ekkor például a képletben szereplő cr, szubjektív valószínűséget legjobb lecserélni a lo, logikai valószínűségre, vagyis az elvet az objektív és a logikai valószínűség közötti koordinációs viszonynak tekinteni. A harmadik lehetőség, hogy az elvet egyszerűen az objektív valószínűség definíciójának tekintjük: az objektív valószínűség bármi, ami kielégíti a *Principal Principle*-t. Akkor azonban az elv nem lesz koordinációs elv.

A racionalitási elvnek tekintett *Principal Principle* számára egy további kérdés, hogy (ii) mit értünk azon, hogy az elv racionalitási elv. A szokásos válasz valamilyen ún. *racionális stratégia*, pl. *Dutch book*-argumentum felmutatása, amelyből az elv következik. De valóban védhető-e a *Principal Principle* ilyen stratégiával, és ha igen, mit jelent az, hogy egy stratégia „racionális”?[7]

Egy harmadik, konkrét kérdés, hogy (iii) mit jelent az a kitétel az elvben, hogy az X szubjektum a objektív valószínűségén túl *nem rendelkezik egyéb releváns információval* a bekövetkezésére vonatkozóan. Nyilvánvaló, hogy ha X tudja, hogy a bekövetkezett, akkor az a-ba vetett hitének mértéke nem lesz egyenlő a valószínűségével (hacsak a bekövetkezett események valószínűségét nem vesszük egynek). Vagyis X-nek nem lehet információja arról, hogy a bekövetkezett-e. Kérdés: mi másról nem lehet még?

A *Principal Principle* és a hasonló koordinációs elvekkel együtt azonban egytől-egyig feltételezik, hogy a valószínűségnek létezik legalább két egyformán helyes interpretációja, amely interpetációk viszonya magyarázatra szorul. Hogy ez így van-e, az természetesen nem válaszolható meg azelőtt, mielőtt a konkrét interpretációkat egyesével meg nem vizsgáltuk.

3.5. A Simpson-paradoxon és a közös ok

A valószínűséggel kapcsolatos filozófiai viták nagyon gyakran a valószínűség egyfajta intuitív értelmezése mellett zajlanak. Ezek a viták olykor olyan súlyos filozófiai kérdések körül forognak, mint a valószínűség és a kauzalitás vagy a valószínűség és a magyarázat viszonyának kérdése. Ebben az alfejezetben két példán, a Simpson-paradoxon illetve a reichenbachi közös ok példáján azt szeretnénk illusztrálni, hogy ezek a filozófiai kérdések nem dönthetők el azelőtt, mielőtt ne nyilatkoznánk a valószínűség természetét illetően.

A *Simpson-paradoxon* egy a californiai Berkeley-n 1973-ban történt eset után került az érdeklődés homlokterébe, ahol is az egyetemre felvettek nemi arányában diszkrimináció mutatkozott a fiúk javára. A közelebbi vizsgálatok során azonban az derült ki, hogy az

[7] A *Principal Principle* racionalitási elvként való *Dutch book*-védelmével kapcsolatban lásd (Szabó, 2010).

egyes szakokra vetítve a diszkrimináció éppen a lányok javára áll fent. Modellezzük a szituációt az alábbi táblázattal:

	lány	fiú
történelem	$\frac{5}{55}$	$\frac{1}{45}$
földrajz	$\frac{45}{45}$	$\frac{50}{55}$
egyetem	$\frac{50}{100}$	$\frac{51}{100}$

A táblázatban első két sorában megadott törtek azt jelentik, hogy az adott szakra jelentkezett lányok illetve fiúk közül hányat vettek fel, pl. történelem szakra 55 lányból 5-öt; az utolsó sor pedig azt jelenti, hogy a teljes egyetemre (ez esetben a két szak valamelyikére) az összes 100 jelentkező lányból illetve fiúból hányat vettek fel. A számokból nyilvánvalóan látszik, hogy az egyetemre több fiút vettek fel, mint lányt, jóllehet mindkét szakon a lányok felvételi aránya a magasabb. A paradoxon magyarázata abban a felismerésben áll, hogy a lányok nagyobb mértékben jelentkeztek a nehezebb történelem szakra, és az itt elvérzett felvételizőket nem tudta kompenzálni a könnyebb szakra felvettek száma sem.

A fenti megfogalmazás statisztikai. A valószínűség nyelvén a paradoxon az alábbi módon fogalmazható meg. Legyen az a, b és c események valamint ortokomplementereik az alábbi módon definiálva:

$$\begin{aligned}
a &: \quad \text{fiú} \\
a^\perp &: \quad \text{lány} \\
b &: \quad \text{felvették} \\
b^\perp &: \quad \text{nem vették fel} \\
c &: \quad \text{történelem szakra jelentkezett} \\
c^\perp &: \quad \text{földrajz szakra jelentkezett}
\end{aligned}$$

Az eseménytér legyen a fenti események által generált nyolc atomos algebra, ahol az atomok mértékét a fenti táblázatból olvashatjuk le nyilvánvaló módon:

$$\begin{aligned}
p(a \wedge b \wedge c) &= \frac{1}{200} & p(a \wedge b \wedge c^\perp) &= \frac{50}{200} \\
p(a^\perp \wedge b \wedge c) &= \frac{5}{200} & p(a^\perp \wedge b \wedge c^\perp) &= \frac{45}{200} \\
p(a \wedge b^\perp \wedge c) &= \frac{44}{200} & p(a \wedge b^\perp \wedge c^\perp) &= \frac{5}{200} \\
p(a^\perp \wedge b^\perp \wedge c) &= \frac{50}{200} & p(a^\perp \wedge b^\perp \wedge c^\perp) &= 0
\end{aligned}$$

Ekkor könnyen látható, hogy az a és b események között az alábbi korrelációt kapjuk:

$$\frac{51}{200} = p(a \wedge b) \quad > \quad p(a)p(b) = \frac{1}{2} \cdot \frac{101}{200}$$

vagyis az egyetemre felvettek és a férfiak között pozitív korreláció áll fenn. Ha azonban a valószínűségeket kondicionáljuk a c illetve a $c^\perp$ eseményre, vagyis a szakokra, a korreláció előjelet vált:

$$\begin{aligned} \frac{1}{100} = p(a \wedge b|c) \quad &< \quad p(a|c)p(b|c) = \frac{45}{100} \cdot \frac{6}{100} \\ \frac{1}{2} = p(a \wedge b|c^\perp) \quad &< \quad p(a|c^\perp)p(b|c^\perp) = \frac{55}{100} \cdot \frac{95}{100} \end{aligned}$$

A példa tehát azt mutatja, hogy két esemény közötti korreláció előjele megfordítható az eseménytér egy adott két elemű partíciójára való kondicionálás révén.

Előfordulhat az is, hogy egy a és b esemény közötti (pozitív vagy negatív) korrelálót egy harmadik c eseményre illetve annak $c^\perp$ ortokomplementumára való kondicionálás függetlenné tesz, vagyis teljesül az alábbi két összefüggés:

$$\begin{aligned} p(a \wedge b|c) \quad &= \quad p(a|c)p(b|c) \\ p(a \wedge b|c^\perp) \quad &= \quad p(a|c^\perp)p(b|c^\perp) \end{aligned}$$

Ilyenkor azt mondjuk, hogy a és b *feltételesen függetlenek* c-re nézve.[8] Ha e két tulajdonságon kívül továbbá az is teljesül, hogy a c és az a illetve a c és a b események között korreláció pozitív, akkor a c eseményt Reichenbach (1956) nyomán az a és b esemény közötti korreláció közös okának nevezzük.

12. Definíció. Legyen (Ω, Σ, p) valószínűségi mértéktér, és legyen a és b két pozitíven korreláló esemény Σ-ban. Egy $c \in \Sigma$ eseményt az a és b esemény közötti korreláció *(reichenbachi) közös okának* nevezzük, ha az a, b, c és $c^\perp$ eseményekre az alábbi összefüggések fennállnak:

$$p(a \wedge b|c) = p(a|c)p(b|c) \tag{3.1}$$
$$p(a \wedge b|c^\perp) = p(a|c^\perp)p(b|c^\perp) \tag{3.2}$$
$$p(a|c) > p(a|c^\perp) \tag{3.3}$$
$$p(b|c) > p(b|c^\perp) \tag{3.4}$$

Egy közös okot valószínűségelméletileg tehát két feltétel körvonalaz: egyrészt két korreláló eseményt feltételesen függetlenné tesz, másrészt növeli mindkét esemény valószínűségét. Reichenbach (1956) a közös okra vonatkozóan a következő példákkal szolgál:

> ”Tegyük fel egy szobában két lámpa egyszerre kialszik. Azt, hogy a két izzó véletlenül aludjon ki egyszerre, valószínűtlennek találjuk, és kiégett biztosíték után nézünk, vagy az áramellátás ingadozását tesszük felelőssé. A valószínűtlen

[8] A függetlenségre, feltételes függetlenségre vonatkozó tételeket lásd a D. Függelékben.

> koincidencia így egy közös ok produktumaként nyer magyarázatot. Vagy tegyük fel egy közös társulatnál néhány játszó színész megbetegszik ételmérgezési tüneteket mutatva. Ilyenkor feltesszük, hogy a romlott étel egy közös forrásból, például a menzáról származik, és a koincidenciát egy közös ok segítségével magyarázzuk." (157. o.)

Reichenbach koincidenciát használ korreláció helyett, de a példák valószínűségi átírása nyilvánvaló. A színészek együttes rosszullétének valószínűsége nagyobb, mint az egyedi rosszullétek valószínűségének szorzata, vagyis korreláció van a rosszullétek között. A közös ok, a menzán elfogyasztott étel mellett azonban a valószínűségek függetlenednek: aki evett a romlott ételből, az megbetegszik (determinisztikus eset), vagy nagy valószínűséggel megbetegszik (indeterminisztikus eset), mindenesetre az együttes megbetegedések valószínűsége – feltéve, hogy ettek a romlott ételből – a külön-külön megbetegedések szorzata lesz. Hasonlóképpen a közös ok komplementere is függetleníti a valószínűségeket: akik nem ettek a romlott ételből, azok (más közös ok híján) egymástól függetlenül betegszenek meg. Továbbá a közös ok növeli az okozatok valószínűségét: aki romlott ételből eszik, az inkább rosszul lesz, mint aki nem. Vagyis az okok és az okozatok kielégítik a fenti valószínűségi kritériumokat.

A fenti kritériumok a közös ok fogalmának természetesen csak valószínűségelméleti jellemzését adják, vagyis ahhoz, hogy c az a és b esemény *valódi* közös okának számítson, még egyéb, nem valószínűségi (téridőbeli, stb.) feltételeket is ki kell elégítenie. Röviden, a fenti kritériumok csak szükséges feltételek.

A közös ok reichenbachi definíciója tehát két követelmény együttes érvényesítése. A (3.1)-(3.2) egyenleteket az irodalomban *leárnyékolási (screening-off) kritériumoknak* nevezik, (3.3)-(3.4) egyenlőtlenségeket pedig *pozitív statisztikus relevanciának*. Reichenbach vélhetően két okból ragaszkodott ez utóbbi egyenlőtlenségekhez a közös ok definíciójában. Egyrészt a közös ok definíciójából így következni fog, hogy az a és a b események pozitívan korrelálnak, vagyis fennáll a következő egyszerű lemma.

Lemma. Legyen (Ω, Σ, p) valószínűségi mértéktér, és legyen a, b és c olyan események Σ-ban, amelyekre (3.1)-(3.4) fennállnak. Ekkor a és b pozitívan korrelálnak.

Az állítás bizonyítása egyszerű, csupán az alábbi összefüggést kell használni:

$$p(a \wedge b) - p(a)p(b) \quad = \quad p(c)p(c^{\perp})[p(a|c) - p(a|c^{\perp})][p(b|c) - p(b|c^{\perp})], \tag{3.5}$$

ahol is kihasználtuk (3.1)-(3.2)-t, vagyis azt, hogy a és b feltételesen függetlenek c-re és $c^{\perp}$-re nézve. Mivel $p(c)$ és $p(c^{\perp})$ nem negatív, ezért (3.3)-(3.4)-et feltételezve (3.5) jobb oldala és így bal oldala is pozitív, vagyis a és b között pozitív korreláció lesz. A közös ok léte tehát a hempeli értelemben „megmagyarázza" a korrelációt. A két egyenlőtlenséget elengedve azonban a pozitív korreláció nem volna levezethető, így (3.3)-(3.4) elengedhetetlen része a magyarázatnak.

A másik ok (3.3)-(3.4) inkorporálására véltetően az volt, hogy a közös ok fogalma túl szorosan kötődött a hagyományos ok fogalmához. Ha a közös ok egyben ok is, és az ok elégséges feltétel, akkor a közös ok bekövetkezésének is maga után kell vonnia okozatainak bekövetkezését. Indeterminista esetben azonban az ok mint elégséges feltételnek átadja a helyét a pozitív statisztikus relevenciának: az ok, ha nem vonja is maga után, de legalább növeli az okozat bekövetkezési esélyét. Vagyis (3.3)-(3.4) az ok mint elégséges feltétel természetes általánosítása.

Egyik megfontolás sem indokolja azonban, hogy ragaszkodjunk a közös ok eredeti reichenbachi definíciójához. Egyrészt amint (3.5)-on látható, akkor is pozitív korrelációt kapnánk, ha (3.3)-(3.4) ellenkező előjellel teljesülnének, vagyis ha nem c, hanem $c^{\perp}$ lenne az a és b között korreláció közös oka. Így tehát (3.3)-(3.4) előírása felesleges: ha az a és b közötti korreláció pozitív, és c illetve $c^{\perp}$ kielégítik (3.1)-(3.2)-t, akkor vagy kielégítik (3.3)-(3.4)-t is, és ekkor c lesz a közös ok; vagy kielégítik (3.3)-(3.4) ellenkezőjét, és ekkor $c^{\perp}$ lesz közös ok.

Másfelől az is kiderült, hogy a közös ok fogalma nem kötődik szorosan az ok hagyományos fogalmához, vagyis tömören szólva c nem a és b együttes oka, hanem az a és b közötti *korreláció* oka. Ezért mára a szakirodalomban a közös ok definíciójából rendszerint el is hagyják az utolsó két feltételt.

A közös ok fogalma könnyen általánosítható két elemű partícióról finomabb partíciókra. Tekintsük a fenti Reichenbach-idézet másik példáját, amely kialvó izzókra vonatkozik. Tegyük fel, hogy egy a és a b izzó elsötétedése között pozitív korrelációt tapasztalunk az alábbi módon:

$$0,485 = p(a \wedge b) \quad > \quad p(a)p(b) = 0,6 \cdot 0,6$$

Tegyük fel továbbá, hogy az áramingadozásokra kondicionálás során azt látjuk, hogy a korreláció nem tűnik el, azaz az áramingadozásokat c_1-gyel jelölve az alábbi egyenlőtlenségek állnak fenn:

$$\begin{aligned} 0,725 = p(a \wedge b|c_1) \quad &> \quad p(a|c_1)p(b|c_1) = 0,85 \cdot 0,85, \\ 0,245 = p(a \wedge b|c_1^{\perp}) \quad &> \quad p(a|c_1^{\perp})p(b|c_1^{\perp}) = 0,35 \cdot 0,35, \end{aligned}$$

vagyis az áramingadozás nem lehet közös oka a korrelációnak. Tegyük fel továbbá, hogy figyelmesek leszünk egy másik releváns körülményre, mondjuk arra, hogy a szobában ingadozik a hőmérséklet, és ez esetleg oka lehet az izzók elsötétedésének. Azt tapasztaljuk azonban, hogy az izzók kialvásai pusztán a hőmérsékletingadozásokat tekintve sem válnak függetlenné, hanem mondjuk a következő egyenlőtlenségek teljesülnek:

$$\begin{aligned} 0,65 = p(a \wedge b|c_1) \quad &> \quad p(a|c_1)p(b|c_1) = 0,8 \cdot 0,8, \\ 0,32 = p(a \wedge b|c_1^{\perp}) \quad &> \quad p(a|c_1^{\perp})p(b|c_1^{\perp}) = 0,4 \cdot 0,4. \end{aligned}$$

Észrevesszük ugyanakkor, hogy a két ok együttes figyelembevétele mellett az okozatok közötti korreláció eltűnik. Más szóval képezve a két okból a $\{c_1 \wedge c_2, c_1 \wedge c_2^\perp, c_1^\perp \wedge c_2, c_1^\perp \wedge c_2^\perp\}$ négy elemű partíciót, a partíció bármely elemére kondicionálva a korreláció eltűnik, azaz

$$\begin{aligned}
0,81 = p(a \wedge b|c_1 \wedge c_2) &= p(a|c_1 \wedge c_2)\, p(b|c_1 \wedge c_2) = 0,9 \cdot 0,9, \\
0,64 = p(a \wedge b|c_1 \wedge c_2^\perp) &= p(a|c_1 \wedge c_2^\perp)\, p(b|c_1 \wedge c_2^\perp) = 0,8 \cdot 0,8, \\
0,49 = p(a \wedge b|c_1^\perp \wedge c_2) &= p(a|c_1^\perp \wedge c_2)\, p(b|c_1^\perp \wedge c_2) = 0,7 \cdot 0,7, \\
0 = p(a \wedge b|c_1^\perp \wedge c_2^\perp) &= p(a|c_1^\perp \wedge c_2^\perp)\, p(b|c_1^\perp \wedge c_2^\perp) = 0 \cdot 0,
\end{aligned}$$

Ekkor a $\{c_1 \wedge c_2, c_1 \wedge c_2^\perp, c_1^\perp \wedge c_2, c_1^\perp \wedge c_2^\perp\}$ partíciót – a fenti terminológia általánosításaként – az a és b közötti korreláció közös okrendszerének nevezzük. Megadva a partíció elemeinek valószínűségét, a példát modellező valószínűségi mértéktér könnyen megszerkeszthető.

A példa általánosításaként egy $\{c_i\}_{i=1}^n$ partíciót az a és b közötti korreláció ***közös okrendszerének*** nevezzünk (természetesen a fenti statisztikus értelemben), ha partíció minden c_i elemére teljesül, hogy:

$$p(a \wedge b|c_i) \;=\; p(a|c_i)p(b|c_i).$$

Végül szeretnénk hangsúlyozni, hogy sem a Simpson-paradoxonban, sem a reichenbachi közös ok fenti definíciójában a valószínűség nincs interpretálva; a fogalmat csupán intuitív értelemben használtuk. A közös ok definíciójának meggyőző ereje azonban nagyban függ attól, hogy a benne szereplő valószínűségnek mi is a természete. Ha a valószínűség pusztán intencionális elmeállapot, akkor korántsem világos, hogy a reichenbachi kritériumokat kielégítő események miért is jelentenék egy korreláció kauzális magyarázatát, amely korreláció így maga is elmeállapotok közötti viszony. Vagy másképp fogalmazva, ha a valószínűséget szubjektívan értelmezzük, akkor miért is előbbrevalóak a magyarázat szempontjából a független események a korreláló eseményeknél. De ha a valószínűség objektív fogalom is, akkor sem mindegy, hogy minek is a tulajdonsága. Ha egyedi eseményeké, akkor szintén nehezen látható, hogy miben is állna ilyen egyedi események közötti korreláció közös ok típusú kauzális magyarázata. Jól látható tehát, hogy a valószínűség fogalmát használó filozófiai problémák nem kerülhetik el, hogy mondjanak valamit arról, hogy milyen valószínűségfogalommal dolgoznak.

4. fejezet

A klasszikus interpretáció

A klasszikus valószínűség annak a felvilágosodásnak a terméke, amely tudományos világképének középpontjában a determinizmus áll. A determinizmus egyik legismertebb megfogalmazása így hangzik:

> „Az Univerzum jelenlegi állapotát úgy kell tekintenünk, mint a megelőző állapot következményét, és a rákövetkező okát. Ha adva volna egy értelem, amely ismerné az összes erőt, amellyel a természet rendelkezik és az őt alkotó létezők helyzetét – egy elegendően hatalmas értelem tehát, amely elemezni lenne képes mindezt az adathalmazt – minden mozgást az Univerzum legnagyobb égitestétől annak legkisebb atomjáig egyazon képletbe lenne képes foglalni; számára semmi sem lenne bizonytalan, és a jövő csakúgy, mint a múlt, jelen lenne szemei előtt." (4. o.)

Az idézet Pierre-Simon de Laplace *Essai philosophique sur les probabilités* című 1814-es monumentális munkájából származik, amely egyben a valószínűség klasszikus interpretációjának is forrása. Laplace mindentudó démonja számára tehát sem a múlt, sem a jövő nem rejteget bizonytalanságot; számunkra, halandók számára azonban ez az ideális tudás elérhetetlen:

> „Minden esemény, még azok is, amelyek jelentéktelen voltuknál fogva nem látszanak követni a természet nagy törvényeit, ugyanolyan szükségszerű következménye azoknak, mint a Nap forgása. Mivel azonban nem ismerjük azokat a szálakat, amelyek az ilyen eseményeket az univerzum teljes rendszeréhez fűzik, célokoktól és a véletlentől tesszük őket függővé, attól függően, hogy szabályosan történnek illetve ismétlődnek-e, avagy mindenféle rend nélkül; azonban ezek a képzelt okok a tudás körének bővülésével fokozatosan veszítenek jelentőségükből, és teljességgel megszűnnek a helyes filozófia színe előtt, amely bennük csupán az igazi okokra vonatkozó tudatlanságunk kifejeződését látja." (3. o.)

A valószínűség tehát a tudásnak az emberi természetből fakadó korlátozottságának korrelátuma. „A valószínűség részben a tudatlanságunkra vonatkozik, részben a tudásunkra.” (6. o.) Laplace ezt a következő példával világítja meg:

> „Tegyük fel, hogy van például három urnánk, A, B és C, amelyek közül az egyik csak fekete golyókat tartalmaz, míg a másik kettő csak fehéreket. A C urnából húzunk egy golyót. Szeretnénk meghatározni, hogy mekkora a valószínűsége annak, hogy a kihúzott golyó fekete. Ha nem tudjuk, hogy a három urna közül melyikben vannak a fekete golyók, vagyis nincs okunk inkább a C urnát feltételezni, mint az A-t vagy a B-t, úgy ez a három hipotézis egyformán lehetségesnek fog tűnni, és mivel fekete golyó csak az első hipotézis mellett húzható, így ennek valószínűsége egy harmad lesz. Ha viszont azt tudjuk, hogy az A urna csak fehér golyókat tartalmaz, a határozatlanság csak a B és C urnákra terjed, így annak valószínűsége, hogy a C urnából fekete golyót húzunk, egyketted lesz. Ez a valószínűség végül bizonyossággá változik, ha biztosak vagyunk benne, hogy mind az A mind a B urna csak fehér golyókat tartalmaz.” (8. o.)

Laplace példája az *elégtelen ok elvének* illusztrációja, amely szerint ha nincs okunk két vagy több esemény tekintetében inkább az egyik bekövetkezésében hinni, mint a másikéban, akkor az eseményeket egyenlően valószínűnek kell tartanunk. Az elnevezés Jacob Bernoullitól származik, vélhetően Leibniz *elégséges ok elvének* egyfajta kontrapozíciójaként. Később az elvet Keynes *az indifferencia elvének* keresztelte, és ezen a néven is vonult be a szakirodalomba. Az elégtelen ok vagy indifferencia elve tehát a klasszikus valószínűség metafizikai alapja. De mi is ez a klasszikus valószínűség? Laplace meghatározása így hangzik:

> „A véletlen elmélete abban áll, hogy minden azonos fajtájú eseményt egy bizonyos számú egyenlően lehetséges esetre vezetünk vissza, azaz olyanokra, amelyek létezésének tekintetében egyformán határozatlanok vagyunk, és meghatározzuk a kérdéses esemény szempontjából kedvező esetek számát. Ennek a számnak az aránya az összes lehetséges esethez viszonyítva lesz a valószínűség mértéke, amely így egyszerűen egy tört, amelynek számlálója a kedvező esetek száma, nevezője pedig az összes lehetséges eset száma.” (6-7. o.)

Egy esemény valószínűsége tehát *az eseményt megvalósító kedvező esetek száma osztva az összes egyenlően lehetséges eset számával.* Ez a valószínűség klasszikus és egyben történetileg az első meghatározása.

A klasszikus valószínűségfogalom felé vezető lépések Laplace gondolatmenetében tehát a következők:

1. Mivel a világban determinizmus uralkodik, ezért a valószínűség csak tudásunk hiányával állhat kapcsolatban.

2. Az ignoranciával, a tudás hiányával kapcsolatos ún. episztemikus valószínűség egyben szubjektív valószínűség is.

3. Ennek a szubjektív valószínűségnek az értékei az indifferencia elve segítségével származtathatók.

Az alábbiakban megmutatjuk, hogy a fenti három kijelentés egyike sem tartható. Vagyis determinista világban nem csak episztemikus valószínűségek létezhetnek; az episztemikus valószínűségek nem feltétlenül szubjektívek; és az indifferencia elve nem nyújt alapot a valószínűség értelmezéséhez. Az első alfejezetben az első két kérdést tisztázzuk, majd kitérésként bemutatjuk a klasszikus valószínűség egyik fontos alkalmazását, Laplace rákövetkezési szabályát. Végül a fejezet végén visszatérünk a harmadik kérdéshez, a definíció magvát jelentő indifferencia elvének értelmezéséhez.

4.1. Episztemikus valószínűség és indeterminizmus

Tekintsük tehát a laplace-i gondolatmenet első két pontját. A két megállapításban két fogalompár szerepel, amelyekkel kapcsolatban a szakirodalomban ma sem ritkák a félreértések. Az egyik fogalompár az ún. episztemikus valószínűség és az előző fejezetben (50. oldal) definiált szubjektív valószínűség, a másik fogalompár a valószínűség és a determinizmus. Kezdjük az első fogalompárral.

Episztemikus és szubjektív valószínűség. A fizikai irodalomból elterjedve *episztemikusnak* szokás nevezni azt a valószínűséget, amely korlátozott tudásunkkal vagy megfigyelőképességünkkel kapcsolatos. Ahogy az a fenti idézetekből kitűnik, Laplace szerint minden valószínűség episztemikus, „csupán az igazi okokra vonatkozó tudatlanságunk kifejeződései", amelyek „teljességgel megszűnnek a helyes filozófia színe előtt". De következik-e abból, hogy egy valószínűség episztemikus, az is, hogy szubjektív, azaz, hogy az elme valamilyen mentális állapotának tulajdonsága?

Egyáltalán nem. Hogy mennyire abszurd volna a korlátozott episztemikus hozzáférésből arra következtetni, hogy a valószínűség valamiféle mentális tulajdonság, azt könnyen megvilágíthatjuk a statisztikus fizika példáján. A statisztikus fizika szerint egy gáz valamely makroszkopikus tulajdonságát (hőmérsékletét) a gázrészecskék mikroszkopikus tulajdonságai (kinetikus energiája, stb.) határozzák meg. A meghatározás átlagolással történik: mikroállapotoknak valamilyen a priori súlyt adunk, majd ezekkel a súlyokkal a mikroszkopikus tulajdonságokból átlagokat számolunk, végül az így nyert átlagokat azonosítjuk az ismert termodinamikai mennyiségekkel. A feladat az a priori eloszlások meghatározásának nehézségén túl persze rengeteg fizikai és filozófiai kérdést felvet, az a priori választás fizikai indoklásától (ergodhipotézis), az átlagok termodinamikai azonosításának igazolásáig (ld. Sklar, 1993). A koncepció azonban világos: a makroszkopikus mennyiségek az objektívnak tekintett mikroszkopikus mennyiségek átlagolása révén adódnak.

Hogyan értelmezzük azonban a fenti gondolatmenetben a mikroállapotoknak tulajdonított a priori súlyt? Itt lép be a gondolatmenetbe az episztemikus értelmezés. A mikroállapotokhoz a részecskék mérete és száma miatt nem férünk hozzá közvetlenül, ezért a róluk alkotott tudásunk csak valószínűségi jellegű. A valószínűségek pontos meghatározásához különféle metafizikai elveket hívhatunk segítségül, például az indifferencia elvét (ld. alább), vagy megpróbálhatjuk az a priori valószínűséget az ergodhipotézisből származtatni (vagyis az alulfekvő dinamikából időátlagolással), vagy egyszerűen próbálkozhatunk egy tetszőleges valószínűségi *ansatz*-cal, amelyet az átlagolás végén a makroszkopikus mennyiségekre adott jóslatain keresztül ellenőrzünk. Mindenesetre ezek az episztemikus valószínűségek lesznek azok a súlyok, amelyekkel átlagolva elméletünk makroszkopikus mennyiségeit származtatjuk. És itt jön a döntő lépés. Ha most azt állítanánk, hogy ezek *az episztemikus valószínűségek szubjektív valószínűségek*, azaz az elme hiteinek mérőszámai, akkor az átlagolással adódó makroszkopikus mennyiségekről is azt lennénk kénytelenek állítani, hogy azok (legalábbis részben) szubjektív mennyiségek: a hőmérséklet a gáznak szubjektív tulajdonsága, mivel statisztikus levezetésébe beszüremkednek episztemikus(=szubjektív!) elemek is.

Ezt állítani nyilvánvalóan abszurdum. A gondolatmenet két dolgot mos egybe. Az egyik az a kétségtelen tény, hogy a statisztikus leírásra valóban azért van szükségünk, mert megfigyelőképességünk korlátozott. Ennyiben a valószínűségi leírásra való *igény*, ha úgy tetszik, episztemikus eredetű. Ugyanakkor azonban az átlagolással nyert mennyiségek ugyanannyira objektívek lehetnek, mint az átlagolandók – egy család átlagjövedelme ugyanúgy objektív, a családra jellemző mennyiség, mint az egyes tagok keresete. Vagyis egy rendszerről adott leírás objektivitása nem feltételezi, hogy a rendszerről *ne lehessen finomabb felbontású igaz leírást adni.* Az átlagos kinetikus energiaként származtatott hőmérséklet mindaddig objektívnek tekinthető, ameddig értéke megegyezik a hőmérőről leolvasott értékkel. Ezen a tényen pedig semmit sem változtat a mikrorészekre vonatkozó tudásunk állapota. Ha a mikrorészecskék viselkedését egyedileg követni tudnánk, attól még a hőmérséklet továbbra is ugyanolyan objektív mennyiség maradna, legfeljebb nem volna okunk tovább használni.[1]

Röviden, élesen meg kell különböztetnünk egy jelenségkör valószínűségi leírását igazoló episztemikus okokat attól, hogy a szóban forgó valószínűségek szubjektívek-e, azaz egy mentális állapot tulajdonságai. Ez utóbbi egy erős metafizikai állítás a valószínűség természetéről, míg az előbbi pusztán a valószínűség alkalmazásával összefüggő ismeretelméleti előfeltételek megjelölése.[2] A félreértéseket elkerülendő a továbbiakban mi nem használjuk az „episztemikus” kifejezést, hanem ragaszkodunk az előző fejezetben felállított distinkcióhoz:

[1] Az episztemikus valószínűség használatából származó fogalmi zavarokról lásd még (Szabó, 2009).

[2] Az objektívról a szubjektívra való áttérés van Fraassen (1980) szerint a következő ún. *statisztikus szillogizmus* révén történik.

- „Az 1944-ben besorozottak 73 százaléka még él.
- Jonest 1944-ben sorozták.
- Nincs egyéb releváns információm arról, hogy vajon Jones él-e még, vagy sem.
- Ezért, annak valószínűsége (számomra), hogy Jones él-e, 73 százalék.” (165. o.)

Az érv első két premisszája tisztán objektív állítás; szubjektivista konklúziót a harmadik premissza eredményezi, amely tudásállapotomról szól. Enélkül a premissza nélkül a Jonesra vonatkozó valószínűség nem lenne végül szubjektív – írja van Fraassen.

az objektív valószínűség a külvilág valamely szinguláris eseményének vagy eseménytípusának tulajdonsága, a szubjektív valószínűség pedig egy eseményre vonatkozó intencionális elmeállapot tulajdonsága. Ebben a terminológiában a statisztikus fizika valószínűségei nem szubjektívek (legalábbis az előző téves gondolatmenet alapján nem azok). Hogy kijelenthessük, hogy objektívak, ahhoz persze még meg kell mondanunk, hogy mit is értünk objektív valószínűségen. Most azonban térjünk át a második fogalompár vizsgálatához.

Determinizmus és valószínűség. Amint a fenti idézetekből kitűnik, Laplace az általános determinizmusból eredezteti az episztemikus valószínűségek gondolatát: „Az Univerzum jelenlegi állapotát úgy kell tekintenünk, mint a megelőző állapot következményét, és a rákövetkező okát", vagyis a világban determinizmus uralkodik. Mi, halandó lelkek „azonban nem ismerjük azokat a szálakat, amelyek az ilyen eseményeket az univerzum teljes rendszeréhez fűzik, [ezért] célokoktól és a véletlentől tesszük őket függővé" – vagyis a valószínűség pusztán episztemikus természetű. Nézzük az idézetben szereplő fogalmakat részletesen.

Először is a modern megközelítés szétválasztja az okság és a determinizmus kérdését, vagy legalábbis a determinizmus fogalmát nem terheli tovább az okság amúgy is igen nehéz fogalmával. A determinizmust a következőképpen szokás megfogalmazni: a világban akkor uralkodik determinizmus, ha a világ állapota egy adott időpontban a természettörvényekkel együtt rögzítik a világ további sorsát. A meghatározás természetesen számos ponton magyarázatot kíván: például, hogy mik a természettörvények, mit értünk „rögzíteni" alatt stb. Ezeket itt most nem tárgyaljuk, mivel a meghatározás intuitíve világos.[3] A kérdés az, hogy hogyan függ össze a világ determinizmusa a valószínűségek létével és típusával.

A válasz: sehogy – mind determinista, mind indeterminista világban létezhet egyaránt szubjektív és objektív valószínűség. Lássuk az eseteket.

Ha a világban *in*determinizmus uralkodik, vagyis a világ soron következő állapotai nem rögzítik egymást egyértelműen, akkor az objektív valószínűség lehet ennek a meghatározatlanságnak valamilyen fokmérője, vagyis az a (meta)fizikai mennyiség, amely valamiképpen megadja a fizikailag lehetséges jövőbeli állapotok „eloszlását". Az ilyen értelemben vett objektív valószínűség verifikálhatósága persze kérdéses, de ettől most tekintsünk el egy pillanatra. Szubjektív valószínűségen ugyanebben a világban érthetjük például az elmeállapotoknak azt a tulajdonságát, hogy a világ objektíve eldöntetlen jövőbeli állapotaiban éppen a nekik megfelelő objektív valószínűségnek megfelelő mértékben hisz. Természetesen a szubjektív valószínűségnek nem kell ilyen előre elrendelt harmóniában lennie az objektív valószínűséggel; a világra vonatkozó hitek éppenséggel lehetnek más szempontok szerint is meghatározva. A lényeg, hogy indeterminista világban mindkét fajta valószínűség könnyen értelmezhető.

Kérdés azonban, hogy van Fraassen helyesen magyarázza-e meg az objektív valószínűségről a szubjektív valószínűségre való áttérést. Milyen értelemben következik ugyanis a konklúzió a premisszákból? Logikailag nyilván nem. A szillogizmus legfeljebb egy az elmeállapotokra vonatkozó empirikus törvény lehet, egyfajta koordinációs elv az objektív és szubjektív valószínűségek között (Ld. 52. oldal).

[3] A determinizmus kérdéséhez jó bevezetőt nyújt (Earman, 1986).

Mi a helyzet a valószínűséggel egy determinista világban? Az objektív valószínűség itt nyilván nem lehet a világ objektív meghatározatlanságát mérő mennyiség. Lehet azonban relatív gyakoriság. Így persze a szinguláris eseményeknek nem lesz valószínűsége, csak az eseményosztályoknak, de ettől a valószínűség még tökéletesen objektív lesz. A szubjektív valószínűségek elhelyezése egy ilyen világban szintén nem jelent problémát: ha a világ eseményei rögzítve vannak is előre, mi még nem feltétlenül tudjuk, hogy mit hoz a jövő – vagyis hitek továbbra is skálázhatók a meggyőződés foka mentén, és ezt a részleges hitet nevezhetjük szubjektív valószínűségnek. Vagyis azt látjuk, hogy az objektív és a szubjektív valószínűség elhelyezésével egy determinista világban sincs probléma.

Hogy a világ egymást követő állapotai meghatározzák-e egymást, hogy létezik-e a világban objektív valószínűség, illetve hogy szubjektív hitállapotaink hogyan követik a világ történéseit – ez három független kérdés. Az objektív valószínűség csak akkor lehetetlen egy determinista világban, ha ragaszkodunk ahhoz, hogy az objektív valószínűséget egyfajta kvantifikált modalitásnak tekintsük, a lehetséges történések fokmérőjének. Ez azonban nem szükségképpen következik az objektivitás fogalmából. Másfelől az a két tény, hogy világunk indeterminista, és hogy tudásunk korlátozott, együtt sem azt nem involválja, hogy a világban léteznie kell valószínűségnek, sem azt, hogy amennyiben létezik, akkor az szubjektív. Lehetséges, hogy valószínűség egyáltalán nem létezik ebben a világban, és előfordulhat az is, hogy csak objektív valószínűség létezik.

Összefoglalva megállapíthatjuk, hogy a valószínűség klasszikus definíciójához vezető laplace-i gondolatmenet első két pontja – ti., hogy 1. egy determinista világban minden valószínűség pusztán az ismerethiányból fakadhat, és 2. így szubjektív – hamis. Ettől természetesen a klasszikus meghatározás, amely szerint a valószínűség a kedvező és az összes egyenlően lehetséges eset aránya, még helyes lehet. Mielőtt magát a definíciót megvizsgálnánk, nézzük az elv egy fontos alkalmazását, *Laplace rákövetkezési szabályát*.

4.2. A rákövetkezés szabálya

A rákövetkezés szabálya szerint ha egy esemény ezidáig egymástól függetlenül n-szer bekövetkezett, és az esemény valószínűségéről semmit sem tudunk, azaz ha az esemény valószínűségére vonatkozó minden feltevésünk egyformán valószínű, akkor annak valószínűsége, hogy a következő alkalommal is bekövetkezik $\frac{n+1}{n+2}$. A szabály matematikai alapja az alábbi tétel jelenti.

3. Tétel. Legyen (Ω, Σ, p) valószínűségi mértéktér, és legyenek $g_1, g_2, \ldots g_{n+1}$ feltételesen független Bernoulli-változók a h egyenletes eloszlású valószínűségi változóra nézve. Ekkor

$$p(g_{n+1} = 1 \mid g_1 + \cdots + g_n = s) = \frac{s+1}{n+2}.$$

Bizonyítás. Lásd a D. Függelékben.

Az s helyébe n-et írva megkapjuk a rákövetkezés szabályát. A szabályt Laplace híres példájában a felkelő Nap esetére alkalmazza: ha a Nap 5000 éves történelmünk során eddig minden reggel felkelt, akkor annak valószínűsége, hogy holnap is fel fog kelni $0,9999994$.

Bár a fenti tétel pusztán egy matematikai tétel, és így mentes mindenfajta metafizikai megfontolástól, mégis érdemes megvizsgálni, hogy a bizonyítás matematikája miképpen sugallt egy bizonyos metafizikai értelmezést. Ehhez tekintsük a bizonyításnak egy egyszerűbb formáját, rögtön az $s = n$ esetre.

Tegyük fel, hogy egy urnában fekete és fehér golyók vannak ismeretlen arányban. Jelöljük h_x-szel azt a hipotézist, hogy ez az arány x. Az egyszerűség kedvéért engedjük meg, hogy ez az x arány bármely 0 és 1 közötti valós szám lehessen, vagyis, hogy a golyók száma az urnában legyen nagyon nagy. Mivel az arányról semmit sem tudunk, ezért az indifferencia elve alapján minden $x \in [0,1]$-hez tartozó h_x hipotézis $p(h_x)$ episztemikus valószínűsége azonos. Ezt a valószínűségeket Laplace óta a hipotézis *a priori* valószínűségének nevezzük.

Mekkora a valószínűsége annak, hogy az urnából a következő húzásra fekete golyót kapunk, ha a h_x hipotézis igaz, vagyis ha a fekete és fehér golyók aránya valóban x? Nyilvánvalóan ez a $p(g_1|h_x)$ feltételes valószínűség egyenlő x-szel. Most tekintsük azt a g_n eseményt, hogy az urnából egymás után n-szer fekete golyót húzunk, ismét csak feltéve, hogy az urnában a fekete golyók aránya a fehérekhez x. Mivel a húzások egymástól függetlenek, így az g_n esemény $p(g_n|h_x)$ feltételes valószínűsége x^n. A levezetéshez még egy feltételes valószínűségre van szükségünk. Jelölje g_{n+1} azt az eseményt, hogy az $n+1$-ik húzás eredménye is fekete. Tekintsük a $p(g_{n+1}|g_n \wedge h_x)$ feltételes valószínűséget, amely azt fejezi ki, hogy milyen valószínűséggel kapunk fekete golyót erre az $n+1$. húzásra, feltéve, hogy a fekete golyók aránya a fehérekhez az urnában x, és az eddigi n húzás során csak fekete golyót kaptunk. Mivel a húzások függetlenek, ezért ez a valószínűség x lesz.

Ezek után tekintsük a $p(h_x|g_n)$ úgynevezett *a posteriori* valószínűséget. Az elnevezés ismét csak Laplace-tól származik, és a h_x hipotézis „új” valószínűségét fejezi ki, *miután* megbizonyosodtunk arról, hogy az összes eddig kihúzott n golyó fekete. Ezt a feltételes valószínűséget a Bayes-tétel folytonos alakját alkalmazva (pontosabban lásd D. Függelék) a következő módon határozhatjuk meg:

$$p(h_x|g_n) = \frac{p(g_n|h_x)p(h_x)}{\int_0^1 p(g_n|h_y)p(h_y)dy} = \frac{p(g_n|h_x)}{\int_0^1 p(g_n|h_y)dy} = \frac{x^n}{\int_0^1 y^n dy} = (n+1)x^n,$$

ahol a második egyenlőség a különböző h_x hipotézisek egyenlő valószínűsége miatt teljesül, a harmadik egyenlőtlenség pedig a fenti $p(g_n|h_x) = x^n$ összefüggés miatt. Végül tekintsük a keresett $p(g_{n+1}|g_n)$ kifejezést, amely az $n+1$. fekete golyó kihúzásának valószínűségét adja meg feltéve, hogy az addigi n húzás mind fekete volt. Alkalmazzuk $p(g_{n+1}|g_n)$-re a teljes valószínűség tételének folytonos alakját:

$$p(g_{n+1}|g_n) = \int_0^1 p(g_{n+1}|g_n \wedge h_x)p(h_x|g_n)dx = \int_0^1 x(n+1)x^n dx = \frac{n+1}{n+2},$$

ahol is a második egyenlőségnél a $p(g_{n+1}|g_n \wedge h_x) = x$ és a $p(h_x|g_n) = (n+1)x^n$ összefüggéseket használtuk. Az eredmény a bizonyítani kívánt rákövetkezés elve: ha eddig n-szer húztunk fekete golyót, akkor – semmit sem tudva az urnában levő fekete és fehér golyók arányáról – annak valószínűsége, hogy a következő húzásra feketét kapunk $\frac{n+1}{n+2}$. Ha eddig még nem húztunk az urnából, akkor a fekete golyó húzásának valószínűsége $\frac{1}{2}$, ha húztunk már egy feketét, akkor $\frac{2}{3}$, ha kettőt, akkor $\frac{3}{4}$; vagyis minél több fekete golyót húztunk eddig annál biztosabbak lehetünk abban, hogy a következő húzásra is feketét kapunk.

A bizonyítás, legalábbis ebben a metafizikai hangszerelésben három dolgot használt ki: (i) Létezik egy x objektív valószínűség, amely az urnából való húzás (a Nap felkelésének) valószínűsége; (ii) Az urnából történő húzások független események; (iii) Az x objektív valószínűség értékére vonatkozó h_x hipotézis valószínűsége minden $x \in [0,1]$ valós számra vonatkozóan egyforma. Ebben a fejezetben mi csak a harmadik feltevésre, az indifferencia elvére térünk ki, a 6. fejezetben azonban visszatérünk a másik két feltevésre is.

4.3. Az indifferencia elve

A rákövetkezés szabálya nyilvánvalóan abszurd.[4] De miben is rejlik az abszurditás?

A rákövetkezés szabálya bizonyításának sarokpontját az a feltevés jelentette, hogy az urnában a fehér és fekete golyók minden x arányára vonatkozó hipotézis egyenlően valószínű (vagy a fenti tételben, hogy f valószínűségi változó egyenletes eloszlású). Ezt a feltételezést a klasszikus elképzelés azzal támasztja alá, hogy a hipotézisek „egyenlően lehetségesek". A valószínűség tehát a lehetőségeket méri. A lehetőség itt – a valószínűséghez hasonlóan – háromféleképpen érthető:

1. *Objektív lehetőség.* A lehetőségek az objektív világ eseményeire vonatkoznak.

2. *Szubjektív lehetőség.* A lehetőségek az empirikus elme hitállapotaira vonatkoznak.

3. *Logikai lehetőség.* A lehetőségek az ideális (racionális) elme hitállapotaira vonatkoznak.

Az objektív lehetőségeket kétféleképpen értelmezhetjük: vagy modálisan vagy aktuálisan. Modálisan értve az objektív lehetőség az indeterminista világnak a természettörvények által megengedett, de mégsem aktualizálódott történései. Ha a világ törvényei nem tiltják, hogy a kockadobás kimenete hatos legyen, de az történetesen ötös, akkor a hatos egy objektív lehetőség. A lehetőséget azonban aktuálisan is érhetjük: akkor lehetséges, hogy a dobás hatos legyen, ha az aktuális világ (hasonló körülmények között elvégzett) dobásai között van hatos. Mint majd látni fogjuk, a lehetőségnek ez a két felfogása a *propensity*-interpretációhoz illetve a relatív gyakoriság-interpretációhoz vezet.

[4] Popper az elven eképpen élcelődik. Képzeljük el, hogy valamilyen kozmikus oknál fogva egy szép napon (éjszakán) egyszer csak leáll a Föld forgása, és reggel nem kel fel a Nap. Mi a valószínűsége, hogy a következő napon ismét felkel? A rákövetkezés szabálya alapján számolva $0,9999989$.

Szubjektív és logikai lehetőségen az empirikus illetve az ideális elme számára szóba jövő hitállapotokat értjük. Ebben az értelemben „egyenlően lehetséges” két esemény akkor, ha az empirikus illetve az ideális elme nem képes közöttük dönteni.

Laplace idézeteiből úgy tűnhet, hogy a klasszikus elmélet „lehetőség” alatt szubjektív vagy logikai lehetőséget értett. Daston (1988) szerint azonban az 1840-es évek előtti asszocionista ismeretelmélet, amely egy benyomás erejét és tisztaságát a benyomás gyakoriságával azonosította, feleslegessé tette az objektív és a szubjektív/logikai oldal elkülönítését. Mint írja:

> „A hitet és a valószínűséget a tapasztalat generálta az érzékelés ismételt korrelációi során, amelyet azután az elme ideák asszociációjában reprodukált. Minél szilárdabb és gyakoribb a megfigyelt korreláció, annál erősebb a mentális kapcsolat, amely viszont a valószínűséget és a hitet erősíti. Így a tapasztalat objektív valószínűsége és a hit szubjektív valószínűsége egy rendezett elmében egymás tükörképei lettek. Ezért lehetett megbízni a széles tapasztalatokra épített intuitív ítéletekben. Ha a klasszikus probabilisták a racionális embert tették meg standardnak, ez részben azért történt, mert racionalitása lényegében valószínűségi jellegű volt.” (197. o.)

A valószínűség a 19. század közepéig tehát egyfajta „Janus-arcú” (Hacking, 1975) fogalom volt: egyaránt mutatott aleatorikus és szubjektív jegyeket is. Az asszocionista ismeretelmélet leáldozásával azonban a „lehetőség” és vele együtt a „valószínűség” fogalma is a fenti elágazó értelmezések alá került, és a klasszikus interpretációt egyre inkább a szubjektív-logikai oldalról kezdték olvasni. Ezt a folyamatot koronázta meg az a mozzanat, amikor Keynes (1921) az elégségtelen ok laplace-i elvét az *indifferencia elvének* keresztelte át – mutatva ezzel, hogy a kulcsfogalom itt az elme pártatlansága. A meghatározása így hangzik:

> „Az indifferencia elve azt állítja, hogy amennyiben nincs *ismert* okunk alternatívák közül az egyiket előnyben részesíteni a másikhoz képest, akkor erre a tudásra vonatkozóan mindegyik alternatívának egyenlő a valószínűsége.” (42. o.)

A dobókocka példájára lefordítva az elv azt mondja ki, hogy amennyiben a kockát és az eldobás körülményeit minden szempontból megvizsgáltuk (vagy az ideális elme minden szempontból megvizsgálta), és egyik kimenetet sem vagyunk képesek kitüntetni a többivel szemben, akkor mind a hat kimenetet egyenlő, azaz egy hatod valószínűségűnek kell tartanunk. De miféle elvről van itt szó? Mit jelent az, hogy a kockát minden szempontból megvizsgáltuk, és hogy nem tudtuk egyik kimenetet sem kitüntetni? És ha ez így van is, miért kellene ezek után az összes kimenetet egyenlően valószínűnek tartanunk?

Haladjunk sorban! Az indifferencia elvének érvénytelenségéről általában egyetértés uralkodik az irodalomban, abban azonban koránt sincs egyetértés, hogy miért is érvénytelen ez az elv. A szokásos cáfolatok matematikai jellegűek. Ehhez első lépésben az indifferencia elvének egyfajta matematikai olvasatát adják az alábbi módon. Vegyük a kocka eseményterének megfelelő Boole-algebrát. *Tegyük fel*, hogy az empirikus vizsgálódásokon túl vagyunk, és az algebra elemeivel reprezentált eseményeket *fizikailag* valamiképp „egyenlően lehetségesnek” találtuk. Hogy ez mit jelent, arra a későbbiekben még visszatérünk. Ekkor az események valószínűségének meghatározásához nem kell mást tennünk, mint megszámolni az algebra atomjait, és ennek a számnak a reciprokát mint valószínűséget rendelni minden atomi eseményhez.

A szabály megszámlálhatóan végtelen esetben a σ-additivitás miatt nem alkalmazható, mivel az atomokhoz rendelt minden véges valószínűség sérti a mérték normáltságát. Ezen azonban segíteni lehet a σ-additivitás feladásával.[5] Nem megszámlálhatóan végtelen esetben azonban – a bevett nézet szerint – az elv szépen gyümölcsöztethető matematikai valószínűségi feladatok megoldásához. A történetileg elsőként megoldott ilyen „geometriai valószínűségi” probléma Buffon tű-problémája, amely Georges-Louis Leclerc de Buffontól származik 1777-ből.

Buffon tű-problémája. Tegyük fel, hogy egy párhuzamos parkettacsíkokból álló padlóra „véletlenszerűen” tűket dobálunk. A csíkok szélessége és a tű hossza legyen egyaránt L. Mi a valószínűsége annak, hogy a tű ráesik valamelyik, két parkettacsíkot elválasztó vonalra?

A feladat megoldásához szokás bevezetni két valószínűségi változót: X-et, amely a tű középpontjának távolságát méri a parkettacsíkokat elválasztó legközelebbi vonaltól, és Φ-t, amely pedig a tűnek a parkettacsíkokat elválasztó vonallal bezárt szögét. Követeljük meg továbbá, hogy ezek a valószínűségi változók függetlenek legyenek valamint egyenletes eloszlásúak a megfelelő intervallumokon. Ezek után a feladat megoldása egyszerű kétváltozós integrálás,[6] amelynek eredménye $\frac{2}{\pi}$.

Az indifferencia elve a feladatban állítólagosan a valószínűségi változók egyenletességének követelményében van kihasználva, vagyis ott, hogy a valószínűségi változók sűrűségfüggvényére előírjuk az alábbiakat: $f_X(x) = \frac{2}{L}$ minden $x \in [0, \frac{L}{2}]$-re, és $f_\Phi(\phi) = \frac{2}{\pi}$ minden $\phi \in [0, \frac{\pi}{2}]$-re. Más szóval az indifferencia elve ebben a matematikai olvasatban egyszerűen azt jelenti, hogy a feladatot jellemző valószínűségi változók sűrűségfüggvényei a megfelelő intervallumon konstansok.

Előfordulhat azonban, hogy a valószínűségi változók egy feladatban nincsenek vagy nincsenek egyértelműen megadva. Tipikusan ez a helyzet azokban a matematikai feladatokban, amelyekben egy kezdeti geometriai problémát hirtelen valószínűségi problémaként kezdünk

[5] Ehhez ld. még de Finetti érveit a σ-additivitás ellen.

[6] Vagyis:

$$\begin{aligned} p(X < \frac{L}{2}\cos\Phi) &= \iint_{x < \frac{L}{2}\cos\phi} f_X(x) f_\Phi(\phi)\, dx d\phi = \frac{4}{\pi L}\int_0^{\frac{\pi}{2}}\int_0^{\frac{L}{2}\cos\phi} dx d\phi \\ &= \frac{4}{\pi L}\int_0^{\frac{\pi}{2}} \frac{L}{2}\cos\phi\, d\phi = \frac{2}{\pi} \end{aligned}$$

értelmezni. És pontosan ezek azok az esetek, amelyeket a szakirodalom az indifferencia elvének cáfolataiként tart számon. A legismertebb ilyen „paradoxon” Buffon tű-problémájának egyfajta transzformációja, a Bertrand-paradoxon.

Bertrand-paradoxon. A paradoxon Joseph Bertrand-tól származik, aki *Calcul des probabilités* (1889) című munkájában a paradoxont annak illusztrálására vezette be, hogy egy esemény valószínűsége egy feladatban rosszul definiált is lehet; vagyis, hogy egy geometriai feladat vázolása még nem ad eligazítást arra nézve, hogy a feladatban szereplő mennyiségekhez egy valószínűségelméleti olvasatban milyen valószínűségeket rendeljünk. A filozófiai szakirodalomban azonban a Bertrand-paradoxon az indifferencia elvének és így a valószínűség klasszikus interpretációjának cáfolataként terjedt el.[7]

A feladatban szintén tűket dobálunk „véletlenszerűen”, de ezúttal nem egy csíkozott padlóra, hanem egy szabályos körlapra. Csak azokat az eseteket tekintjük, amikor a tű két részre vágja a kört, vagyis egy húrt képez a körön. Bertrand kérdése ezek után a következő: mi a valószínűsége annak, hogy egy ilyen dobás során létrejött húr hosszabb lesz, mint a körbe írható szabályos háromszög oldala?

A paradoxon abban áll, hogy – Buffon tű-problémájával ellentétben – a „véletlenszerűen” kifejezés itt többféleképpen is értelmezhető, és a különböző értelmezések mentén a fenti kérdésre három helyes, ugyanakkor egymásnak ellentmondó válasz is adható. Nézzük őket sorban. Mindhárom válasz valamilyen egyenletes eloszlásra hivatkozik. Véletlenen egyfelől érthetjük azt, hogy a húrok távolsága a kör középpontjától egyenletes eloszlású. Ekkor a kérdésre a válasz $\frac{1}{2}$ lesz, mivel a szabályos háromszög oldalánál hosszabb és rövidebb húrokat épp a háromszögnek a középponttól $\frac{r}{2}$ távolságra levő oldala választja el egymástól (ld. 4.1. ábra). Másrészt érthetjük a véletlent úgy is, hogy a húr és a körvonal közötti szög véletlen,

[7] Hasonló paradoxonok tucatja kering az irodalomban. Itt csak Keynes (1921) két ismert paradoxonát említjük:

Könyv-paradoxon. Egy sosem látott könyv színét próbáljuk eltalálni.

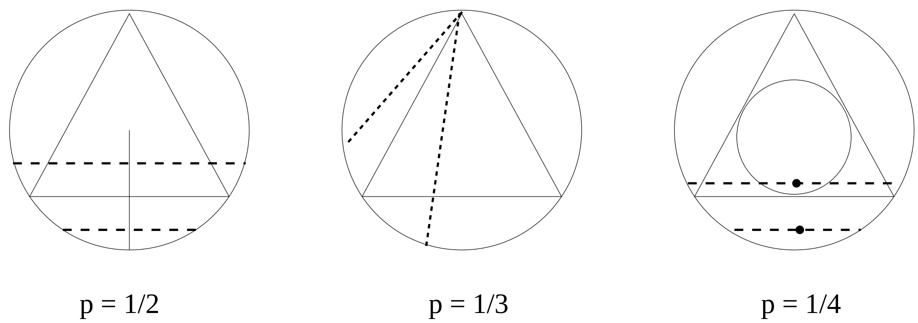
> „Ha például nincs releváns bizonyítékunk a könyv színét illetően, akkor $\frac{1}{2}$ valószínűséggel következtethetünk arra, hogy 'A könyv piros'. De éppolyan joggal következtethetünk $\frac{1}{2}$ valószínűséggel arra is, hogy 'A könyv fekete' és, hogy 'A könyv kék'. Így azonban három olyan kizáró alternatívánk lesz, amelyek közül mindegyik ugyanolyan valószínűséggel igaz, mint amennyire nem.” (46. o.)

Bor/víz-paradoxon. Egy bor-víz keverékről csak annyit tudunk, hogy egyik folyadék sem háromszor több a másiknál. Mekkora a valószínűsége annak, hogy a bor/víz arány kisebb kettőnél?

Az állítólagos paradoxon abból adódik, hogy a kérdésre két különböző válasz adható. A bor/víz arány a feladat kikötése szerint az $[\frac{1}{3}, 3]$ intervallumból vehet fel értékeket. Az indifferencia elve alapján az intervallum minden értéke egyenlően lehetséges, azaz a valószínűségi eloszlás az intervallumon uniform. Ezen az intervallumon belül a 2-nél kisebb bor/víz arányhoz tartozó értékek mértéke $\frac{5}{6}$.

Tekintsük azonban a feladatot fordítva, a víz/bor arány szerint. Ennek lehetséges értékei ismét csak az $[\frac{1}{3}, 3]$ intervallumba esnek, az eloszlás pedig ismét csak uniform. A bor/víz arány akkor és csak akkor kisebb kettőnél, ha a víz/bor arány nagyobb $\frac{1}{2}$-nél. Az $[\frac{1}{3}, 3]$ intervallumban az $\frac{1}{2}$-nél nagyobb intervallum mértéke azonban $\frac{15}{16}$. Vagyis ugyanannak az eseménynek a feladat értelmezésétől függően két különböző valószínűsége is van. (48. o.)

azaz egyenletes eloszlású. Ekkor a kérdésre a válasz $\frac{1}{3}$, mivel a $(0, \pi)$ szögtartományból csak a $(\frac{\pi}{3}, \frac{2\pi}{3})$ rész esetében lesz a húr hosszabb a szabályos háromszög oldalánál. Ha pedig véletlenen azt értjük, hogy a húrok középpontjai, amelyek egy-egyértelműen megfeleltethetők a húroknak, egyenletesen oszlanak el a körlapon, azaz a területmértéket követik, akkor a válasz $\frac{1}{4}$, mivel az etalonnál hosszabb húrhoz tartozó középpontok egy fele akkora sugarú nyílt körlapon helyezkednek el. A Bertrand-paradoxon, hangzik az érvelés, tehát azért cáfolja az indifferencia elvét, mivel pusztán abból az információból, hogy a dobások véletlenszerűek, vagyis egyiket sem tudjuk preferálni közülük, a kérdéses esemény valószínűségére nem kapunk egyértelmű választ. Más szóval a lehetséges eseményekkel szembeni ignoranciánk nem vezet egyértelmű valószínűségekre az események tekintetében.



4.1. ábra. Bertrand-paradoxon

A paradoxonnal foglalkozó szerzők többsége, Boreltől von Misesig azon a véleményen volt, hogy a probléma rosszul definiált, vagyis egy olyan matematikailag alulhatározott feladattal állunk szemben, mintha mondjuk háromszöget kellene szerkesztenünk két oldal ismeretében. E. T. Janes (2003) azonban amellett érvelt, hogy a feladat igenis jól definiált, és a három megoldás közül csak az első helyes. Érvelésében Janes az indifferencia elvére hivatkozik. Azt állítja ugyanis, hogy a probléma vázolásakor közelebbről meg nem határozott részletek a feladat szimmetriatulajdonságaira utalnak, vagyis a tudás hiánya egyfajta *invarianciatulajdonságként* értelmezhető. A feladatnak három ilyen szimmetriatulajdonsága van. Mivel a feladat nem beszél a kör orientációjáról, a megoldásnak nyilvánvalóan forgásszimmetrikusnak kell lennie. A forgásszimmetria azonban nem dönt a kérdésben, mivel mindhárom megoldás (pontosabban a megoldásokhoz vezető eloszlásfüggvény) forgásszimmetrikus. A másik nem specifikált tulajdonság a kör mérete: ha létezik megoldás, akkor annak skálainvariánsnak kell lennie. Ez a szimmetriakövetelmény kizárja a második esetet. A harmadik ignoranciánk a kör helyére vonatkozik, amelyből eltolási invariancia adódik. Ez a szimmetria már egyértelműen kiválasztja a helyes megoldást (eloszlásfüggvényt), sőt pusztán ezt az invarianciát használva a másik kettő nélkül is kizárhatjuk a $\frac{1}{3}$ és $\frac{1}{4}$ válaszokat. A feladat hallgatólagos feltevései tehát csoportelméleti alapon meghatározzák a helyes megoldást – állítja Janes.

A gondolatmenet pikantériáját azonban a fejezet utolsó paragrafusa (393-394. o.) adja, ahol is Janes miután igen kifinomult csoportelméleti módszerekkel meggyőzi olvasóját az első valószínűségi eloszlás helyességéről, mintegy az előzőek empirikus bizonyítékaként még azt is hozzáteszi, hogy egy álló helyzetből egy öt hüvelyk átmérőjű a padlóra rajzolt körlapra szalmaszálakat dobáló kísérlet „zavarba ejtő” pontossággal megerősítette nézetét.

De miért is van szüksége Janes-nek erre az empirikus verifikációra? Ha az ignoranciának ez az invarianciába való átfordítása ilyen egyértelműen kiválasztja a helyes valószínűségeket (és amint Janes jósolja, minden valószínűségi paradoxonnal ez a helyzet), akkor mi keresnivalója van itt még a tapasztalatnak? Miért kell egyáltalán szalmaszálat vennünk a kezünkbe?

Valójában ez az utolsó kis paragrafus roppant árulkodó az indifferencia elvének szokásos cáfolatai és a cáfolatok semlegesítésére irányuló törekvések elhibázottságát illetően. Miről is van szó?

Való igaz, hogy az a heurisztika, amely alapján adott halmazalgebrához úgy rendelünk valószínűséget, hogy az algebra minden atomjához azonos mértéket rendelünk, véges esetben egyértelmű megoldáshoz vezet, és ezt éppenséggel tekinthetjük valamilyen értelemben „természetes”, „szimmetrikus” módszernek. Mindezt azonban úgy értelmezni mint az indifferencia elvének érvénytelenségét vagy érvényességét, teljességgel jogtalan. Az indifferencia elve – akár igaz, akár nem – *fizikai eseményekről* állít valamit, nem pedig geometriai pontokról és húrokról. Ha ezeket a húrokat tűknek vagy szalmaszálaknak *nevezzük*, és a feladat matematikai jellegét elkendőzzük holmi szöveges feladattal, attól a feladat még matematikai feladat marad. Vagyis az indifferencia elve nem bizonyítható vagy cáfolható azzal, hogy bizonyos algebrákban létezik vagy nem létezik egyértelműen valamilyen „természetesnek” mondott mérték. A Bertrand-paradoxon semmivel sem cáfolja jobban az indifferencia elvét, mint Buffon tű-problémája. Mindkét esetben, ha tetszik, végtelen sok különböző eloszlásfüggvényt választhatok, amelyek mindegyike egy-egy valószínűségi feladatot határoz meg, és egy adott megoldáshoz vezet. Az pedig, hogy Buffon tű-problémájánál ezek közül a valószínűségi feladatok közül az egyiket *intuitíve* a véletlen dobások helyes modelljének tartom, míg a Bertrand-paradoxon esetében az intuícióm nem súg semmit, nem jelent sem érvet, sem ellenérvet az indifferencia elvével szemben.

Az indifferencia elve tehát nem igazolható és nem is cáfolható *matematikailag*. De akkor hogyan dönthető el az igazsága?

Az indifferencia elve a valószínűség fogalmát a lehetőség fogalmához köti: két eseménynek akkor egyezik meg a valószínűsége, ha a két esemény „egyenlően lehetséges”. Most feledkezzünk meg egy pillanatra arról a kérdésről, hogy ez az elv segítségével hogyan is tudnánk *nem* egyenlően lehetséges eseményekhez valószínűséget rendelni, és tegyük fel rögtön a lényegi kérdést: mit is jelent az elvnek az a központi kifejezése, hogy két esemény „egyenlően lehetséges”? A válaszokat tekintsük végig a „lehetőség” fogalmának az alfejezet elején adott három értelmének megfelelően!

Ha „lehetőség” alatt objektív lehetőséget értünk, akkor az „egyenlően lehetséges” értelmezésére a következőkkel próbálkozhatunk. (Ld. Szabó 2002, 67. o.) Két esemény akkor egyenlően lehetséges, ha *minden fizikai tulajdonságukban* megegyeznek. Ez az értelmezés azonban túlságosan erős. Ha ugyanis két esemény minden tulajdonságában megegyezik, akkor a Leibniz-elvnek megfelelően nem is tudjuk megkülönböztetni őket. Elvégre a kockán legalább a pöttyök számának különbözni kell ahhoz, hogy tudjuk, milyen dobásról is van szó.

Próbáljuk meg gyengíteni az értelmezést. Két esemény egyenlően lehetséges, ha *fizikai tulajdonságaik egy meghatározott halmazán* megegyeznek. A kockadobás két kimenete akkor egyenlően lehetséges, ha a kocka geometriai formája, tömegeloszlása vagy az eldobás egyéb körülményei stb. nem tüntetik ki egyik oldalt sem a másikkal szemben. Másképp szólva, ha a kocka szimmetriái a relevánsnak tekintett fizikai tulajdonságoknak is szimmetriái. De mi történik akkor, ha egy releváns tulajdonságaiban szimmetrikusnak tartott kocka esetében azt tapasztaljuk, hogy mondjuk minden dobásra hatost kapunk. Vajon továbbra is kitartanánk amellett, hogy a kockadobások „egyenlően lehetségesek"?

Most térjünk át a „lehetőség" második, szubjektív értelmére. Itt két esemény akkor egyenlően lehetséges, ha *az empirikus elme egyik eseményt sem preferálja a másikkal szemben*, egyik esemény bekövetkezésére vonatkozóan sincs több tudása, mint a másikére. Hogy az „egyenlően lehetséges" fogalmának meghatározását az empirikus ágensre bízni nem a legjobb döntés, álljon itt példaként az alábbi két paradoxon.

Három kártya-paradoxon. Adva van három kártya, az egyik mindkét oldalán piros, a másik mindkét oldalán kék, a harmadik az egyik oldalán piros, a másik oldalán kék. Az osztó megkeveri a kártyákat, kihúz egyet, majd elénk teszi az asztalra. A kártya felső lapja piros, a túloldal színét nem látjuk. Ezek után megkérdezi tőlünk, hogy a túloldal színét illetően inkább a kékre vagy a pirosra szavazunk-e?

A tapaszlatok szerint a megkérdezettek többsége így okoskodik: a kihúzott kártya nyilván nem lehet a kék-kék, így vagy a piros-kék vagy a piros-piros. Vagyis a túloldali kék és piros „egyenlően lehetséges" esetek.

A helyes válasz azonban következő. A fenti eset háromféleképpen tud megvalósulni. Vagy a piros-piros kártya egyik oldala fekszik előttünk, vagy a piros-piros kártya másik oldala, vagy a piros-kék kártya piros oldala. Feltételezve a húzás véletlenszerűségét, a túloldali piros kétszer gyakrabban fordul elő a túloldalon, vagyis a két eset nem egyenlően lehetséges.

Monty Hall-paradoxon. A paradoxon döntéselméleti változata annak az amerikai TV-show műsorvezetőjéről kapta a nevét, amely show-ban először fogalmazták meg. A show-ban három ajtót mutatnak egy játékosnak. A három ajtó közül az egyik mögött egy autó van elrejtve, a másik kettő mögött pedig egy-egy kecske. A játékos kiválaszt egy ajtót, amely azonban továbbra is zárva marad. Ezután a játékvezető a maradék két ajtó közül kinyitja az egyiket, mindig úgy, hogy az ajtó mögött kecske legyen. Ezt a játékos választásától függetlenül mindig megteheti. Ezután megkérdezi a játékost, hogy továbbra is kitart-e a választott ajtó mellett, vagy áttér a harmadik ajtóra. A játékosok zöme szerint mindegy volt, hogy maradnak-e eredeti választásuknál vagy áttérnek, mivel a két zárt ajtó közül bármelyikben egyenlő lehetséges az autó. A helyes stratégia azonban az áttérés. Ennek racionalitása akkor látszik jól, ha növeljük az ajtók számát. Szerepeljen a játékban három helyett 100 ajtó. Miután egyet kiválasztottunk, a játékvezető kinyit 98 ajtót, amelyek mögött kecske van. Mivel 100-ból 99 esetben elsőre nem a nyereményt rejtő ajtót választottuk, ezért az áttérés 99 esetben a nyereményhez vezet, és csak egy esetben veszítjük el a nyereményt.

Természetesen mindkét paradoxon esetében az empirikus ágens becslésének és a kombinatorikailag kiszámolt lehetőségeknek az eltérése csak annyiban jelent ellenérvet az „egyenlően lehetséges" szubjektív definíciója ellen, amennyiben feltételezzük, hogy a kombinatorikailag egyenlő lehetőségek tükröződnek a relatív gyakoriságokban; magyarán a kártyákat jól megkevertük, illetve a nyereményt véletlenszerűen tettük valamelyik ajtó mögé. De az empirikus ágens még ebben az esetben is ragaszkodhat az „egyenlően lehetséges" szubjektív definíciójához, ha hajlandó beletörődni abba, hogy valószínűségfogalma nem fog kvadrálni a relatív gyakoriságokkal.

Az „egyenlően lehetséges" szubjektív értelmezésével azonban nem ez a legnagyobb baj. A legnagyobb probléma az, hogy ha a valószínűség fogalmát az „egyenlően lehetséges" fogalmából, ez utóbbit pedig empirikus elme pártatlanságából eredeztetjük, akkor az indifferencia elvével kapcsolatos vesződséget akár meg is spórolhatjuk magunknak, és a valószínűséget egyszerűen magával az empirikus elme hitének erejével azonosíthatjuk. Ebben az esetben azonban elhagyjuk a klasszikus interpretációt, és átlépünk a szubjektív interpretációba.

Maradt a „lehetőség" fogalmának utolsó értelmezése, a logikai. Erre az értelmezésére támaszkodva az indifferencia elve így hangzik: Két esemény egyenlően lehetséges akkor, ha *az ideális (racionális) elme egyiket sem preferálja a másikkal szemben.* A definíció Akhilleusz-pontja az „ideális elme" kifejezés. Mi is az az ideális elme? Hogyan is verifikálható, hogy két esemény egy ideális elme számára egyenlően lehetséges. Az indifferencia elve ebben a logikai esetben leginkább egyfajta normatív állításként értendő: ha nincs releváns információd két esemény bekövetkezésére vonatkozóan, akkor a két eseményt *tartsd* egyenlően lehetségesnek. Hogy ez az állítás milyen szerepet tölt be a normatív állítások körében, azt nem tudjuk; mindenesetre a valószínűség fogalmának empirikus meghatározásához nem visz bennünket közelebb.

Összefoglalva, az indifferencia elvében szereplő „egyenlően lehetséges" fogalmát a „lehetőség" sem objektív, sem szubjektív, sem pedig logikai fogalmával nem lehet értelmessé tenni. Ez viszont azt jelenti, hogy a klasszikus interpretáció nem tartható.

5. fejezet

A logikai interpretáció

A valószínűség logikai interpretációja a klasszikus interpretáció örököseként szintén abból az elgondolásból indul ki, hogy a valószínűség a lehetőségek terének a priori vizsgálatával meghatározható. A logikai interpretáció az indukció formális tárgyalását célzó törekvések nyomán született. A dedukció pontos fogalmának tisztázása a formális nyelvekben a századfordulón azzal kecsegtetett, hogy alkalmasint az indukció is megfogalmazható hasonló szabatossággal, mintegy a dedukció általánosításaként. A deduktív következtetéshez hasonlóan, amely viszonyt teremt a nyelv bizonyos mondatai között, egy olyan $c(h, e)$ függvény megadására törekedtek, amely azt az intuíciónkat fejezné ki, hogy a nyelv egy e mondata, az evidencia, adott mértékben konfirmál egy másik h mondatot, a hipotézist. A konfirmáció mértéke 0 és 1 közé esik, határesetben a deduktív következtetést is magában foglalva. A valószínűség ezek után a konfirmációs függvényből származtatott egyfajta induktív súly a nyelv mondatain (pontosabban a nyelv Lindenbaum–Tarski-algebráján). A kérdés tehát az volt, hogy vajon kiróhatóak-e a konfirmációs függvényre az indukció intuitív sajátosságait szem előtt tartó olyan „természetes” követelmények, amelyek azután a nyelven mintegy visszamenőleg valószínűségi mértékhez vezetnek.

A valószínűség logikai interpretációja egészen Leibnizig nyúlik vissza.[1] Tizenkilencedik századi előfutárai August De Morgan (1848), George Boole (1852) és William Stanley Jevons (1973), huszadik századi képviselői pedig John Maynard Keynes, Harold Jeffreys és Rudolf Carnap. Keynes 1913-ban fogott hozzá a *Treatise on Probability* írásához, de a könyv a

[1] Ian Hacking (1971) így ír erről:

> „Leibniz volt az első, aki ragaszkodott ahhoz az elképzeléshez, hogy a valószínűségelmélet a logika egyik ágaként a dedukció elméletéhez hasonló szerepet tölthet be. Ő volt az első, aki a valószínűséget mint önálló tudományt axiomatizálni kívánta. Felismerte, hogy az általánosított játékelmélet hogyan szolgálhat alapul az olyan helyzetekben való kvantitatív döntésekhez, ahol inkonkluzív bizonyítékok alapján kell cselekednünk. Jócskán megelőzve korát azt hirdette, hogy a valószínűség relációs tulajdonság, és hogy a valószínűségi ítéletek a rendelkezésre álló adatokra nézve relatívak.” (597. o.)

háború miatt csak 1921-ben jelent meg. Jeffreys 1939-ben adta közre a *Theory of Probability* című művét, amely sok tekintetben a modern bayesianizmus alapköve. Carnap a 40-es évektől kezdve foglalkozott a valószínűség megalapozásával. Fő műve a *Logical Foundations of Probability* 1950-ben jelent meg, amelynek gondolatmenetét két évre rá Carnap a *The Continuum of Inductive Methods*-ben folytonos esetekre is kiterjesztette. A valószínűségről alkotott elképzeléseit legteljesebb formában az 1971-ben és 1980-ban két részletben posztumus kiadott *A Basic System of Inductive Logic* című művéből ismerhetjük meg.

A korai logisták, Keynes, Jeffreys, de a valószínűséghez hasonló szellemben közelítő *Tractatus* (1922) Wittgensteinje is cambridge-iek. A mai szemmel nézve kissé idegennek tűnő logikai interpretáció megértéséhez mindenekelőtt tudatosítanunk kell azt a platonista filozófiai hátteret, amely az első világháború előtti Cambridge-et jellemezte. A korszak két fő művét, Russell *Principia Mathematicá*-ját és Moore *Principia Ethicá*-ját egyaránt ez a realista elkötelezettség hatotta át, egyik esetben a matematika törvényei, másik esetben a jó önálló, nem természeti léte iránt. Keynes, Jeffreys és Wittgenstein művei is ebben filozófiai klímában születtek – feltételezve, hogy a valószínűségi viszonyok valamilyen ideális nyelv vagy logikai tér a priori törvényei, amelyekkel közvetlen ismeretség (acquaintance) vagy intuíció révén állunk kapcsolatban. Ez az a priori logikai viszony a kijelentések közötti deduktív következtetés kiterjesztése egyfajta részleges következtetésre. Keynes (1921) így fogalmaz:

> „Amennyiben azt feltételezzük, hogy néha közvetlenül megítélhető, hogy egy konklúzió *következik* egy premisszából, úgy nem túlzott kiterjesztése ennek a feltételezésnek azt gondolni, hogy néha az is felismerhető, hogy egy konklúzió *részlegesen következik* egy premisszából, illetve, hogy a valószínűség viszonyában áll a premisszával." (52. o.)

Vagy ugyanez Wittgenstein (1956) szabatosabb nyelvén:

> „Ha 'I_r' az 'r' kijelentés igazságalapjainak száma, 'I_{rs}' pedig az 's' kijelentés azon igazságalapjainak száma, amelyek egyben 'r' igazságalapjai is, úgy az "$I_{rs} : I_r$' viszonyt az 'r' kijelentés által 's' kijelentésnek kölcsönzött *valószínűségi mértéknek* nevezzük. ...Ha p következik q-ból, úgy a 'q' kijelentés 1 valószínűséget kölcsönöz a 'p' kijelentésnek. A logikai zárótétel bizonyossága a valószínűség egyik határesete." (5.15-5.152.)

– ahol is az igazságalap a kijelentés azon „igazságlehetőségei, amelyek igazolják a kijelentést." (5.101.) Tömören szólva: pusztán a logikai tér vizsgálata eligazítást nyújt a kijelentések valószínűségi viszonyai tekintetében. Éppen ez az, ami miatt a valószínűség logikai fogalom.

Keynes Wittgensteinnél tágabban értelmezi a valószínűség fogalmát. Bár „a logikai viszonnyal közvetlen ismeretségben vagyunk" (13. o.), ez a közvetlen ismeretség azonban nem jelenti azt, hogy feltétlenül a numerikus viszonyokkal, vagy akár csak a kijelentések között valamiféle rendezéssel is tisztában lennénk. Valószínűségi „észleléseink” tágabbak a valószínűségi kalkulussal megragadható tartománynál. A numerikus viszonyok meghatározásához nem kerülhetjük el az indifferencia elvének alkalmazását, amelynek ellentmondásaira, amint arra az előző fejezetben (ld. 69. oldal) két tőle származó példával már utaltunk, maga Keynes mutat rá elsők között.

A Bécsi Körre gyakorolt hatása ellenére Wittgenstein valószínűségi koncepciója nem talált kedvező fogadtatásra a Körben. A logikai pozitivisták – Waismann-t leszámítva – mindannyian a valószínűség frekventista értelmezését részesítették előnyben a logikaival szemben. Amikor Carnap a 40-es évektől kezdve von Mises és Feigl hatására foglalkozni kezdett a valószínűség megalapozásával, a logikai és frekventista megközelítések egymással viaskodó értelmezéseknek számítottak. Carnap nem fogadta el ezt a szembeállítást, hanem mindkét koncepciót legitimnek és tudományosan hasznosnak tekintette. A logikai valószínűségre a valószínűség$_1$, a statisztikai valószínűségre a valószínűség$_2$ jelölést vezette be. A valószínűség$_1$ kijelentések részleges implikációt fejeznek ki egy nyelv mondatai között, és ennyiben a priori és analitikus állítások. Ezzel szemben a valószínűség$_2$ kijelentések empirikus tartalommal bírnak: egy statisztikus tényállást fejeznek ki. Az első típusú kijelentések a tudomány metanyelvéhez tartoznak, míg a második típusú kijelentések magához a tudomány nyelvéhez. Ezért lehetséges például olyan logikai valószínűségi kijelentést tenni egy statisztikus valószínűségi kijelentésről, hogy egy bizonyos statisztikus tényállás ilyen és ilyen mértékben konfirmál egy tudományos hipotézist. Mi a továbbiakban csak Carnap valószínűség$_1$ értelmezésére szorítkozunk.

5.1. Konfirmáció és valószínűség

A *Logical Foundations of Probability* elején Carnap a következő pontokban foglalja össze a mű programját:

> „Az itt kifejtett elmélet a következő alapfogalmakkal jellemezhető: (1) minden induktív érvelés, vagyis tág értelemben nem-deduktív és nem-demonstratív érvelés, valószínűségi terminusokban való érvelés; (2) így az induktív logika, az induktív érvelés elveinek elmélete ugyanaz, mint a valószínűségi logika; (3) a valószínűség fogalma, amelyen az induktív logika alapul, egy logikai viszony két kijelentés vagy propozíció között, egy hipotézis (vagy következmény) konfirmációjának foka egy bizonyíték (vagy premissza) alapján; (4) a valószínűség ún. frekvencia-fogalma, ahogyan azt a statisztikus vizsgálatokban használják, saját jogán fontos tudományos fogalom, de az induktív logika alapfogalmaként nem használható; (5) az induktív logika összes elve és tétele analitikus; (6) így az induktív érvelés érvénye nem függ semmilyen szintetikus előfeltevéstől, úgy mint például a világ uniformitásának feltevésétől.” (v. o.)

Az induktív viszony és az alapjául szolgáló valószínűség$_1$ tehát logikai, és nem empirikus fogalom. Hogy egy ilyen fogalom egy empirista számára is elfogadható lehet, azt Carnap a következőképpen indokolja:

> „Számos empirista szerző azon az alapon utasította el a valószínűség$_2$-től különböző logikai valószínűség$_1$ fogalmát, hogy az sértené az empirizmus elvét, és így csak a valószínűség$_2$ fogalma megengedhető az empirizmus és következésképpen a tudomány számára. Az egyik érv emellett a következő. A valószínűség$_1$ fogalma abban az esetben is alkalmazható, amikor a h hipotézis egy egyedi eseményre vonatkozó jóslat, mint például, hogy holnap esni fog, vagy hogy a következő kockadobás hatos lesz. Néhány filozófus úgy véli, hogy a fogalom alkalmazása megsérti a verifikálhatósági (vagy konfirmálhatósági) elvet. Ezt mondhatják például: 'Hogyan verifikálható az az állítás, hogy ⌜a meteorológiai bizonyítékok fényében annak valószínűsége, hogy holnap esni fog, $\frac{1}{5}$⌝? Holnap vagy esni fog, vagy nem, de semmi olyat nem látunk, ami az $\frac{1}{5}$ értéket igazolhatná.' Ez az ellenvetés azonban a valószínűség$_1$ kijelentés félreértésén alapszik. Ez a kijelentés nem a holnapi esőhöz rendel $\frac{1}{5}$ valószínűséget, hanem egy bizonyos logikai viszonyhoz az eső jóslata és az időjárásjelentés között. Mivel a viszony logikai, ezért az állítás, ha igaz, L-igaz [logikai igazság]; így verifikációjához nincs is szükség a holnapi időjárás vagy bármilyen más tényállás megfigyelésére. A helyzet világosabbá tehető a deduktív logikával való összevetéssel. Legyen h az a hipotézis, hogy 'holnap esni fog', és j az a mondat, hogy 'holnap esni fog és a szél is fúj'. Tegyük fel, hogy valaki azt a deduktív következtetést állítja, hogy 'h logikailag következik j-ből'. Nyilvánvalóan senki sem fogja őt apriorizmussal vádolni csak azért mert ezt állította, vagy mert az állítás verifikációja nem igényel faktuális tudást. Annak az állításnak azonban, hogy 'A h hipotézis valószínűség$_1$-e az e evidencia mellett $\frac{1}{5}$' ugyanaz a formális karaktere, mint az előző állításé; így nem is sértheti jobban az empirizmust, mint az előző. Mindkét állítás egy pusztán logikai viszonyt fejez ki két állítás között. A különbség közöttük mindössze ennyi: amíg az első egy teljes logikai következtetést állít, addig a második úgymond egy részleges következtetési viszonyt; így amíg az első a deduktív logikához tartozik, addig a második az induktív logikához." (30-31. o.)

Carnap logikai programjának vázlatos ismertetéséhez tekintsünk egy egyszerű nyelvet, amely megszámlálható a_i individuumkonstanst és monadikus predikátumot tartalmaz. A monadikus predikátumok *családokba* tömörülnek. Egy családon belül a predikátumok minden individuumra kizárják egymást. Ezek az egymást kizáró prediátumokból álló predikátumcsaládok hivatottak reprezentálni az olyan modalitásokat, mint az alak, a szín, a szubsztancia, stb. A nyelv atomi mondatai ilyen alakúak: $P_j^k a_i$, ahol a predikátumok felső indexe a predikátumcsaládra utal. A továbbiakban válasszunk ki egy predikátumcsaládot, és hagyjuk el a predikátumok felső indexét. Nevezzük az individuumok egy véges halmazát *mintának* (pl. $\{a_1, a_3\}$), és nevezzük *mintaleírásnak* atomi mondatok olyan konjunkcióját, amelyben az individuumok egy adott minta elemei, a predikátumok pedig valamennyien egy adott predikátumcsalád predikátumai (pl. $P_4a_1 \,\&\, P_5a_3$). Ezek után Carnap az indukció intuitív sajátosságait szem előtt tartva az S mintaleírásokon definiál egy $c : S \times S \to [0,1]$ kétargumentumú ún. *konfirmációs függvényt*, amely (többek között) az alábbi tulajdonságokkal rendelkezik. (Ahol a függvénynek csak egy argumentuma van, ott a második argumentum a tautológia, azaz üres minta mintaleírása.)

(i) *Szimmetria.* A c függvény invariáns az individuumkonstansok permutációjára nézve, vagyis pl. $c(P_4a_1 \,\&\, P_5a_3) = c(P_4a_3 \,\&\, P_5a_1)$.

(ii) *Instanciális relevancia.* Azonos predikátum terjedelmébe tartozó individuumok relevánssak egymásra nézve, vagyis minden $i \neq i'$-re $c(P_ja_i, P_ja_{i'}) > c(P_ja_i)$.

(iii) λ*-feltétel.* Ha egy S mintaleírás nem vonatkozik egy a individuumra, akkor $c(P_ja, S)$ nem függhet mástól, csak az S alá tartozó minta elemszámától, és az S-ben P_j tulajdonságú individuumok számától, vagyis pl. $c(P_4a_2|P_4a_3 \,\&\, P_5a_1) = c(P_4a_2|P_4a_6 \,\&\, P_7a_8)$.

A szimmetria követelménye azt reprezentálja, hogy a konfirmációs függvény nem érzékeny arra, hogy egy adott predikátum mely individuumokra igaz. A mintaleírásnak ez a permutációszimmetriája a bayesianizmusban majd a felcserélhetőség nevet viseli. Jelentőségéről a 6.3. alfejezetetben szólunk. Az instanciális relevancia követelménye az indukció legfontosabb sajátosságát, a tapasztalatból való tanulást hivatott modellezni: ha egy tulajdonságot egy individuum már instanciált, akkor ez konfirmálni fogja azt az állítást, hogy a tulajdonságot egy másik individuum is instanciálja. Végül a λ-feltétel azt az induktív következtetési heurisztikát fejezi ki, hogy egy új individuum adott tulajdonságára a tulajdonság eddigi relatív gyakorisága alapján következtetünk.

Mindezekből a feltételekből (és még néhány természetesnek tartott követelményből, pl. $c(S,S) = 1$, $c(S) > 0$) Carnap levezeti az ún. $\lambda\gamma_j$-tételt, amely a következőket állítja:

4. Tétel. ($\lambda\gamma_j$-tétel.) Legyen $c : S \times S \to [0,1]$ olyan kétargumentumú függvény a fent definiált egyszerű nyelv mintaleírásain, amely kielégíti az (i)-(iii) követelményeket. Ekkor létezik egy $\lambda \geqslant 0$ valamint minden P_j predikátumhoz egy $\gamma_j \in (0,1)$ valós szám úgy, hogy

$$c(P_ja, S) \;=\; \frac{n_i + \lambda\gamma_j}{n + \lambda} \tag{5.1}$$

ahol a egy tetszőleges individuumkonstans, amelyre nem vonatkozik S, n_j az S-ben levő azon individuumok száma, amelyről S azt mondja, hogy P_j; n pedig az S-beli összes individuum száma.

Az, hogy (5.1) kielégíti az (i)-(iii) követelményeket, közvetlenül látható. Az ellenkező irányt bizonyítani, vagyis, hogy minden (i)-(iii)-t kielégítő c függvény (5.1) alakú, rövid meggondolást igényel, amelyet az olvasóra hagyunk.

Mi a (5.1) képletben szereplő λ és γ_j jelentése Carnap szerint? Ha a minta az üreshalmaz, vagyis S analitikus mondat, akkor n_j és n is 0, így $c(P_ja) = \gamma_j$, vagyis γ_j annak a priori valószínűsége, hogy az a individuum P_j tulajdonságú. λ meghatározásához alakítsuk át (5.1) jobb oldalát az alábbi módon:

$$\frac{n_i + \lambda\gamma_j}{n+\lambda} = \frac{n}{n+\lambda}\left(\frac{n_j}{n}\right) + \frac{\lambda}{n+\lambda}\gamma_j.$$

$c(P_j a, S)$ tehát két kifejezés konvex kombinációja. Az első kifejezés $\frac{n_j}{n}$, ami egy empirikus tényező, az S mintában a P_j tulajdonság relatív gyakoriságát jelenti. A második kifejezés pedig γ_j, amely, mint láttuk, a P_j tulajdonság a priori valószínűsége. A λ tényező értéke azt fejezi ki tehát, hogy az empirikus és az a priori faktorok milyen arányban keverednek, azaz az induktív tanulás milyen „gyors": milyen mértékben támaszkodik a tapasztalatra, és milyen mértékben ragaszkodik az a priori tényezőhöz. Ha λ kicsi (≈ 0), akkor az empirikus tényező dominál, vagyis a tanulás gyors, ha λ nagy ($\approx \infty$), akkor az a priori tényező, vagyis a tanulás lassú. λ tehát a „logikai" tényező súlyát adja meg.

> „Úgy tűnik, hogy egy megfigyelő szabadon választhat egy megengedhető λ értéket, és ezzel egyszersmind egy induktív módszert is. Ha azt látjuk, hogy az X személy nagyobb λ értéket választ az Y személynél, akkor megállapíthatjuk, hogy X óvatosabb Y-nál, vagyis X megfigyelési adatoknak egy nagyobb osztályát fogja megvárni Y-nál, amíg hajlandó lesz eltérni a relatív gyakoriságokra vonatkozó becsléseiben az a priori értékektől." (Carnap, 1950, 75. o.)

A γ_j értékek megválasztása a megfelelő modalitás a priori vizsgálatával történik. Pl. a színteret felosztjuk a szivárvány színeinek vagy az elektromágneses spektrum tartományainak megfelelően. Ha jobb ötletünk nincs, a predikátumcsalád minden elemének egyenlő a priori valószínűséget tulajdonítunk az indifferencia elvének megfelelően. Laplace rákövetkezési szabályát például úgy kapjuk meg, ha minden családba csak egy predikátum tartozik, és γ_j értékét az indifferencia elvével összhangban $\frac{1}{2}$-nek, λ értékét pedig 2-nek választjuk. Mivel a rákövetkezési szabályánál az eddigi individuumok mind instanciálták P_j-t, ezért $n_j = n$, és így

$$\frac{n_i + \lambda\gamma_j}{n+\lambda} = \frac{n+1}{n+2}.$$

Végül lássunk egy egyszerű példát arra, hogyan is működik a fent vázolt indukció (Salmon et al. 1992). Álljon a nyelvünk három a, b és c individuumkonstansból, valamint egy P predikátumból. A nyelv összesen nyolc mintaleírást tartalmaz. Az (i)-(iii) követelményeket kielégítő konfirmációfüggvények serege két paraméterrel, λ-val és a γ-val paraméterezhető, ahol $\lambda \in [0, \infty)$ és $\gamma \in (0, 1)$. Válasszuk a paramétereket a fenti laplace-i módon: $\lambda = 2$ és a $\gamma = \frac{1}{2}$. Könnyű látni, hogy ezek a paraméterek a nyolc mintaleíráson az alábbi a priori súlyokat fogják meghatározni:

$$\begin{aligned} c(Pa \,\&\, Pb \,\&\, Pc) &= \frac{1}{4}, \\ c({\sim}Pa \,\&\, Pb \,\&\, Pc) = c(Pa \,\&\, {\sim}Pb \,\&\, Pc) = c(Pa \,\&\, Pb \,\&\, {\sim}Pc) &= \frac{1}{12}, \\ c(Pa \,\&\, {\sim}Pb \,\&\, {\sim}Pc) = c({\sim}Pa \,\&\, Pb \,\&\, {\sim}Pc) = c({\sim}Pa \,\&\, {\sim}Pb \,\&\, Pc) &= \frac{1}{12}, \\ c({\sim}Pa \,\&\, {\sim}Pb \,\&\, {\sim}Pc) &= \frac{1}{4}. \end{aligned}$$

Ezekkel a valószínűségi súlyokkal Pc-re az alábbi konfirmációs értékek adódnak:

$$\begin{aligned} c(Pc) &= \frac{1}{2}, \\ c(Pc, Pa) &= \frac{2}{3}, \\ c(Pc, Pa \,\&\, Pb) &= \frac{3}{4}, \end{aligned}$$

amelyek épp a laplace-i szukcesszió számai. Annak a valószínűsége tehát, hogy a c individuum P tulajdonságú, miután sorra meggyőződünk arról, hogy a is, valamint b is P tulajdonságú, egyre növekszik. Carnap interpretációjában ez fejezi ki a tapasztalatból való induktív tanulás tényét.

A λ és a γ paramétert persze választhatnánk másképpen is, pl. így: $\lambda \approx \infty$ és a $\gamma = \frac{1}{2}$. Ez a paraméterezés továbbra is figyelembe veszi az indifferencia elvét, ugyanakkor az a priori valószínűségekhez való megátalkodott ragaszkodást modellezi. Ekkor a mintaleírásokon uniform valószínűségi súlyeloszlás adódik, azaz minden mintaleírás valószínűségi súlya $\frac{1}{8}$.[2] Ekkor Pc konfirmációira azt kapjuk, hogy

$$c(Pc) = c(Pc, Pa) = c(Pc, Pa \,\&\, Pb) = \frac{1}{2},$$

vagyis a tapasztalatból nem lehet tanulni: akármennyi individuum volt eddig P tulajdonságú, ezek nem konfirmálják az állítást, hogy a következő individuum P tulajdonságú lesz.

A $\lambda = 2$, $\gamma = \frac{1}{2}$ illetve a $\lambda \approx \infty$, $\gamma = \frac{1}{2}$ esetre a fizikában mint Bose–Einstein-statisztika illetve mint Maxwell–Boltzmann-statisztika szokás hivatkozni. Mindkettő egyszerű kombinatorikai feladattal példázható: hányféleképpen dobhatok három pénzérmével? Ha az érméket nem különböztetjük meg, akkor a válasz 4, ha megkülönböztetjük, akkor 8. Ha az eseteket egyenlően valószínűnek vesszük, megkapjuk a Bose–Einstein- illetve a Maxwell–Boltzmann-eloszlást.

[2] Ez egyébiránt Wittgenstein súlyfüggvénye a *Tractatusban*: „Két elemi kijelentés $\frac{1}{2}$ valószínűséget kölcsönöz egymásnak." (5.152)

A konfirmációs skála másik végén a $\lambda = 0$, $\gamma = \frac{1}{2}$ eset áll. Ekkor az a priori valószínűségek rögtön az első bizonyítékkal törlődnek, vagyis

$$\begin{aligned} c(Pc) &= \frac{1}{2}, \\ c(Pc, Pa) &= 1, \\ c(Pc, Pa \,\&\, Pb) &= 1. \end{aligned}$$

Más szóval az új eset konfirmációját pusztán az eddigi relatív gyakoriságok határozzák meg a kezdeti a priori faktortól teljesen függetlenül. A 0 és ∞ értékek közötti λ-értékek ennek a két szélsőséges esetnek, a tisztán empirikus és tisztán a priori esetnek a keverékei.

A nyolc mintaleíráson az a priori súlyokat tetszőleges módon kioszthatjuk. Tekintsük pl. a következő elosztást:

$$\begin{aligned} c(Pa \,\&\, Pb \,\&\, Pc) &= \frac{1}{20}, \\ c({\sim}Pa \,\&\, Pb \,\&\, Pc) = c(Pa \,\&\, {\sim}Pb \,\&\, Pc) = c(Pa \,\&\, Pb \,\&\, {\sim}Pc) &= \frac{3}{20}, \\ c(Pa \,\&\, {\sim}Pb \,\&\, {\sim}Pc) = c({\sim}Pa \,\&\, Pb \,\&\, {\sim}Pc) = c({\sim}Pa \,\&\, {\sim}Pb \,\&\, Pc) &= \frac{3}{20}, \\ c({\sim}Pa \,\&\, {\sim}Pb \,\&\, {\sim}Pc) &= \frac{1}{20}. \end{aligned}$$

A súlyokból visszaszámolt paraméterek a következők: $\lambda = -6$ és $\gamma = \frac{1}{2}$, vagyis λ negatív, és így nem esik a fenti tétel hatálya alá. Az eset egy ún. antiindukciós sémát, ahol a konfirmációs értékek az alábbiak:

$$\begin{aligned} c(Pc) &= \frac{1}{2}, \\ c(Pc, Pa) &= \frac{2}{5}, \\ c(Pc, Pa \,\&\, Pb) &= \frac{1}{4}, \end{aligned}$$

vagyis minél több P tulajdonságú individuumunk van, annál *valószínűtlenebb*, hogy a következő is az lesz.

5.2. A konfirmáció mint interpretáció

Amint a példából látható a mintaleírásokon értelmezett $c(\cdot\,,\,\cdot)$ konfirmációs függvény *matematikailag* egy $p(\cdot\,|\,\cdot)$ feltételes valószínűséget határoz meg; c második argumentumába tautológiát írva pedig egy p valószínűségi mértéket. A kérdés már most az, hogy ez a megfelelés mennyiben tekinthető a valószínűség interpretációjának.

Kezdjük néhány technikai észrevétellel. Először is vegyük észre, hogy a fenti nyelv egy primitív nyelv. Bonyolultabb esetekben egyáltalán nem garantálható – amint azt Carnap hasztalan erőfeszítései a kontinuum esetben jól mutatják –, hogy a konfirmációs függvény és így a valószínűség fogalma értelmesen általánosítható. És akkor még nem beszéltünk az olyan szubtilitásokról, hogy a konfirmációs függvény révén nyert valószínűségi mérték vajon σ-additív-e?

De tegyük félre a formális megfontolásokat, és összpontosítsunk a logikai interpretációnak pusztán arra a központi gondolatára, hogy a valószínűség a konfirmáció súlya. Helyes interpretációja-e ez a valószínűségnek? A válaszhoz természetesen azt kell tisztáznunk, hogy mit értünk konfirmáció alatt. A fenti egyszerű példánál maradva milyen mértékben konfirmálja például az az állítás, hogy egy pénzérmével eddig kétszer fejet kaptunk, azt az állítást, hogy harmadszorra is fejet kapunk? Semmilyen mértékben: harmadszorra éppen annyi az esélyünk a fejre, mint ha előtte nem dobtunk volna kétszer fejet. – Honnan tudjuk ezt? – Tapasztalatból. – Mit értünk itt tapasztalaton? – Mondjuk, hogy egy kellően hosszú sorozatban azon fej dobások relatív gyakorisága, amelyek előtt két fej dobás áll közelítőleg ugyanannyi, mint maguknak a fej dobásoknak a relatív gyakorisága a sorozatban. A tapasztalat alapján választjuk ki tehát a fejdobásokra vonatkozó lehetséges konfirmációfüggvények közül éppen a $\lambda = 2$, $\gamma = \frac{1}{2}$ paraméterűt. – Rendben, tegyük fel, hogy létezik valamilyen eljárás, amelynek segítségével egy konfirmációs függvény relatív gyakoriságokra vezethető vissza. Ekkor azonban a következő problémába ütközünk. Ha a valószínűség fogalmát a logikai interpretáció jegyében a konfirmációra építve szeretnénk értelmezni, a konfirmációt pedig a relatív gyakoriságra vezetjük vissza, akkor *a konfirmáció közbülső fogalmát akár ki is hagyhatnánk*, és a valószínűséget közvetlenül relatív gyakoriságként is interpretálhatnánk. Természetesen ugyanez a helyzet a konfirmáció bármely más fizikai értelmezése esetén is: ha a konfimáció fizikailag értelmezhető, akkor eliminálható közvetítőfogalom csupán.

De miért kellene a konfirmációt fizikai fogalmakban értelmezni? És valóban Carnap (1963) a kérdésben igen határozottan azon az állásponton van, hogy az konfirmáció törvényei a priori törvények:

> „Nem osztom azt az igen széles körben elterjedt nézetet, hogy az induktív módszer racionalitása faktuális tudáson múlik, mondjuk, sikerén a múltban. Úgy gondolom, hogy a racionalitás kérdését pusztán a priori meggondolásokkal kell megválaszolni.” (981. o.)

Vagyis a konfirmáció fenti leírása, nem a tapasztalatból való tanulás *tényleges* folyamatának valamilyen empirikus modellje, hanem az összes lehetséges induktív következtetés általános formája, az indukció úgymond általános elmélete. Az olyan szavak pedig, mint „tanulni a tapasztalatból”, „ilyen és ilyen arányban súlyozni az empirikus és az a priori faktort” stb. pusztán metaforikus kifejezések, egy interpretálatlan formalizmus minden empirikus jelentést nélkülöző kifejezései.

Ha azonban konfirmáció alatt ezt értjük, akkor a konfirmációt mint a valószínűség interpretációját sem kell komolyan vennünk. Amint azt az interpretáció általános követelményeit taglaló 3. fejezetben említettük, interpretáció alatt minden esetben *fizikai* interpretációt kell érteni; vagyis nem elégedhetünk meg pusztán azzal, ha a valószínűségnek mint normált mértéknek megadjuk egy másik matematikai modelljét. A logikai interpretáció esetében azonban pont ez történik. A valószínűséget úgy vezetjük vissza a konfirmáció fogalmára, hogy ez utóbbinak semmilyen empirikus jelentést nem adunk. A logikai interpretáció tehát nem tekinthető a valószínűség interpretációjának.

6. fejezet

A szubjektív valószínűség

A valószínűség szubjektív interpretációja a logikai interpretáció kritikájából nőtt ki Frank Ramsey és Bruno de Finetti munkássága nyomán. Ramsey az 1926-ban megírt, de csak halála után 1931-ben megjelent *Truth and probability* című nagy horderejű munkájában fejtette ki szubjektivista nézeteit; de Finetti pedig a 30-as évek elejétől jelentette meg sorra írásait, amelyek közül a legismertebb az 1937-es *La prévision: ses lois logiques, ses sources subjectives.*[1]

Ramsey értelmezése annak a Keynes-től származó elképzelésnek a tagadásával kezdődik, amely szerint a kijelentések közötti logikai-valószínűségi viszony minden racionális individuum számára valamilyen intuitív értelemben egyformán érzékelhető. Ez az elképzelés nyilvánvalóan tarthatatlan, legalábbis Ramsey a maga részéről vitatja, hogy rendelkezne vele:

> „Én nem érzékelem, és ha valaki meg akar győzni, hogy létezik, annak érvekre lesz szüksége; sőt mi több, az a gyanúm, hogy mások sem érzékelik, hiszen olyan csekély egyetértés van közöttük abban, hogy mi is két adott kijelentés viszonya." (161. o.)

Ramsey fenntartja Keynes álláspontját abban a tekintetben, hogy a kijelentések között létezik valamiféle valószínűségi viszony, de megengedi, hogy ez a viszony egyénről-egyénre változhasson. Azonban hogyan lehet tudomást szerezni a logikai viszony egyéni mértékéről, vagy ahogy Ramsey fogalmaz, a kijelentésbe vetett hit mértékéről? Ramsey számára nyilvánvaló, hogy a hit mértéke introspektíve nem mérhető,[2] így aztán arra a konklúzióra jut, hogy „a hit mértéke az a kauzális tulajdonság, amelyet úgy fejezhetünk ki, mint annak mértékét, amennyire készek vagyunk cselekedni általa". Az a hitem, hogy a csapvíz iható, abban jut kifejezésre, hogy habozás nélkül hajlandó vagyok inni belőle. A hitet tehát azoknak a

[1] Ramsey nem tudott de Finettiről, de Finetti pedig 1937-ig nem ismerte Ramsey munkáját.

[2] Eljátszik azonban a gondolattal, hogy talán egy „pszichogalvanométerrel" mérni lehet a hazugságvizsgálók mintájára (amelyet 1924-ben alkalmaznak először).

cselekedeteknek a vizsgálata révén tanulmányozhatjuk, amelyekben az manifesztálódik. Ez a *par excellence* cselekedet Ramsey számára a fogadás: „Egy személy hite mérésének legősibb módja az, hogy fogadást ajánlunk neki, és megnézzük, hogy melyek azok a fogadási arányok, amelyeket elfogad. Ezt a módszert én alapvetően helyesnek tartom." Ramsey számára minden cselekedet fogadás ("Mindahányszor kimegyünk az állomásra, fogadunk, hogy a vonat valóban jár"), így a fogadás a cselekedetekben manifesztálódó hit mérésének alkalmas eszközévé válik.

Ramsey-vel ellentétben de Finettit a logikai interpretációnak elsősorban nem az a feltételezése zavarta, hogy a kijelentések között létezne egy minden racionális személy számára észlelhető részleges implikáció. Számára, aki a laplace-i tradícióban álló francia szerzők (Bertrand, Borel, Lévy, Poincaré) munkáin nevelkedett, a klasszikus és logikai interpretáció elfogadhatatlan pontja az indifferencia elve volt. Az indifferencia elvének kritikájától ő is hamar eljut a fogadás fogalmához. A *locus classicus* a fogadási szituáció karakterizálására de Finetti 1937-es írásában olvasható:

> "Tegyük fel, hogy egy individuumot arra kötelezünk, hogy becsülje meg azt a p arányt, amelyre hajlandó volna felcserélni egy tetszőleges S összeg (pozitív vagy negatív) birtoklását, feltéve hogy egy adott E esemény bekövetkezik, a pE összeg birtoklásáért; azt mondjuk, hogy definíció szerint ez a p szám annak a valószínűségnek a mértéke, amelyet az individuum az E eseménynek tulajdonít, vagy, még egyszerűbben, p E-nek a valószínűsége."

A fogadást a valószínűséggel a Ramsey–de Finetti-tétel kapcsolja össze, amely azt mondja ki, hogy kijelentések egy halmazára történő fogadások kvóciensei akkor és csak akkor nem eredményeznek a fogadó számára biztos veszteséget, ha azok matematikai értelemben valószínűségek. Azokat a filozófiai érveket, amelyek első lépésben a hit mértékét a fogadáshoz kötik, majd második lépésben a fogadást a Ramsey–de Finetti-tétellel a valószínűséghez, összefoglalóan *Dutch book*-argumentumoknak nevezzük. Az első alfejezetben a *Dutch book*-argumentumokat mutatjuk be, a másodikban pedig ezeknek egy 1954-ben Savage által továbbfejlesztett változatát, amely a valószínűséget döntéselméleti keretben a várható hasznosság maximalizálásából mint elvből származtatja. Mivel a valószínűség szubjektív értelmezésének kezdeti törekvései mára teljes mértékben beleolvadtak a bayesianizmusba, ezért a fejezetet a bayesianizmus ismertetésével folytatjuk, végül pedig rátérünk a bayesianizmus empirikus alapjainak kérdésére, és felsorolunk néhány olyan heurisztikát, amely az empirikus pszichológia szerint valószínűségi becsléseinket irányítja.

6.1. *Dutch book*-argumentumok

A *Dutch book* kifejezés eredetéről csak találgatások vannak. Van, aki a holland hajózási társaságok bonyolult kereszt- és viszontbiztosítási rendszereire, mások a holland kereskedők tradicionálisan jó üzleti érzékére vezetik vissza, ismét mások, éppen fordítva, az angol nyelvben gyakori, a hollandokra vonatkozó negatív kifejezésnek ("Dutch uncle", "going Dutch") tartják, hogy a *book* miért *Dutch*. A mai értelemben a *book* kifejezést először Keynes (1921), majd pedig Ramsey (1926) használja híres cikkében, az összetett *Dutch book*ot pedig Lehman (1955).

A valószínűség szubjektív interpretációjának tézismondata szerint „a valószínűség a hit mértéke”. De milyen értelemben lehetséges *részlegesen* hinni valamiben? Ha csak egy kicsit hiszek abban, hogy elérem a vonatot, akkor végső soron nem abban hiszek-e, hogy kicsi az esélyem, hogy elérem? Vagyis léteznek-e *nem teljes* hitek? A nem teljes hitek léte mellett bizonyíték gyanánt a kételkedést szokás felhozni, vagyis egy olyan elmeállapotot, amely valamiképpen a hit és a nem hit között fekszik. Bármi legyen is az indok, fogadjuk el kiindulásként, hogy léteznek parciális hitek.

A szubjektív interpretáció a valószínűséget tehát hitekhez vagy pontosabban hitek propozicionális tartalmához, azaz kijelentésekhez rendeli. Egy (nulladrendű) nyelv kijelentései azonban önmagukban nem alkotnak Boole-algebrát, csak a logikai ekvivalenciával való faktorizálás után – így hát nem világos, hogy milyen értelemben lehetséges rajtuk valószínűségi mértéket definiálni. Az alábbiakban megmutatjuk, hogy a *Dutch book*-argumentum hogyan kényszeríti ki a nyelv faktorizációját, majd pedig azt, hogy az így nyert Boole-algebrán a „hitek mértéke” miért valószínűségi mérték.

Az alapító atyák elgondolása szerint a valószínűségi mérték egy fogadási helyzet révén definiálható. A fogadási helyzetben egy X játékos egy A kijelentés igazságára fogad egy másik, Y játékossal. Az X játékos megad egy $q(A)$ számot, a *fogadási kvócienst*, amely azt reprezentálja, hogy X milyen mértékben hisz A igazságában. Ezután Y választ egy $S(A)$ összeget (mondjuk forintban mérve), a *fogadási tétet*, amely negatív is lehet. Az X játékos nevezési díjként befizet $q(A)S(A)$ összeget Y-nak. Ha A igaznak bizonyul, akkor X elnyeri az S tétet Y-tól; ha A hamisnak bizonyul, akkor X nem kap semmit. X össznyereménye tehát $(1-q(A))S(A)$, ha A igaz, vesztesége pedig $q(A)S(A)$, ha A hamis. Az $(1-q(A)):q(A)$ arány megadja, hogy milyen mértékben hajlandó az X játékos fogadni A igazságára $\sim A$ ellenében.

Annak a szerepe, hogy $S(A)$ negatív is lehet, a következő. Ha a játékvezető csak pozitív tétet választhatna, akkor X-nek érdekében állna minél kisebb kvócienst választani, hiszen így olcsóban nevezhetne a fogadásban, A igazságának „esélye” pedig úgysem változna. S negatív választása éppen ezt a lehetőséget zárja ki. Amennyiben S negatív, a szerepek megcserélődnek a játékban: Y játékos fizet $q(A)|S(A)|$ összeget X-nek – ha A igaz, Y megnyeri az $|S|$ tétet, ha A nem igaz, Y nem kap semmit. X-nek tehát nem érdemes sem lefelé, sem felfelé torzítania fogadási hajlandóságát, különben Y a tét előjelének megválasztásával nyerhet ellene.

Általános esetben az X játékos a nyelv minden A kijelentéshez megad egy $q(A)$ fogadási kvócienst, vagyis megad egy $q : Form \to \mathbb{R}$ kvóciensfüggvényt, amely X-nek a fogadási hajlandóságát reprezentálja a nyelv különböző mondataival szemben. A q kvóciensfüggvény ismeretében Y megad egy $S : Form \to \mathbb{R}$ tétfüggvényt, amely minden A mondathoz egy $S(A)$ tétet rendel. Az A mondatra történő fogadásnál az X játékos nyereménye az I

interpretációban:

$$G^I(A) \;=\; \big(I(A) - q(A)\big)S(A).$$

Fogadni azonban egyszerre több kijelentésre is lehet. Ha a fogadás a kijelentések valamely $F \subseteq Form$ halmazára történik, akkor X össznyereménye egy adott I interpretációban

$$G^I(F) \;=\; \sum_{A\in F} G^I(A).$$

Hogy a fogadás milyen kijelentésekre történjék, azt mindig Y választja. Az össznyeremény segítségével definiálhatjuk a *Dutch book* és a konzisztencia fogalmát.

13. Definíció. X-nek egy q kvóciensfüggvényével szemben Y-nak egy S tétfüggvényét *Dutch book*nak nevezzük formulák egy adott $F \subseteq Form$ halmazán, ha $G^I(F) < 0$ minden interpretációban; vagyis, bármi történjék, X veszít az F-beli mondatokra való együttes fogadáson. Ha q ellen nem választható *Dutch book* F-en, akkor q-t *F-konzisztensnek* mondjuk. q-t *konzisztensnek* mondjuk, ha F-konzisztens $Form$ minden F részhalmazán.

A fenti fogalmak segítségével a *Dutch book*-argumentum a következőképpen fogalmazható meg:

1. premissza: Annak a hitnek a mértéke, amelyet egy individuum egy kijelentés igazságába (valójában, mint majd látni fogjuk, egy esemény bekövetkezésébe) vet, jól operacionalizálható azzal a kvócienssel, amellyel a kijelentés igazsága (esemény bekövetkezése) mellett hajlandó fogadni egy fogadási szituációban.

2. premissza: Egy racionális individuum egy nyelv kijelentéseire (egy eseményalgeba elemeire) nem fogad úgy, hogy bármi történjék, garantáltan veszítsen.

Konklúzió: Egy racionális individuum hitének mértéke valószínűség.

Egyelőre fogadjuk el a fenti két premisszát, vagyis tegyük fel, hogy az X játékosunk fogadási kvóciensei jól reprezentálják hitének mértékét, és hogy X racionális, tehát nem fogad úgy, hogy garantáltan veszítsen. Feladatunk azt megmutatni, hogy X fogadási kvóciensei valószínűségek. Első lépésben megmutatjuk, hogy X ellen mindig adható *Dutch book*, ha q különböző értékeket rendel logikailag ekvivalens formulákhoz, vagyis ha $q(A) \neq q(A')$ két olyan A és A' kijelentésre, ahol $A \Leftrightarrow A'$. Tegyük fel ugyanis, hogy $A \Leftrightarrow A'$, de $q(A) \neq q(A')$. Fogadjon Y ekkor X-szel az A és A' kijelentésekre, azaz legyen $F = \{A, A'\}$, és válassza a tétfüggvényt úgy, hogy

$$\begin{aligned} S(A) &= \operatorname{sgn}\big(q(A) - q(A')\big) \\ S(A') &= \operatorname{sgn}\big(q(A') - q(A)\big), \end{aligned}$$

ahol sgn a szignum-függvény, és legyen S értéke 0 minden más $B \notin F$ kijelentésre. Ezzel a választással X össznyereménye $G^I(F) = -|q(A) - q(A')| < 0$ bármely interpretációban, vagyis bármi történik, X veszít.

Ez viszont azt jelenti, hogy konzisztens kvóciensfüggvényként csak azok a függvények jöhetnek szóba, amelyek respektálják a nyelv logikai ekvivalenciáját, vagyis az ekvivalenciaosztályokhoz azonos értéket rendelnek. Más szóval, a konzisztens kvóciensfüggvények a nyelv ekvivalenciaosztályain vannak értelmezve, ami egy Boole-algebra. Ezzel egy lépéssel közelebb léptünk a kvóciensfüggvények valószínűségként való értelmezéséhez. Ennek fényében az alábbiakban $Form$-ról a nyelv Σ Lindenbaum–Tarski-algebrájára módosítjuk q-nak és S-nek az értelmezési tartományát. Σ elemeire (vagyis az $a \equiv [A]$ ekvivalenciaosztályokra) egyszerűen mint eseményekre fogunk hivatkozni. X és Y tehát események bekövetkezésére fogad; a kvóciensfüggvény egy $q : \Sigma \to \mathbb{R}$, a tétfüggvény egy $S : \Sigma \to \mathbb{R}$ hozzárendelés; F pedig Σ részhalmaza. Az interpretációkból $I : \Sigma \to \{0, 1\}$ háló-homomorfizmusok lesznek (amelyeket továbbra is interpretációknak nevezünk).

Az ekvivalens kijelentésekre adott eltérő fogadási kvóciensek azonban nem merítik ki az inkonzisztens kvóciensfüggvények körét. Nézzük erre egy példát (immár az eseményeket használó terminológiában).

1. Példa. Tegyük fel, hogy az X játékos $0,6$ mértékben hisz abban, hogy egy adott pénzérmével való dobás fejet fog eredményezni, és *ugyanakkor* $0,6$ mértékben abban is, hogy a dobás írás lesz. X fogadási kvóciensei inkonzisztensek, ellenük *Dutch book* adható.

Legyen tehát $F = \{\{f\}, \{i\}\}$ az pénzdobások $\Sigma = \{\emptyset, \{f\}, \{i\}, \Omega\}$ eseményterének részhalmaza (partíciója), és legyen q a következő:

$$q(\{f\}) = 0,6; \;\; q(\{i\}) = 0,6$$

Válassza Y a kvóciensfüggvénnyel szemben az alábbi tétfüggvényt:

$$S(\{f\}) = 1; \;\; S(\{i\}) = 1$$

Ekkor az X játékos $0,2$ összeget fog veszíteni akár a fej, akár az írás következik be, mivel

$$G^I(F) \;\;=\;\; \sum_{a \in F} \big(I(a) - q(a)\big) S(a) = 1 - 0,6 - 0,6 = -0,2$$

Ugyanakkor ha X a kvócienseket

$$q(\{f\}) = 0,4; \;\; q(\{i\}) = 0,4$$

értékűnek választaná, akkor könnyen látható, hogy Y a

$$S(\{f\}) = -1; \;\; S(\{i\}) = -1$$

tétfüggvénnyel szintén $0,2$ összeget nyerne tőle fogadásonként.

A *Dutch book*-argumentum szempontjából központi jelentőségű a racionális ágens fogadási kvócienseit és a valószínűségeket egymásnak megfeleltető *Ramsey–de Finetti-tétel.*

5. Tétel. Egy $q : \Sigma \to \mathbb{R}$ kvóciensfüggvény akkor és csak akkor konzisztens, vagyis nem adható ellene *Dutch book*, ha valószínűségi mérték.

Bizonyítás. Lásd a C. Függelékben.

A szubjektivisták tehát a fogadás fogalmára építve képesek megmutatni, hogy a hit mértékét jellemző kvóciensfüggvény értelmezési tartománya miért lesz Boole-algebra, és hogy a kvóciensfüggvény miért lesz valószínűségi mérték. A szubjektivisták ambíciózusabb irányzata, köztük de Finetti, azonban a fogadás segítségével nemcsak a valószínűségi mérték fogalmát kívánja levezetni, hanem magát a Bayes-tételt is. Korábban említettük, hogy miután a klasszikus valószínűségi elméletben *definiáltuk* a feltéles valószínűséget, a Bayes-tétel az elmélet egyszerű tétele lesz. Így a *Dutch book*-argumentum indirekt módon a Bayes-tételt is igazolja. A feltételes valószínűség elméleti definícióját azonban a szubjektivisták nem érzik eléggé motiváltnak, ezért a feltételes valószínűséget az alábbi fogadási szituációval értelmezik. Legyen a és b két esemény, és értelmezzük a „fogadás a-ra, feltéve, hogy b" jelentését az alábbi módon. Az X játékos megad egy $q(a|b)$ fogadási kvócienst. Ehhez a kvócienshez Y választ egy $S(a|b)$ tétet, X pedig befizeti a kívánt $q(a|b)S(a|b)$ nevezési díjat. Ha b nem következik be, akkor a fogadást törlik, és X visszakapja a pénzét. Ha azonban b bekövetkezik, akkor minden úgy zajlik, ahogy egy a eseményre történő fogadás esetében: ha a bekövetkezik, X nyer $S(a|b)$ összeget, ha a nem következik be, akkor X nem nyer semmit. Ha S negatív, akkor a szerepek megcserélődnek.

Általános esetben X megad egy $q(\cdot|\cdot) : \Sigma \times \Sigma \to \mathbb{R}$ feltételes kvóciensfüggvényt, Y pedig megad egy $S(\cdot|\cdot) : \Sigma \times \Sigma \to \mathbb{R}$ feltételes tétfüggvényt. Jelöljük $G^I(a|b)$-vel X-nek az I interpretációbeli nyereményfüggvényét az a-ra történő fogadáskor a b feltétel mellett. Nyilvánvalóan

$$G^I(a|b) \quad = \quad \big(I(a \wedge b) + (I(b^\perp) - 1)\, q(a|b)\big)\, S(a|b).$$

Most tegyük fel, hogy X egyszerre megadja a p és a $q(\cdot|\cdot)$ kvóciensfüggvényeket, vagyis fogad feltételesen is, és feltétel nélkül is. Tegyük fel továbbá, hogy p valószínűségi mérték, vagyis konzisztens. Reprezentáljuk a fogadás tárgyát képező eseményeket, és a feltételes fogadás eseménypárjait egy $F = F_1 \cup F_2$ halmazzal, ahol $F_1 \subseteq \Sigma$ és $F_2 \subseteq \Sigma \times \Sigma$, vagyis legyen F a Σ eseménytér elemeiből és rendezett elempárjaiból álló halmaz. Ekkor X össznyereménye az I interpretációban

$$G^I(F) \quad = \quad \sum_{a \in F_1} G^I(a) + \sum_{(b,c) \in F_2} G^I(b|c). \tag{6.1}$$

Az alábbiakban kiterjesztjük a konzisztencia és a *Dutch book* fogalmát.

14. Definíció. X-nek egy p valószínűségi mértékkel és $q(\cdot|\cdot)$ feltételes kvóciensfüggvényével szemben Y-nak egy S és $S(\cdot|\cdot)$ tétfüggvényét *Dutch book*nak nevezzük F-en, ha $G^I(F) < 0$ minden interpretációban. Ha p és $q(\cdot|\cdot)$ ellen nem választható *Dutch book* F-en, akkor őket *feltételesen F-konzisztensnek* mondjuk. p-t és $q(\cdot|\cdot)$-t *feltételesen konzisztensnek* mondjuk, ha konzisztens $\Sigma \times (\Sigma \times \Sigma)$ minden részhalmazán. Ilyenkor $q(a|b)$-t – amennyiben $p(b) \neq 0$ – a-nak a b-re vonatkozó *feltételes valószínűségének* nevezzük, és $p(a|b)$-val jelöljük.

Nézzünk egy példát (konzisztens, de) *feltételesen* inkonzisztens fogadásra:

2. Példa. Legyen Σ a kockadobások eseménytere, és válassza X p-nek az uniform valószínűségi mértéket Σ-n, vagyis legyen $p(\{i\}) = \frac{1}{6}$ minden $1, \ldots, 6$ eseményre. Mivel p valószínűségi mérték, nem adható ellene *Dutch book*. A p és $q(\cdot|\cdot)$ ellen adott *Dutch book*ok csak $q(\cdot|\cdot)$ inkonzisztenciájára épülhetnek.

Legyen $a \equiv \{6\}$ a hatos dobás, és $b \equiv \{2, 4, 6\}$ a páros dobás eseménye. Nyilvánvalóan $p(a) = \frac{1}{6}$ és $p(b) = \frac{1}{2}$. Válassza már most X a $q(\cdot|\cdot)$ feltételes kvóciensfüggvényt olyannak, hogy $q(a|b)$ nagyobb, mint $\frac{p(a \wedge b)}{p(b)} = \frac{1}{3}$; legyen mondjuk $q(a|b) = \frac{1}{2}$. Ekkor Y az alábbi *Dutch book*ot tudja adni X ellen az $F_1 = \{b^\perp, a \wedge b\}$ és az $F_2 = \{(a, b)\}$ halmaz uniójának választott F halmazon. Legyen a tétfüggvény

$$\begin{aligned} S(b^\perp) &= -\frac{1}{2} \\ S(a \wedge b) &= -1 \\ S(a|b) &= 1 \end{aligned}$$

Ekkor látható, hogy X – a dobás kimenetétől függetlenül – $\frac{1}{12}$ összeget fog veszíteni. Az X játékos nevezési díja $-\frac{1}{2} \times \frac{1}{2} - 1 \times \frac{1}{6} + 1 \times \frac{1}{2} = \frac{1}{12}$; az összes kifizetés pedig, amint az alábbi táblázat mutatja, minden interpretációban 0.

Interpretációk	$b^\perp$	$a \wedge b$	a, ha b	Összesen
I(a)=1, I(b)=1	0	-1	1	0
I(a)=1, I(b)=0	$-\frac{1}{2}$	0	$\frac{1}{2}$	0
I(a)=0, I(b)=1	0	0	0	0
I(a)=0, I(b)=0	$-\frac{1}{2}$	0	$\frac{1}{2}$	0

Ugyanezt az eredményt (6.1)-be helyettesítve közvetlenül is megkapjuk.

Ha már most X a két esemény feltételes kvóciensének $\frac{1}{3}$-nál kisebb értéket választana, mondjuk $q(a|b) = \frac{1}{6}$-ot, akkor a

$$\begin{aligned} S(b^\perp) &= \frac{1}{6} \\ S(a \wedge b) &= 1 \\ S(a|b) &= -1 \end{aligned}$$

tétfüggvényválasztással X ismét csak $\frac{1}{12}$ összeget veszítene minden esetben.

A kérdés tehát, hogy X hogyan válassza meg a $q(\cdot|\cdot)$-t, hogy az feltételesen is konzisztens legyen, azaz feltételes valószínűséget definiáljon. Az alábbi tétel erre a kérdésre ad választ.

6. Tétel. p és $q(\cdot|\cdot)$ akkor és csak akkor feltételesen konzisztens, ha minden $a, b \in \Sigma$ elemre $(p(b) \neq 0)$ fennáll, hogy

$$q(a|b) \quad = \quad \frac{p(a \wedge b)}{p(b)}. \tag{6.2}$$

Bizonyítás. Lásd a C. Függelékben.

Nyilvánvaló, hogy a Bayes-tétel (6.2) közvetlen következménye. A feltételes valószínűség operacionalista definíciójával tehát a Bayes-tétel közvetlenül igazolható *Dutch book*-argumentummal. A szubjektivisták legfőbb törekvése tehát teljesülni látszik: a fogadási szituációkban való racionális viselkedés elvezet a valószínűség és a feltételes valószínűség szubjektív fogalmához.

Itt az idő azonban, hogy megvizsgáljuk a *Dutch book*-argumentum premisszáinak plauzibilitását. Kezdjük a 2. premisszával (88. oldal). Vajon irracionális-e minden esetben az a fogadás, amelyik inkonzisztens, vagyis, amely ellen *Dutch book* adható? Vegyük az alábbi Patrick Mahertől (1993) származó példát:

> „Tegyük fel, hogy éjszaka busszal akarsz hazamenni a városból, azonban csak 60 cent van a zsebedben, míg a buszjegy 1 dollárba kerül. A bukméker, látva szorult helyzetedet, a következő fogadást ajánlja: ha odaadod neki a 60 centedet, feldob egy érmét, és ha az fej, ad neked 1 dollárt, különben elveszted a 60 centet. Ha elfogadod a bukméker ajánlatát, $50 - 50\%$ lesz az esélyed, hogy busszal tudsz hazamenni, míg ha elutasítod, kénytelen leszel gyalog hazasétálni. Mindezt látva, úgy érezheted, hogy az ajánlat elfogadható. Feltehetően az ajánlat ugyanúgy elfogadható lenne, ha fej helyett írásra fogadnátok; nincs okod előnyben részesíteni az egyiket a másikhoz képest.
>
> Már most ha a szubjektív valószínűséget *Dutch book*-argumentummal definiáljuk, a fej szubjektív valószínűsége $0,6$, és az írásé is ugyanannyi, és így a valószínűségeid sértik a valószínűségi kalkulust. A *Dutch book*-argumentum azt állítja tehát, hogy irracionális vagy. Mégis a kínos helyzetre tekintettel preferenciáid teljesen ésszerűnek tűnnek."

A példára a szubjektivisták négyféle módon reagálhatnak. (i) a racionalitást egyszerűen mint *Dutch book*-mentességet *definiálják*, (ii) újabb metaérveket adnak arra, hogy a racionalitás miért is azonosítható a *Dutch book*-mentességgel, (iii) úgy módosítják racionalitás fogalmukat, hogy az a példában vázolt szituációt is magába foglalja, (iv) elfogadják, hogy a fenti második premissza igazsága (akár csak az elsőé) empirikus kérdés. Az (i) álláspont (amely egyébként a kezdeti, naiv korszak álláspontja volt) nyilvánvalóan önfelszámoló, mivel a szubjektivisták kizárják magukat minden érdemi vitából a *Dutch book*-argumentumok helytállósága tekintetében. A (ii) álláspontnak nem ismerem képviselőit, de a Maheréhez hasonló példák miatt nehéz is volna képviselni. Mivel a (iv) álláspont, azaz az empíriának való meghódolás az egész szubjektivista konstrukció végét jelentené (hiszen ennyiben akár azt is kérdezhetjük minden *Dutch book*-argumentum nélkül, hogy hiteink *valóban* valószínűségek-e), ezért a szubjektivisták utolsó menedéke a (iii) álláspont, vagyis a racionalitás fogalmának tágítása. Az alábbiakban a leghíresebb ilyen irányú kísérletet, Savage (1954) döntéselméleti keretbe ágyazott elgondolását ismertetjük.

6.2. A várható hasznosság maximalizálása

A döntéselmélet három alapfogalma a *cselekedet*, az *állapot* és a *következmény*. Egy ágens bizonyos cselekedetek közül választhat, amelyek általa nem kontrollálható állapotokban eltérő következményekhez vezetnek. Az elmélet központi kérdése az optimális cselekedetek meghatározása. A kérdés kezelésének egyik legígéretesebb tárgyalása a várható hasznosság fogalmára épít. Az elgondolás középpontjában az a maxima áll, hogy az ágens cselekvési stratégiáját a következmények várható hasznosságának maximalizálása vezérli.

Az első döntéselméleti problémának Pascal *fogadását* tekintik. Azt az elképzelést, hogy az individuumok a vagyon helyett a hasznosság várható értékenek maximalizálására törekszenek, a *pétervári paradoxon* vetette fel először. A várható hasznosság legkorábbi axiomatikus kezelése Ramseytől (1926) származik, széles érdeklődést kiváltó axiomatikus tárgyalás Neumanntól és Morgensterntől (1944), majd ennek alkalmazása a kockázatviselésre Arrowtól és Pratt-től (1974), a reprezentációs tétel döntéselméleti bizonyítása pedig Savage-től (1954). Bennünket most ez a reprezentációs tétel érdekel, amely szerint ha egy ágens preferenciája kielégít bizonyos racionálisnak mondott követelményeket, akkor ezek a preferenciák várható hasznosságként lesznek reprezentálhatók. Savage munkája erősen technikai (rövid, jó összefoglalót ad (Luce, Raiffa 1957, p. 302-304.)), így csak az alapvető fogalmakat és posztulátumokat ismertetjük, és elhagyjuk a technikai részleteket.

Az alapfogalmak reprezentációja Savage-nél roppant egyszerű: mind az állapotok (feltételek, körülmények), mind a következmények halmazok, a cselekedetek pedig – lévén hogy a világ állapotainak függvényében bizonyos következményekhez vezetnek – az állapotok tere és a következmények tere közötti leképezések. Formálisan tehát az állapotokat egy (Ω, Σ), a következményeket egy (Θ, Ξ) mérhető térrel reprezentáljuk, a cselekedetek pedig az $\Omega \to \Theta$ mérhető függvények egy D halmazával, amelyekre az alábbiakban rovunk majd ki kikötéseket. Az illusztráció kedvéért tekintsük az alábbi Savage-től (1954, 13. o.) származó példát:

> „A feleséged éppen az ötödik tojást üti bele egy tálba, amikor betoppansz, és ajánlkozol, hogy befejezed az omlettet. A hatodik tojás, amelyet valamilyen oknál fogva vagy az omlettbe kell tenned, vagy kidobnod, a tál mellett hever

töretlenül. El kell döntened, hogy mit teszel ezzel a tojással. Talán nem túlzott egyszerűsítés azt mondani, hogy három cselekedet közül választhatsz: vagy beleütöd a tálba, amelyben már ott van az előző öt, vagy egy másik csészébe ütöd, hogy megvizsgáld, jó-e, vagy eldobod mindenféle vizsgálat nélkül. A tojás állapotától függően, mindhárom cselekedet valamilyen következményhez vezet, mondjuk azokhoz, amelyeket az alábbi táblázat mutat.

Cselekedetek	Állapot: a tojás jó	Állapot: a tojás rossz
A tálba töröd	omlett hat tojásból	nincs omlett, öt tojás elpocsékolva
A csészébe töröd	omlett hat tojásból és egy mosogatni való csésze	omlett öt tojásból és egy mosogatni való csésze
Eldobod	omlett öt tojásból és egy jó tojás elpocsékolva	omlett öt tojásból”

A cselekedetek tehát az állapotok és a következmények közötti leképezések. A reprezentációs tétel premisszái a cselekedeteknek erre a D függvényterére rónak ki bizonyos *racionális* feltételeket. Nézzük a legfontosabbakat.

Összefüggési feltétel: A cselekedetek között létezik egy összefüggő gyenge $\preccurlyeq$ preferenciareláció, azaz bármely $e, f \in D$-re $e \preccurlyeq f$ vagy $e \succcurlyeq f$ vagy mindkettő.

Tranzitivitási feltétel: A preferenciareláció tranzitív, azaz minden $e, f, g \in D$-re, ha $e \preccurlyeq f$ és $f \preccurlyeq g$, akkor $e \preccurlyeq g$.

Vagyis, ha jobb döntés inkább elővigyázatosan a csészébe ütni a tojást, mint a tálba, és ez utóbbi jobb, mint kidobni a tojást, akkor a csészébe ütni is jobb, mint kidobni.

A következő feltétel megvilágításához módosítsuk Savage példáját. Tegyük fel, hogy egy negyedik választásod is van: a döntés súlyától megijedve fogod mind a hat tojást, és a szemétbe hajítod. A döntési mátrix ekkor a következőképpen fest:

Cselekedetek	Állapot: a tojás jó	Állapot: a tojás rossz
A tálba töröd	omlett hat tojásból	nincs omlett, öt tojás elpocsékolva
A csészébe töröd	omlett hat tojásból és egy mosogatni való csésze	omlett öt tojásból és egy mosogatni való csésze
Eldobod a hatodikat	omlett öt tojásból és egy jó tojás elpocsékolva	omlett öt tojásból
Mind eldobod	nincs omlett, hat tojás elpocsékolva	nincs omlett, öt tojás elpocsékolva

Ha a tojás rossz, akkor az első és a negyedik cselekedet következményei megegyeznek: mindkét esetben öt tojást pocsékolsz el. A reprezentációs tétel következő premisszája azt mondja ki, hogy a két cselekedet közötti preferenciasorrend nem függhet attól, hogy pontosan mi is ez a következmény. Vagyis, ha az egyes cselekedetet előnyben részesíted a negyedikkel szemben, akkor ez akkor is így lenne, ha az öt elpocsékolt tojást még tetézné mindkét esetben feleséged szidalma is. A feltétel pontos megfogalmazásához vezessük be az alábbi jelölést: $e, e' \in D$-re $e =_a e'$ illetve $e =_{a^\perp} e'$ jelentse azt, hogy e és e' következményei megegyeznek az $a \in \Sigma$ illetve az $a^\perp \in \Sigma$ állapotban, vagyis $e(x) = e'(x)$ minden $x \in a$-ra. Ezzel a jelöléssel a harmadik feltétel így fest:

Függetlenségi feltétel: Minden $e, e', f, f' \in D$ cselekedetre és $a \in \Sigma$ állapotra, ahol

$$\begin{aligned} e =_a e' \quad &\text{és} \quad f =_a f', \\ e =_{a^\perp} f \quad &\text{és} \quad e' =_{a^\perp} f', \end{aligned}$$

ha $e \preccurlyeq f$, akkor $e' \preccurlyeq f'$.

A függetlenségi elv tehát azt mondja ki, hogy ha két cselekedetnek ugyanaz a következménye valamely állapotban, akkor a cselekedetek közötti preferenciák nem függhetnek ettől a következménytől. Erre az elvre a fejezet végén még visszatérünk. A fenti kikötéseket még néhány technikai kikötés kíséri, amelyek együtt alkotják a reprezentációs tétel premisszáit. Ezek után a reprezentációs tétel így hangzik.

7. Tétel. Legyen adva az állapotok (Ω, Σ) és a következmények (Θ, Ξ) mérhető tere, valamint a cselekedetek mint $\Omega \to \Theta$ mérhető függvények D halmaza a fenti gyenge rendezéssel, amely kielégíti a fenti kikötéseket (és még néhányat). Ekkor (Ω, Σ)-n létezik egy p valószínűségi mérték, (Θ, Ξ)-n pedig egy u hasznosságfüggvény, úgy hogy

$$e \preccurlyeq f \quad \text{akkor és csak akkor, ha} \quad Exp(u \circ e) \leqslant Exp(u \circ f)$$

ahol $u \circ f$ az u és f függvények kompozíciója. A p valószínűség egyértelmű, az u hasznossági függvény pedig pozitív affin transzformáció erejéig egyértelmű.

A reprezentációs tétel szerint tehát egy racionális ágens preferenciái egyértelműen reprezentálhatók egy valószínűségi mértékkel és hasznossági függvénnyel, amelyekre a várható hasznosság visszaadja a preferenciarelációt. A valószínűség tehát általános döntési helyzetekben való racionális viselkedésből eredeztethető. A valószínűségnek ez az ún. preferencia-interpretációja a *Dutch book*-argumentumok általánosításának tekinthető, amely a Maher-típusú példák racionalitásának kimutatására törekszik. Ha nagyon szeretnél busszal hazajutni, racionális lesz fogadnod, még akkor is, ha a fogadás nem *fair*.

De vajon előbbre jutottunk-e fenti kérdésünkben, hogy ti. miben is áll a racionalitásként elfogadott feltételek racionalitása. A *Dutch book*-mentesség követelményét most hét posztulátum váltotta fel, amelyek maguk után vonják a szubjektív valószínűség létét. De mire alapozzuk a posztulátumok igazságát? Itt a szubjektivistákat ismét csak az a veszély fenyegeti, hogy kénytelenek lesznek elismerni, hogy a kérdés empirikus. Néhányan nem adják fel a harcot, és a fenti premisszákat újabb racionális elvből vezetik le (pl. szinkron szeparálhatóság, lásd alább). Ezeket a törekvéseket nem követjük, hanem elfogadjuk, hogy mind Savage posztulátumai, mind az eredeti *Dutch book*-argumentumok premisszái empirikus elvek. Mielőtt azonban megvizsgálnánk, hogy empirikusan adekvátak-e, nézzük meg teljes fegyverzetében azt az elméletet, ami a szubjektivista valószínűségfelfogásra épül. Ez a bayesianizmus.

6.3. Bayesianizmus

A bayesianizmus huszadik századi episztemológiai irányzat, amelyben központi szerepet játszik a valószínűség fogalma. Az eredetileg a konfirmáció problémáinak tanulmányozására kifejlesztett módszer azonban mára vezető szerephez jutott mind a döntéselméletben, mind a tanuláselméletben vagy az ún. Bayes-hálók elméletében. Az elnevezés a 18. századi angol teológusra és matematikusra, Thomas Bayesre megy vissza (ld. 15. oldal).

A modern bayesianizmust az a Bécsi kör nyomán létrejött törekvés hozta létre, hogy a deduktív logika apparátusához hasonlóan megalkossák az induktív következtetések formális elméletét. A elméletet alakjához a vezérfonalat a deduktív logikával való analógiák szolgáltatták. Az induktív következtetést a deduktív következtetés mintájára képzelték el, azzal a különbséggel, hogy míg a deduktív esetben a premisszák maguk után vonják a konklúzió igazságát, addig az induktív esetben ez az implikáció csak részleges. A részleges következtetés nyomán a kijelentések igazsága csak részleges igazolást nyer.

A korai szubjektivisták szakítanak azzal az induktív-logikai kontextussal, amelyben a részlegesen igazolt kijelentés fogalma megszületett, de megőrzik a kijelentésekbe vetett hit mértékének a fogalmát. Ismét csak a deduktív logikából merítve a hitek mértékére korlátozásokat rónak ki. Episztemológiai olvasatban a deduktív logika két értelemben korlátozza a racionális hiteket. Egyrészt a logikai konzisztencia követelménye révén, amely szinkron módon korlátozza az egyszerre képviselhető hitek körét, másrészt a logikai következményrelációin keresztül, amely a régi hitekből az újakra való áttérés diakron lépését koordinálja. Ennek a szinkron és diakron racionalitási előírásnak lett az örököse az két elv, amelyeket ma a modern bayesianizmus alapelveiként tartunk számon.

Konzisztencia: A szinkron racionális hitek mértékei valószínűségi mezőt alkotnak.

Kondicionálás: A racionális hitek viselkedését a valószínűségi kondicionálás írja le.

Mindkét elv olyan racionális ágenst tart a szeme előtt, aki egy nyelv kijelentéseihez szubjektív hitének mértékében számértékeket rendel, majd miután teljes vagy részleges bizonyosságot szerez egyik vagy másik kijelentésbe vetett hitét illetően, ezeket a mérőszámokat a bizonyíték alapján változtatja. Kihallható, hogy ez az episztemológiai modell elsősorban a konfirmáció leírására született: hogyan módosítsunk egy eddig csak részlegesen igazolt hipotézist, ha újabb bizonyítékhoz jutunk. A két elv tehát két előírást fogalmaz meg a hitek mértékére vonatkozóan. Az egy időben képviselt hitek mérőszámai elégítsék ki a valószínűség axiómáit; az új bizonyíték szerzésével a mérőszámok pedig frissüljenek a kondicionálás szerint. Az első elvvel az előbbiekben már foglalkoztunk, nézzük most a kondicionálást. Legyen p egy racionális ágens hitfüggvénye egy adott pillanatban, ahol most feltételezzük, hogy p a nyelv Lindenbaum–Tarski-algebráján van értelmezve. A konzisztencia követelménye kimondja, hogy p legyen valószínűségi mérték. Legyen b egy bizonyíték, az eseménytér egy eleme, amelynek az ágens birtokába kerül. Ez a bizonyíték megváltoztatja az ágens hitfüggvényét. Az új p' hitfüggvény szintén valószínűségi mérték lesz; olyan, amelyre $p'(b) = 1$. Ilyen valószínűségi mérték sok van. A bayesiánusok ezek közül az alábbi mértéket választják ki:

Szigorú kondicionálás: Egy b bizonyíték megszerzésével egy racionális ágens valószínűségi mezeje az alábbi kondicionálással változik.

$$p'(\,\cdot\,) \quad = \quad p(\,\cdot\,|b)$$

A kondicionálás feltételét Richard Jeffrey (1983) olyan esetekre is alkalmazta, amikor az ágens nem egy új bizonyítékról szerez tudomást, hanem egymást kizáró lehetőségek valószínűségéről. Vagyis létezik az eseménytérnek egy olyan $\{b_i\}_{i=1}^n$ partíciója, amelyre nézve az új p' valószínűségi változó értékei ismertek. Ekkor az új valószínűségi mértéket Jeffrey szerint az alábbi összefüggés definiálja:

Jeffrey-kondicionálás: Egy $\{e_i\}_{i=1}^n$ partíció új p' valószínűségeinek ismeretében az új valószínűségi mezőt az alábbi összefüggés definiálja:

$$p'(\,\cdot\,) \quad = \quad \sum_{i=1}^{n} p(\,\cdot\,|e_i)p'(e_i)$$

A szigorú kondicionálás a Jeffrey-kondicionális speciális esete abban az esetben, amikor a partíció egyik elemének valószínűsége 1.

A konzisztencia igazolásának fentebb két stratégiáját is láttuk: a *Dutch book*-argumentumot illetve a várható haszon maximalizálását. Maradjunk most a *Dutch book*-argumentumnál. Az argumentum két dolgot előfeltételezett. Egyfelől azt, hogy az ágens hitfüggvénye fogadási hajlandóságot tükröz, és fogadási kvóciensekben operacionalizálható. Másfelől azt, hogy a racionális viselkedés a biztos vesztés elkerülése. Az igazolás a Ramsey–de Finetti-tétel (5. tétel) segítségével történt, illetve a feltételes valószínűség operacionalista bevezetése esetében a Ramsey–de Finetti-tételnek a feltételes valószínűségekre való általánosítása (6. tétel) révén. A konzisztenciát igazoló *Dutch book*-argumentumokat a bayesiánusok *szinkron Dutch book*-argumentumoknak nevezik.

A kondicionálás igazolása szintén *Dutch book*-argumentumokkal történik. Ezek az ún. *diakron Dutch book*-argumentumok. A diakron *Dutch book*-argumentumok ugyanazokkal az előfeltevésekkel élnek, mint a szinkron *Dutch book*-argumentumok. Ha az előfeltevéseket elfogadjuk, akkor megmutatható, hogy az ágens csak akkor jár el racionálisan, ha hitfüggvényeinek kinematikáját a Jeffrey-kondicionálás írja le. A diakron *Dutch book*-argumentumok a következőképpen vannak értelmezve. Tegyük fel, hogy az X játékos hitének mértéke egy b bizonyíték tudomásul vétele során a p kvóciensfüggvényről a p' kvóciensfüggvényre változik. Tegyük fel, hogy p is és p' is valószínűségi mérték, és tegyük fel azt is, hogy mind a $p(\cdot|\cdot)$ mind a $p'(\cdot|\cdot)$ feltételes valószínűségek (6.2) szerint vannak értelmezve. p és p' ellen tehát külön-külön nem adható szinkron *Dutch book*. Kérdés azonban, hogy van-e valamilyen megszorítás a $p \to p'$ átmenetre. A megszorítást éppen a diakron *Dutch book*-argumentum jelenti.

Az X játékos megadja a $p : \Sigma \to \mathbb{R}$ fogadási kvócienseit b tudomásul vétele előtt, és ugyanakkor megadja a $p' : \Sigma \to \mathbb{R}$ fogadási kvócienseit is arra az esetre, ha b bekövetkezne. Ezek ismeretében az Y játékos megad egy $S : \Sigma \to \mathbb{R}$ tétfüggvényt, és kiválaszthatja a kezdeti események egy tetszőleges halmazát, amelyre fogadást ajánl X-nek, majd pedig azt is megmondja, hogy mely további eseményekre illetve milyen S' tétekkel kíván fogadni *abban az esetben, ha b bekövetkezik*. Ezek a fogadások természetesen már X új fogadási kvócienseivel történnek. A diakron *Dutch book*-argumentum azt igyekszik megmutatni, hogy ha X nem a Jeffrey-kondicionálás szerint frissíti valószínűségi mezejét miután egy b bizonyíték birtokába jut, akkor Y képes úgy eseményeket és tételeket választani, hogy bármi történjék, X mindig veszíteni fog. Az alábbiakban a szigorú kondicionálásra mutatunk diakron *Dutch book*-argumentumot, azonban a gondolatmenet Jeffrey-kondicionálásra ugyanúgy érvényes.

Jelöljük a b tudomásul vétele előtti fogadás tárgyát képező eseményeket rendre F-fel, b tudomásul vétele utániakat pedig F'-vel. Legyen H a két halmaz uniója. Ekkor X össznyeresége az I interpretációban

$$G^I(H) \;=\; \sum_{a\in F}\big(I(a)-p(a)\big)S(a)) + \sum_{a'\in F'}\big(I(a')-p'(a')\big)S'(a')\,I(b). \tag{6.3}$$

Mindezek segítségével a diakron konzisztencia így értelmezhető.

15. Definíció. X-nek a p és p' kvóciensfüggvényeivel szemben Y-nak egy S és S' tétfüggvényét *diakron Dutch book*nak nevezzük F-en, ha $G^I(F) < 0$ minden interpretációban. Ha p és p' ellen nem választható diakron *Dutch book* F-en, akkor őket *diakron F-konzisztensnek* mondjuk. p-t és p'-t *diakron konzisztensnek* mondjuk, ha konzisztens $\Sigma \times \Sigma$ minden részhalmazán.

Ismét nézzünk egy példát szinkron konzisztens, de diakron inkonzisztens fogadásra:

3. Példa. Legyen Σ ismét a kockadobások eseménytere, és válassza X p-nek ismét az uniform valószínűségi mértéket Σ-n, vagyis legyen $p(\{i\}) = \frac{1}{6}$ minden $1\ldots 6$ eseményre. Legyen továbbá p' olyan, hogy $p'(\{i\}) = \frac{1}{4}$, ha $i > 2$, és $p'(\{i\}) = 0$, ha $i \leqslant 2$. A feltételes valószínűségek legyenek (6.2) alapján értelmezve p-ben is és p'-ben is. Legyen a bizonyíték $b \equiv \{\{2\} \vee \{4\} \vee \{6\}\}$ a párosdobás eseménye. Miután tehát X tudomást szerez arról, hogy a dobás páros, kezdeti p kvóciensfüggvényét p'-re frissíti, ami nyilvánvalóan nem azonos $p(\cdot|b)$-vel. Legyen továbbá $a \equiv \{6\}$ egy másik esemény, a hatos dobás eseménye. Nyilvánvalóan $p(a) = \frac{1}{6}$, $p(b) = \frac{1}{2}$, $p(a|b) = \frac{1}{3}$ és $p'(a) = \frac{1}{4}$. Ekkor Y az alábbi diakron *Dutch book*ot tudja adni X ellen az $F = \{a \wedge b, b, b^{\perp}\}$ és $F' = \{a\}$ választással. Legyen a tétfüggvény

$$\begin{aligned} S(a \wedge b) &= 1 \\ S(b) &= \frac{1}{12} \\ S(b^{\perp}) &= \frac{1}{3} \\ S'(a) &= -1 \end{aligned}$$

Ekkor az X játékos nevezési díja $-\frac{1}{6} \times 1 - \frac{1}{2} \times \frac{1}{12} - \frac{1}{2} \times \frac{2}{3} + \frac{1}{4} \times 1 = -\frac{3}{24}$; kifizetéseit pedig az alábbi táblázat mutatja:

Interpretációk	$a \wedge b$	b	$b^{\perp}$	a	Összesen
I(a)=1, I(b)=1	1	$\frac{1}{12}$	0	-1	$\frac{1}{12}$
I(a)=1, I(b)=0	0	0	$\frac{1}{3}$	$-\frac{1}{4}$	$\frac{1}{12}$
I(a)=0, I(b)=1	0	$\frac{1}{12}$	0	0	$\frac{1}{12}$
I(a)=0, I(b)=0	0	0	$\frac{1}{3}$	$-\frac{1}{4}$	$\frac{1}{12}$

ahol az a eseményre fogadás b bekövetkezése esetén az új kvóciens és tét szerint történik. Így tehát X nyereménye minden esetben $-\frac{3}{24} + \frac{1}{12} = -\frac{1}{24}$, azaz minden esetben azaz veszíteni fog. Mindez (6.3)-be helyettesítve közvetlenül is látható.

A diakron *Dutch book*-argumentum a következő tételre épül.

8. Tétel. p és p' akkor és csak akkor diakron konzisztens, ha minden $a \in \Sigma$ elemre fennáll, hogy

$$p'(a) = p(a|b).$$

Bizonyítás. Lásd a C. Függelékben.

A kondicionálás tehát diakron *Dutch book*-argumentummal igazolható.

A kondicionálással függ össze egy másik fontos bayesiánus fogalom, az ún. *felcserélhetőség*. Amint azt a fejezet elején említettük, de Finetti mind a klasszikus illetve logikai interpretációt, mind pedig a valószínűség objektív értelmezéseit elutasította; az előbbieket az indifferencia kétes elve, az utóbbit az objektív valószínűség tételezése miatt. Mindezek ellenére de Finetti mégis érvényesnek tartotta mindazokat a valószínűségi gondolatmeneteket, amelyek a tapasztalatból való induktív tanulás lehetőségét támasztották alá, sőt mi több, egyenesen azt állította, hogy "a szubjektív valószínűség az indukció problémáját felcserélhetőség esetén teljesen megoldotta." (Kyburg, Smokler, 1980, 105. o.) A felcserélhetőség fogalmát legjobban a tapasztalatból való tanulás híres példáján, Laplace rákövetkezési szabályán világíthatjuk meg. E szabály szerint, ha egy esemény ezidáig egymástól függetlenül n-szer bekövetkezett, és az esemény valószínűségéről semmit sem tudunk, akkor annak valószínűsége, hogy a következő alkalommal is bekövetkezik $\frac{n+1}{n+2}$. Ha egy fehér és fekete golyókat tartalmazó ismeretlen összetételű urnából ezidáig n-szer fekete golyó húztunk, akkor annak valószínűsége, hogy a következő húzásra is feketét fogunk kapni $\frac{n+1}{n+2}$. A szokásos metafizikai olvasatban (ld. 66. oldal) a szabály bizonyítása három dolgot feltételez:

(i) *Objektív valószínűség.* Létezik egy x objektív valószínűség, amely az urnából való fekete golyó kihúzásának valószínűsége.

(ii) *Függetlenség.* Az urnából történő húzások független események.

(iii) *Indifferencia elve.* Az x objektív valószínűség értékére vonatkozó minden h_x hipotézis episztemikus valószínűsége minden $x \in [0, 1]$ valós számra vonatkozóan egyforma.

De Finetti mindhárom premisszát elutasította. Objektív valószínűség egyáltalán nem létezik, amint azt a *Theory of Probability* elején félreérthetetlen egyértelműséggel kijelenti. Ami az urnából való fekete golyó húzásának ismeretlen valószínűségét illeti, de Finetti erről a következőket írja: „Amit nem ismerünk, az az urna összetétele, nem pedig az összetétel valószínűsége: ez utóbbi ismert és az összetételt illető szubjektív véleményünktől függ." (1995, 214. o.) Az összetétel valószínűsége maga nem egy esemény, így nem is vonatkoztathatunk rá újabb valószínűséget. „Minden események valószínűségére vonatkozó kijelentés pusztán valaki véleményének a kifejezése, nem pedig egy esemény. Ezért nincs értelme megkérdezni, hogy vajon a kijelentés igaz, hamis vagy valószínű-e." (1972, 189. o.) A rákövetkezés szabályának bizonyításában (64. oldal) szereplő h_x kifejezés tehát nem egy hipotézist reprezentál, hanem az urna egy lehetséges összetételét, a $p(h_x)$ kifejezés pedig nem ennek a hipotézisnek a valószínűségét, hanem egy lehetséges összetétel (szubjektív) valószínűségét.

Ennélfogva a $p(h_x)$ kifejezés nem is mindig értelmes. Ha például a hipotézist a napfelkeltékre vonatkoztatjuk, ahogy azt eredetileg Laplace tette, akkor de Finetti szerint az urnás modellel ellentétben nincs okunk különböző valószínűségeket feltételeznünk, minthogy nem léteznek semmilyen objektív – az urna lehetséges összetételeihez hasonló – tényállások, amelyekre vonatkozóan véleményeket, és így valószínűségeket fogalmazhatnánk meg. Az indifferencia elve tehát a hasonló heurisztikákkal együtt de Finetti szemében csupán annak a hiábavaló erőfeszítésnek az eszköze, hogy a valószínűség fogalmát valamilyen egyéb logikai-matematikai fogalomból származtassuk. Valószínűségi jóslataink szükségképpen valamilyen reális fizikai eseményekre vonatkozó kezdeti valószínűségekből származnak, amelyeket aztán a tapasztalatnak megfelelően módosítunk. Ezek a kezdeti valószínűségek azután a tapasztalatok szaporodtával konvergálni fognak, vagy amint a bayesiánus terminológia fogalmaz, az evidenciák „kimossák” a kezdeti valószínűségeket. Ez a konvergencia azonban nem jelenti azt, hogy létezne valamiféle objektív valószínűség, amelyhez a szubjektív valószínűségek konvergálnak. Éppen fordított a helyzet: bizonyos valószínűségeket azért tartunk objektívnak, mert a rájuk vonatkozó szubjektív valószínűségek konvergálnak.

A tétel, amelyre de Finetti értelmezését építi, a híres reprezentációs tétel, amelyet a valószínűségszámításban azóta is az ő neve alatt tartanak számon. A reprezentációs tétel felcserélhető illetve feltételesen független valószínűségi változók között teremt kapcsolatot. Valószínűségi változók egy halmazát akkor nevezzük *felcserélhetőnek*, ha együttes eloszlásuk a változók minden permutációjára invariáns, és akkor nevezzük egy adott valószínűségi változóra nézve *feltételesen függetlennek*, ha függetlenek a valószínűségi változó által generált feltételes valószínűségi mértékben. (A definíciókat és a hozzájuk kapcsolódó tételeket lásd részletesen a D. Függelékben.) Bernoulli valószínűségi változók esetében a két fogalom között a reprezentációs tétel az alábbi összefüggést teremti:[3]

9. Tétel. (De Finetti reprezentációs tétele.) Legyen (Ω, Σ, p) valószínűségi mértéktér, és legyen $g_1, g_2, \dots : \Omega \to \{0,1\}$ felcserélhető Bernoulli-változók egy végtelen sorozata. Ekkor a valószínűségi változók minden véges n elemű halmazának eloszlása előáll mint valamilyen $h : \Omega \to [0,1]$ valószínűségi változóra nézve feltételesen független Bernoulli-változók keveréke, vagyis

$$p(g_1 = \varepsilon_1, g_2 = \varepsilon_2, \dots g_n = \varepsilon_n) \quad = \quad \int_{\mathbb{R}} \prod_{i=1}^{n} f_{g_i|h}(\varepsilon_i \,|\, x)\, dF_h(x) \qquad (6.4)$$

ahol $f_{g_i|h}$ a g_i valószínűségi változónak a h valószínűségi változóra vett feltételes sűrűségfüggvénye, az F_h eloszlásfüggvény pedig a „relatív gyakoriságok” határértéke:

$$F_h(x) \quad = \quad \lim_{n\to\infty} p\left(\frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \leqslant x\right). \qquad (6.5)$$

A tétel a szubjektivista valószínűségi interpretáció központi tétele. Itt csak Bernoulli-változókra mondtuk ki, de érvényessége kiterjed tetszőleges felcserélhető valószínűségi változókra. A tétel azt állítja, hogy felcserélhető valószínűségi változók (legalábbis a végtelen esetben!) mindig előállnak mint feltételesen független valószínűségi változók keverékei. Mivel a viszony visszafelé triviális, azaz feltételesen független valószínűségi változók keverékei felcserélhetők, ezért a két fogalom (széles körben) azonosítható.

Lássuk a tétel értelmezését de Finetti szellemében. Induljunk ismét az urnás példától. Egy ismeretlen arányú fekete és fehér golyókat tartalmazó urnából visszatevéssel golyókat húzunk. Az objektivista – tévesen – abból indul ki, hogy (i) a fekete golyó kihúzásának objektív valószínűsége van; (ii) a húzások erre az ismeretlen objektív valószínűségre nézve

[3] A reprezentációs tétel Birkhoff, Khintchine és Hopf munkássága nyomán mára teljesen beleolvadt az ergodikus dekompozíció problémájába.

feltételesen függetlenek; (iii) a húzások valószínűségét csak valószínűség erejéig ismerjük. A reprezentációs tétel de Finetti szerint azonban éppen azt mutatja meg, hogy ezek a feltevések teljességgel jogosulatlanok. A (6.4) egyenlet jobb oldalán álló kifejezések, pl. a feltételes sűrűségfüggvény vagy a h valószínűségi változó eloszlásfüggvénye csupán matematikai segédfogalmak, amelyek nélkülöznek minden fizikai-metafizikai jelentést, és így nem értelmezhetjük őket úgy, mint „egy hipotézisre nézve független események" vagy „konstans, de nem ismert valószínűségek". Ezek a fogalmak a felcserélhető valószínűségi változóknak csak egy bizonyos reprezentációjában jelennek meg; fizikai jelentése azonban egyedül magának a bal oldalon szereplő felcserélhetőségnek van.[4]

A felcserélhetőség az urnás példa esetében azt jelenti, hogy annak a valószínűsége, hogy n húzásból k fekete golyót kapunk nem függ attól, hogy pontosan mely k húzás során kaptunk fekete golyót. A valószínűség itt természetesen szubjektíven értendő, vagyis az eloszlásfüggvény azért felcserélhető, mert „úgy érezzük, hogy a következő elmeállapotban vagyunk: két húzássorozatot, amelyek csak a sorrendükben különböznek, egyenlően lehetségesnek ítélünk" (1989, 201. o.) Elegendő tehát a felcserélhetőséget feltennünk, és a fenti tétel alapján az összes többi matematikai eszköz a rendelkezésünkre áll, hogy a rákövetkezési elvét és a hasonló indukciós elveket levezessük. A (6.5) egyenlet pedig éppen azt fogja garantálni, hogy tetszőleges kezdeti valószínűségből indulhatunk; n növekedtével a g_i valószínűségi változók aritmetikai átlaga az eloszlásfüggvényhez konvergál. Éppen emiatt a konvergencia miatt vélik az objektivisták – ismét csak tévesen –, hogy létezik valamiféle objektív valószínűség.

Ha tehát felcserélhető kezdeti eloszlásfüggvényből indulunk ki, akkor „az a mentális hajlamunk, hogy a jövőt a múlthoz hasonlónak várjuk" (uo.) igazolhatóvá válik, azaz "az indukció problémáját felcserélhetőség esetén teljesen megoldott[uk]" – hangzik de Finetti ambiciózus kijelentése.

A konzisztencia és a kondicionálás alapelveinek bevezetésével a bayesiánusok egy sereg tudományfilozófiai és ismeretelméleti problémára (Hempel-konfirmáció, holló-paradoxon, stb.) kínáltak választ. Ugyanakkor a szakirodalom számon tart néhány olyan problémát is (univerzális állítások konfirmációja, „a régi bizonyíték" problémája), amelyeket a bayesianizmus sem volt képes kezelni. Hogy ezeknek az ismert bayesiánus problémáknak a kezelése mennyiben sikeres illetve sikertelen, azt itt nem tárgyaljuk (ehhez ld. Earman 1992), hanem ehelyett közvetlenül a bayesianizmus empirikus alapjaira kérdezünk rá.

6.4. Empirikus alapok

Elérkeztünk ahhoz a mindeddig halogatott kérdéshez, hogy miképpen igazolhatók a bayesiánus alapelvek vagy a mögöttük álló *Dutch book*-argumentumok illetve döntéselméleti

[4] A felcserélhetőség fogalma de Finetti radikális szubjektivista valószínűségértelmezésében mindvégig központi szerepet játszott. A fogalmat számtalan módon általánosította (parciális felcserélhetőség, Markov-felcserélhetőség, stb.).

bizonyítások premisszái. Láttuk, hogy a racionalitásra való puszta hivatkozás nem jelent megnyugtató megoldást. Egy út maradt hátra: megvizsgálni, hogy mi a helyzet a premisszák empirikus igazságával. A vonatkozó pszichológiai irodalom óriási; itt csak néhány fontos eredményt közlünk (Kahneman et al., 1982). Kezdjük azokkal az általános ismeretekkel, amelyeket események valószínűségi értékeléséről tudunk. Az empirikus kutatások szerint a szubjektív valószínűségi becsléseket néhány heurisztikus elv irányítja, amelyek a legtöbb esetben nagyon hasznosak, azonban az egyéb észlelési heurisztikákhoz hasonlóan – ahogyan például a távolságérzékelés megtéveszthető a megvilágítás erejének változtatásával – komoly és szisztematikus tévedésekhez vezethetnek. A három legismertebb ilyen heurisztika a *reprezentativitás*, a *hozzáférhetőség* és a *horgonyzás és hangolás*.

1. Reprezentativitás. A reprezentativitás heurisztikája alapján egy szinguláris eseményről olyan valószínűséggel véljük, hogy egy eseménytípusba esik, amennyire az eseményt reprezentatívnak érzékeljük a típus szempontjából. Ez a reprezentativitási heurisztika azonban számos statisztikus faktorral egyáltalán nem számol. Az egyik ilyen figyelmen kívül hagyott tényező az *előzetes valószínűség*. Vizsgált személyeknek rövid személyleírásokat mutatnak, majd arra kérik őket, hogy állapítsák meg, hogy a személyleírás milyen valószínűséggel felel meg inkább egy mérnöknek, mint egy ügyvédnek. A kísérletek azt mutatják, hogy az eredmény gyakorlatilag nem változik akkor, ha az alanyt előzetesen informálják arról, hogy a személyleírások egy meghatározott százaléka mérnökre vonatkozik. A személyleírások teljes hiányában az alanyok figyelembe veszik az előzetes valószínűséget, de a reprezentativitási heurisztika akár semmitmondó leírások esetén is rögtön törli az előzetes valószínűségeket (Kahneman, Tversky, 1973).

A reprezentativitási heurisztika alkalmazásának egy másik bizonyítéka a becsült valószínűség érzéketlensége a *mintaméretre*. A vizsgált személyekhez a következő kérdést intézik. „Egy város nagyobbik kórházában átlagosan 45 gyerek születik naponta, a kisebbikben 15. A gyerekeknek átlagosan a fele fiú. Mindkét kórházban egy éven át rögzítik azokat a napokat, amelyeken az aznap születettek gyerekek több, mint 60%-a fiú. Melyik kórházban rögzítettek több ilyen napot?" (Kahneman, Tversky, 1972). A megkérdezettek döntő többsége azt válaszolja, hogy a kórházak ebből a szempontból egyformák, mivel az eseményeket ugyanaz a statisztika írja le, és ezért ugyanannyira vélik reprezentatívnak az általános populációra nézve.

Ezzel függ össze az a szisztematikus torzító faktor is, hogy a populációk statisztikus tulajdonságait már kis mintákon reprezentatívnak véljük. Véletlen folyamatokkal generált események statisztikai jellemzőiről például az alanyok nagy része úgy gondolja, hogy az már *rövid sorozatokon* is reprezentálódik. Egy szabályos pénzérmével való dobásnál a $fifiif$ sorozatot valószínűbbnek gondolják az $fffiii$ sorozatnál, mivel ez nem véletlen, és a $ffffif$ sorozatnál is, mivel ez nem *fair* (Kahneman, Tversky, 1972). A „kis számok törvényére" hasonló példa a szerencsejátékos tévedése is, aki egy hosszú piros sorozat után a roulette-en a feketét valószínűsíti. A valószínűségre mindegyik esetben egyfajta önkorrigáló folyamatként tekintenek.

2. Hozzáférhetőség. A hozzáférhetőség heurisztikája alapján egy esemény valószínűségét az alapján ítéljük meg, hogy milyen *könnyen* vagyunk képesek felidézni az eseménytípusba tartozó egyedi eseményeket. Ez a heurisztika nyilván hasznos, mivel gyakori események felidézése általában könnyebb; azonban a hozzáférést a frekvencián kívül egyéb faktorok is befolyásolják, így a heurisztika számos esetben torzíthat. Alanyoknak híres emberek listáját olvassák fel, majd azt kérik tőlük, hogy becsüljék meg, hogy hányan voltak a felsoroltak közül nők és hányan férfiak. Az arány jelentősen változik annak függvényében, hogy inkább a nők vagy a férfiak a relatíve ismertebbek (Tversky, Kahneman, 1973).

A hozzáférhetőségre támaszkodó valószínűségi becslés függ a keresési halmaztól is. Arra a kérdésre, hogy vajon az angol (háromnál több betűs) szavak körében az r betűvel kezdődő vagy az r betűt harmadik betűként tartalmazó szavak vannak-e nagyobb számban, a kérdezettek zöme az első mellett dönt (helytelenül), mivel r betűvel kezdődő szavakat könnyebb felidézni (Tversky, Kahneman, 1973).

Olykor a típusba tartozó eseteket nem felidézzük, hanem generáljuk. Itt a generáló algoritmus könnyűsége jelenti a tévedés forrását. Kísérleti személyektől megkérdezik, hogy egy 10 tagú csoportból milyen $2 \leqslant k \leqslant 8$ számra lehet a legtöbb különböző tagokból álló bizottságot létrehozni. A megkérdezettek általában az alacsony k-t preferálják, és a bizottságok számát k-val csökkenőnek ítélik. Feltehetően arról van szó, hogy az alcsoportok (számolás nélküli) mentális konstrukciójában első lépésként a 10-es csoportot diszjunkt részekre particionálják, így a kisebb számok több lehetőséget jelentenek (Tversky, Kahneman, 1973).

Végül események közötti korreláció erősségét az eseménytípusok közötti asszociációs erő alapján ítéljük meg. Alanyoknak képzeletbeli mentálisan sérült betegek diagnózisát illetve személyekről készült rajzait mutatják meg, később pedig arra kérik őket, hogy idézzék fel, mely betegek kapcsán tudták a diagnózisokat a rajzokon látható jegyekkel azonosítani (pl. gyanakvás – különös szemek). Az alanyok a diagnózis és a rajz közötti korrelációt rendszeresen felülbecslik az emberrajzra vonatkozó közkeletű ugyanakkor megalapozatlan interpretációk alapján (Chapman, Chapman, 1969).

3. Hangolás és horgonyzás. Sok esetben a becslés valamilyen számolt vagy készen kapott kezdeti értékből indul, majd pedig az érték újrahangolásával jut a végső válaszhoz. Ez a hangolás és horgonyzás heurisztikája. A kísérleti személyeket arra kérik, hogy bizonyos arányokat becsüljenek meg, pl. az afrikai országok arányát az ENSZ-ben. A becslés a következőképpen történik. A kísérletvezető megforgat egy 1-től 100-ig számozott kereket, majd arra kéri az alanyt, hogy nyilatkozzon, az általa adott becslést nagyobb vagy kisebb-e a kapott számnál, majd pedig lefelé vagy felfelé indulva a számtól adja meg a becslését. A fenti kérdésre a 10-es számról induló alanyok mediánja 25 volt, míg a 65-ről indulóké 45. A készen kapott kezdeti érték tehát jelentősen módosította a végértéket.[5]

[5]Vagy egy hasonló érdekes kísérlet (szintén nem a valószínűségi becslések köréből) a következő: Két csoport előtt a táblára egy-egy szorzat van felírva: az egyikben a számok 1-tól 8-ig, a másodikban 8-től 1-ig vannak összeszorozva. Mindkét csoportban az alanyok csupán pár másodpercet kapnak a szorzás elvégzésére,

A hangolás és horgonyzásnak a valószínűségi becsléseinkben játszott szerepe illusztrálására különösen érdekes az a kísérlet, amelyben egyszerű, konjunktív és diszjunktív események szubjektív valószínűségét tanulmányozták. Kísérleti személyek urnából húzott golyókra vonatkozó események közül vagy egy egyszerű vagy egy konjunktív vagy egy diszjunktív eseményre fogadhattak. Az egyszerű esemény egy piros golyó húzása egy 50-50 százalékban piros illetve fehér golyókat tartalmazó urnából ($p = 0,5$); a konjunktív esemény hét piros golyó húzása egy 90 százalékban piros, 10 százalékban fehér golyókat tartalmazó urnából ($p = 0,48$); a diszjunktív esemény pedig legalább egy piros golyó húzása hét alkalomból egy 10 százalékban piros, 90 százalékban pedig fehér golyókat tartalmazó urnából ($p = 0,52$). Az alanyok szignifikáns többsége a konjunktív eseményt preferálta az egyszerű eseménnyel szemben, és az egyszerű eseményt a diszjunktív eseménnyel szemben, vagyis a konjunktív események valószínűségét fölébecsülte, a diszjunktívokét pedig alá. A magyarázat ez esetben is a horgonyzás és hangolás. Mind a konjunktív, mind pedig a diszjunktív események valószínűségét az egyszerű események valószínűsége alapján becsüljük. Mivel azonban a konjunktív események valószínűsége kisebb, mint a konjunkció tagjaié, míg a diszjunktív események valószínűsége nagyobb, mint a tagoké, így az első esetben fölé-, a második esetben alábecslés történik (Bar-Hillel, 1973). Igen érdekes következményekhez vezet a horgonyzás döntéselméleti kontextusban. Mivel konjunktív események valószínűségi becslése elsősorban a tervezésnél jelentkeznek, ahol sok tényezőnek kell együttesen fennállnia, hogy a terv sikerüljön, így a horgonyzás miatt a tervezésnél hajlamosak vagyunk a terv sikerét optimistán felülbecsülni. Ugyanakkor a kockázatelemzés diszjunktív eseményekre épül, hiszen egy katasztrófa sokféleképpen bekövetkezhet. Ismét csak a horgonyzás miatt a kockázati tényezőt hajlamosak vagyunk alulbecsülni.

Mindezek fényében vizsgáljuk meg végül egy konkrét esetben, hogy mi is tudható empirikusan. Tekintsük a valószínűség Savage-féle döntéselméleti bevezetését. Láttuk, hogy Savage hét, a cselekedetek preferencia-sorrendjére kirótt racionális posztulátumból vezeti le, hogy az állapotok terén létezik egy valószínűségi mérték. Vegyük az egyik legtöbbet vizsgált posztulátumot, a függetlenségi feltételt. A feltétel azt mondja ki, hogy ha két cselekedetnek ugyanaz a következménye valamely állapotban, akkor a cselekedetek közötti preferenciák nem függhetnek ettől a következménytől. Az elv megvilágítására tekintsük a következő Allaistól (1979) származó kísérletet. Egy urnából golyókat kell kihúzni, amelyek $1-100$-ig vannak számozva. A húzáshoz két különböző döntési probléma csatlakozik. Az A problémában egy a_1 és egy a_2 cselekedet közül lehet választani, amelyek a megfelelő számú golyó kihúzása után az alábbi kifizetéseket adják:

A	1	2-11	12-100
a_1	1000000 \$	1000000 \$	1000000 \$
a_2	0 \$	5000000 \$	1000000 \$

ezután pedig nyilatkozniuk kell a szorzat értékéről. A várakozás kettős volt. Egyrészt az elégtelen hangolás miatt a becsült érték kevesebb lesz a ténylegesnél, másrészt a növekvő sorozat értéke a kisebb kezdeti tényezőkre történő extrapoláció miatt kisebb lesz. Az eredmények mindkét várakozást megerősítették.

A B döntési problémában egy b_1 és egy b_2 cselekedet közül lehet választani, amelyek kifizetései az alábbiak:

B	1	2-11	12-100
b_1	1000000 $	1000000 $	0 $
b_2	0 $	5000000 $	0 $

A függetlenségi feltétel tehát azt követeli meg, hogy egy racionális ágens akkor és csak akkor részesítse előnyben az a_2 cselekedetet az a_1 cselekedettel szemben, ha b_2-et is előnyben részesíti b_1-gyel szemben, hiszen a két döntési helyzet mindössze annyiban különbözik, hogy az egyikben a $12 - 100$ számozott golyókra 0$ a kifizetés, míg a másikban 1000000$. A kísérletek tanúsága szerint azonban a megkérdezettek 46% számára $a_1 \succcurlyeq a_2$, de $b_1 \preccurlyeq b_2$, vagyis a megkérdezettek csaknem fele sérti a függetlenségi feltételt. A kísérletet azóta számos formában és helyszínen megismételték hasonló eredménnyel. A felmérések szerint az Allais-paradoxonban a megkérdezettek az A döntési helyzetet általában úgy értelmezik, mint amelyikben a biztos 1000000$-os nyereséget kockáztatják, ha a_1 helyett az a_2-t választják; míg a B döntési helyzetet úgy, hogy egyrészt szinte biztos, hogy nem nyernek, másrészt 5000000$-t ugyanolyan jó eséllyel nyernek, mint 1000000$-t, így inkább választják b_2-t. A bayesiánusok szerint amennyiben ezek a hallgatólagos pszichológiai feltételek (a bizonyosat előnyben részesíteni a bizonytalannal szemben, stb.) is tematizálva lennének a döntési helyzetben, akkor a függetlenségi feltétel empirikusan igazolható lenne.

A helyzetet még érdekesebbé teszi, hogy a függetlenségi feltétel levezethető egy olyan másik elvből, amelyet az empirikus kísérletek meggyőzően alátámasztanak. A kísérletben az Allais-döntést egy döntési fa alapján vázolják az alanyoknak. Vegyünk ismét egy urnát $1 - 100$-ig számozott golyókkal, amelyek közül a játékvezető egyet kihúz. Ha a golyó száma 12 és 100 között van, akkor az alany vagy 0 $-t kap, vagy 1000000 $-t attól függően, hogy egy A vagy egy B kísérletben vesz részt. Ha a száma 1 és 11 között van, akkor két lehetősége van: (i) vagy kap biztosan 1000000 $-t; vagy (ii) 5000000 $-t kap, amennyiben a golyó száma nem 1, és semmit sem kap, ha a szám 1. Az alanynak nyilatkoznia kell, hogy *most*, a játékvezető húzása előtt melyik lehetőséget választaná. A kísérletek azt bizonyítják, hogy az alanyok egyöntetűen vagy az (i) vagy a (ii) választással élnek mindkét döntési fán. Azt az elvet, amely alapján az alanyok cselekednek Maher (1993) *szinkron szeparálhatósági* elvnek nevezi. Az elv szerint jelenlegi preferenciáink arra vonatkozóan, hogy hogyan fogunk választani egy későbbi döntési pontban, nem függnek attól, hogy mi történne akkor, ha nem érnénk el a döntési pontot.

A kísérletek tehát azt mutatják, hogy szinkron szeparálhatósági feltétel empirikusan verifikálható, a függetlenségi feltétel azonban nem, jóllehet az utóbbi logikai következménye az előbbinek. A bayesiánusok számára egy további lehetséges manőver ilyenkor azt állítani, hogy azok az alanyok, aki elfogadják a szinkron szeparálhatósági feltételt, de nem fogadják el a függetlenségi feltétel, egyszerűen irracionálisak. Ez a lépés azonban teljes mértékben kiüresíti a racionalitás fogalmát.

A fenti kísérleti pszichológiai eredmények fényében a valószínűség szubjektív interpretációjáról összefoglalóan tehát a következőket mondhatjuk. A szinkron illetve diakron *Dutch book*-tételek a valószínűség, a feltételes valószínűség illetve a kondicionálás egyfajta játékelméleti reprezentációját adják. Ahhoz, hogy ezeket a modelleket az empirikus ágens hitfüggvényével azonosíthassuk, két empirikus hipotézissel kell élnünk. Egyfelől fel kell tennünk, hogy a hitfüggvények helyesen reprezentálhatók fogadásokkal, másfelől fel kell tételeznünk, hogy az empirikus ágens nem fogad úgy, hogy bármi történik, mindig veszítsen. Az első premisszát számos kísérletben cáfolták rámutatva, hogy a fogadás ténye visszahat a kiváltó hitállapotra, eltorzítva a fogadás előtti hitet. A második premissza ellen a Maher-típusú példák szolgáltatnak ellenpéldát: az empirikus ágens néha a konzisztencia figyelmen kívül hagyásával fogad; ráadásul erre sok esetben igen jó oka is van.

A szubjektivista persze tovább menekülhet, és tágíthatja a *Dutch book*-típusú elveket. Mondhatja azt, hogy az empirikus ágens viselkedését nem a *Dutch book*-mentesség kritériumai írják le jól, hanem valamilyen általánosabb döntéselméleti kritériumok, például a várható haszon maximalizálása. Megadhat ilyen kritériumokat, és ezek matematikai reprezentációjából levezetheti a valószínűséget. A valószínűségnek ezek a levezetései azonban újfent tartalmazhatnak olyan premisszákat, amelyeket az empirikus ágensek viselkedései, vélekedései sértenek. Erre mutattak példát a függetlenségi feltételt sértő Allais-döntések, amelyek útját állták a valószínűség empirikus visszavezetésének a várható haszon maximalizálására. Továbbra is nyitott tehát a kérdés, hogy a valószínűség levezethető-e olyan általános, a hívő hiteire és cselekedeteire vonatkozó premisszákból, amelyek empirikusan kielégítően igazolhatók.

A szubjektivista természetesen megteheti, hogy az interpretáció egy pontján hirtelen átvált az empirikus ágensről a racionális ágensre, és azt állítja, hogy mindezek a premisszák nem az empirikus, hanem a racionális, ideális ágensre igazak. Nagyon gyakori eljárás ez a szubjektivisták között, és ugyancsak nagyon gyakori a két nézet lebegtetése. Amennyiben azonban az empirikus ágenst racionális ágensre cseréljük le, annyiban elhagyjuk a szubjektivista interpretációt is, és visszatérünk a logikai interpretációhoz – mindössze annyi történik, hogy a konfirmáció empirikusan nem értelmezett fogalmát a konzisztencia hasonlóan nem értelmezett fogalmával helyettesítjük. A szubjektív interpretáció tehát nem képes beváltani azt a reményt, amelyet a valószínűség értelmezésével kapcsolatban táplál.

7. fejezet

A relatív gyakoriság-interpretáció

A valószínűség eddigi három interpretációja mind összhangban állt azzal a prefilozofikus elképzeléssel, hogy valószínűsége elsősorban egyedi eseményeknek van. Ez természetesen nem jelentette azt, hogy események egy sorozatához vagy egy eseménytípushoz ne lehetne valószínűséget rendelni. A naív megközelítés szerint azonban az eseménytípusokhoz rendelt valószínűség csupán származtatott fogalom: csak annyiban létezik, amennyiben az egyedi események valószínűséggel rendelkeznek. A kockadobások sorozatában bizonyos tulajdonságoknak, így például annak, hogy a dobások egy hatoda hatos lesz, csak azért van valószínűsége, mivel maguk az egyedi kockadobások rendelkeznek valószínűséggel.

A relatív gyakoriság- vagy más néven frekvenciainterpretáció azonban megfordítja ezt az intuitív sorrendet, és a valószínűséget nem a szinguláris eseményhez, hanem az eseménytípushoz rendeli. A hatos dobás valószínűsége nem más, mint a hatos relatív gyakorisága a kockadobások sorozatában. Az olyan szinguláris valószínűségi kijelentésekre vonatkozóan, mint az „Ennek a kockadobásnak a valószínűsége egy hatod" ezek után két álláspont lehetséges. A toleránsabb álláspont Reichenbach-é (1949):

> „Egy szinguláris eset valószínűségére vonatkozó kijelentést nem tekintem jelentéssel bíró kijelentésnek, hanem pusztán elliptikus beszédmódnak. Ahhoz, hogy egy ilyen kijelentés jelentéssel bírjon, le kell fordítani ismételt történések sorozatában relatív gyakoriságokra vonatkozó kijelentésre. A szinguláris eset valószínűségére vonatkozó kijelentésnek tehát *átvitt jelentése* van, amelyet *az általános esetről a partikuláris esetre való jelentésátvitellel* konstruálunk." (376-77. o.)

A szinguláris valószínűségi kijelentések tehát csak annyiban értelmesek, amennyiben eseménytípusra vonatkozó kijelentésekből származnak. Ez a származtatás azonban rögtön felveti az alábbi problémát:

> „Ha megkérdezik tőlünk egy szinguláris jövőbeli eset valószínűségét, akkor először is az esetet egy megfelelő referenciaosztályba kell sorolnunk. Az egyedi dolgok vagy események azonban sok referenciaosztályba sorolhatók, amelyekből

> különböző valószínűségek származnak. Ezt az ambiguitást nevezzük a referenciaosztály problémájának."[1] (374. o.)

A referenciaosztály problémája tehát bármely olyen frekventista értelmezés problémája, amely a szinguláris kijelentéseknek jelentést tulajdonít. Elkerülni csak úgy lehet, ha a szinguláris valószínűségi kijelentésektől minden értelmet megvonunk. Ez a radikális álláspont von Mises-é (1928):

> „annak a 'valószínűségnek, hogy megnyerjük a csatát' például nincs helye valószínűségelméletünkben, mert nem tudunk elgondolni egy olyan kollektívát, amelyhez ez az esemény tartozna. A valószínűség elmélete ugyanolyan kevéssé alkalmazható erre a problémára, mint amennyire a munka fizikai fogalma alkalmazható a színész munkájára midőn eljátssza szerepét a színpadon."

Egy eseménytípus vagy más szóhasználatban egy tulajdonság valószínűsége tehát a tulajdonságot instanciáló egyedi események bekövetkezési gyakoriságával azonos egy alkalmasan választott referenciaosztályon. A frekventista interpretáció rendszerint a következőképpen halad. Mivel a valószínűséget relatív gyakorisággal azonosítjuk, ezért első lépésben a valószínűség mértékelméleti fogalmait egy másik matematikai diszciplína, a sorelmélet fogalmaival reprezentáljuk. Az alapstruktúra tehát a mérhető tér helyett a sorozat lesz. A második lépésben ezeket a sorozatokat interpretáljuk fizikailag úgy, mint bizonyos eseménytípusba tartozó szinguláris események rendezett összességét. Mindkét lépésnek megvannak a maga nehézségei. A valószínűségelmélet más matematikai struktúrákkal való reprezentációja azt a problémát veti fel, amelyet Salmon megengedhetőségi kritériumnak nevez: az így nyert struktúrák vajon a releváns szempontból izomorfak lesznek-e a valószínűség mértékelméleti megfogalmazásával. A második lépés, a sorozatok fizikai interpretációjának nehézsége abban áll, ami minden interpretáció közös nehézsége, és ami Salmon osztályozásában a kideríthetőség nevet viseli: mit is jelentenek *fizikailag* a sorozatok? A két kritériumot az alábbiakban gondosan elkülönítve tárgyaljuk.

Ami a kideríthetőséget illeti, azt kell tehát megvizsgálnunk, hogy a frekventista iskola hogyan értelmezi az interpretáció matematikai alapjául szolgáló sorozatokat. A sorozatok szinguláris események, megfigyelések rendezett összessége. Ez rögtön nyilvánvalóvá teszi azt a tényt, hogy a sorozatok révén definiált valószínűség érzékenyen fog függni a sorozatokkal reprezentált események körétől. Milyen egyedi eseményeket tekintsünk kockadobásnak?

---

[1] A referenciaosztály problémájának egy korai megfogalmazása John Venntől (1866) származik:

> „Tegyük fel például, hogy tíz angolból kilenc megsérül Madeirában tartózkodása során, de tíz tüdőbetegből kilencre jótékony hatással van ugyanez a hely. Ezek a statisztikák bár képzeletbeliek, de tökéletesen elképzelhetőek és összeegyeztethetőek. John Smith tüdőbeteg angol; vajon ajánljuk neki Madeirát vagy sem? Más szavakkal, milyen következtetést vonhatunk le a halálát illetően? Mindkét statisztikai táblázat illik rá, azonban ellentétes következményekhez vezetnek ... További adat nélkül nem jutunk döntésre." (222–223. o.)

A referenciaosztály problémájával kapcsolatban lásd még (Hájek, 2007).

Milyen magasról, milyen sebességgel kell eldobni a kockát, hogy az dobásnak számítson? A referenciaosztály rögzítése tehát előfeltétele minden frekventista definíciónak. Milyen események alkossák tehát a referenciaosztályt? Ebben a kérdésben két ellentétes álláspont létezik: az *aktuális* és a *hipotetikus* frekvenciainterpretáció.

Az aktuális frekventizmus a szigorú empirista hagyomány szellemében a sorozatokat az aktuálisan megfigyelt események halmazával azonosítja. Mivel ez a halmaz véges, ezért a matematikai struktúra a véges sorozat, amelynek elemei tulajdonságok valamilyen algebrájából vesznek fel értékeket. Egy tulajdonság (eseménytípus) valószínűsége a véges sorozatban vett relatív gyakorisága. Ahogy Venn (1866) fogalmaz: „A valószínűség semmi más, mint arány". A hatos valószínűsége tehát a kockának egy adott véges sorozatában a hatos dobás relatív gyakorisága. Ez az elképzelés jól vizsgázik a kideríthetőség kritériuma szempontjából: minden valószínűségi kijelentés szigorúan verifikálható lesz. Ugyanakkor azonban a valószínűség fogalmának ilyen szigorú operacionalizálása kiüresíti a fogalmat. Egy valószínűségi kijelentés csak az aktuálisan megfigyelt véges számú esetre alkalmazható, azon túl pedig nem általánosítható. Az aktuális frekventizmust ebben a véges formában ezért nem is képviseli senki.

A hipotetikus frekventizmus alapstruktúrája a végtelen sorozat. Az értelmezés nehézségét éppen ennek a végtelennek a fizikai értelmezése jelenti. A „hipotetikus" kifejezés arra a modális elemre utal, amely ezt az iskolát megkülönbözteti szigorúan empirista párjától. Mivel az aktuális világban a megfigyelések száma véges, ezért mondani kell valamit a valószínűséget reprezentáló sorozat végtelen számú nem aktuális eleméről. A klasszikus megoldás valamilyen modális fogalom bevezetése: a sorozat azt az eseményekből álló rendezett összességet reprezentálja, amelyhez aktuális megfigyeléseink tartoz*ná*nak, ha a kísérletet végtelenszer megismétel*nénk*. A modális elem itt a kontrafaktuális kifejezésben jut kifejezésre. A hatos dobás valószínűsége tehát az az érték, amelyet a hatos relatív gyakorisága határértékben felven*ne*, ha a kockával a végtelenségig dobál*nánk*. De hogyan értelmezzük a modális kifejezéseket? A realista megoldás a modális kifejezések mögötti ontológia komolyan vétele. A végtelen sorozatok reálisan léteznek, csak éppen nem az aktuális, hanem egy lehetséges világban. De miképpen léteznek ezek a lehetséges világok?

A gondolatmenetet nem szükséges folytatnunk. Látható, hogy a hipotetikus frekventizmus, és amibe torkollik, a modális realizmus súlyos metafizikai elköteleződéssel jár. Ezért a hipotetikus frekventizmusnak ma, a véges aktualizmushoz hasonlóan szintén nem akad képviselője. A reálisan létező interpretációk a két szélsőség között lavíroznak.

Térjünk át a megengedhetőség kérdésére. A feladat itt az, hogy a valószínűség mértékelméleti megfogalmazásának megadjuk valamilyen izomorf matematikai megfogalmazását, amelynek relatív gyakoriságként való fizikai interpretációja valamilyen okból kézenfekvőbb, mint az eredeti megfogalmazásé. A választott matematikai struktúra a legtöbb ilyen esetben a sorozat, izomorfia alatt pedig a következőt értjük. Legyen (Ω, Σ, p) valószínűségi mértéktér, és legyen $x : \mathbb{N} \to \Sigma$ egy sorozat, amely olyan szinguláris események végtelen rendezett összességét reprezentálja, amelyek mindegyike bizonyos Σ-beli tulajdonsággal jellemezhető. Egy $a \in \Sigma$ tulajdonság relatív gyakoriságának határértékét az x sorozatban jelöljük $r_x(a)$-val, és értelmezzük az alábbi módon:

$$r_x(a) \quad = \quad \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} 1_a(x_k), \tag{7.1}$$

ahol $1_a(x_k)$ karakterisztikus függvény, vagyis

$$1_a(x_k) \quad = \quad \begin{cases} 1, & \text{ha } x_k \subseteq a, \\ 0, & \text{ha } x_k \nsubseteq a. \end{cases}$$

Ekkor a valószínűségnek relatív gyakoriságként való *matematikai* reprezentációját így definiáljuk:

16. Definíció. Egy (Ω, Σ, p) valószínűségi mértéktér rendelkezik *relatív gyakoriság-modellel*, ha létezik egy olyan $x : \mathbb{N} \to \Sigma$ sorozat, ahol minden $a \in \Sigma$ tulajdonságra $r_x(a) = p(a)$.

Hangsúlyozzuk, hogy egy ilyen modell létezése még nem jelenti azt, hogy a valószínűségeknek van relatív gyakoriság-*interpretációja*.[2] Az „interpretáció” kifejezést a fentiekkel összhangban kizárólag a fizikai interpretációra tartjuk fent. A relatív gyakoriság-modell csak az első lépés a relatív gyakoriság-interpretáció felé, amelyet a modell fizikai értelmezésének kell követnie.

Végül szóljunk néhány szót a történeti háttérről és a forrásokról. A relatív gyakoriság-interpretáció először a tizenkilencedik század közepén jelent meg Cambridge-ben Robert Leslie Ellis (1849) és John Venn (1866) munkássága nyomán, mintegy a laplace-i valószínűséginterpretációra adott empirista válaszként. Igazi népszerűségre azonban csak a logikai pozitivizmus kialakulása során a Berlini kör két képviselőjénél, Hans Reichenbach és Richard von Mises révén tett szert. A frekventista tárgyaláson túl von Mises nevéhez fűződik a valószínűségelmélet első axiomatikus tárgyalása is. Az alábbiakban mi kettőjük közül az ő elképzeléseit ismertetjük. Ami von Mises interpretációját az összes többitől megkülönböztette, az az a törekvése, hogy a valószínűség tárgyalásába behozza a véletlenség fogalmát is: a határérték létezésén túl a fenti sorozatoktól von Mises még azt is megkövetelte, hogy véletlenek legyenek. Így elmélete alkalmat ad majd a valószínűség és a véletlen viszonyának megtárgyalására. Von Mises frekventizmusát az 1928-ban megjelent *Wahrscheinlichkeit, Statistik und Wahrheit*,[3] és az 1931-es nagyszabású *Wahrscheinlichkeitsrechnung und ihre*

[2] A „szabályos pénzérme” B. Függelék 1. pontjában tárgyalt (Ω, Σ, p) mértékterének (ahol $\Omega = \{f, i\}$ kételemű halmaz, $\Sigma = \mathcal{P}(\{f, i\})$ az Ω hatványhalmaza, p pedig a $p(\{f\}) = p(\{i\}) = \frac{1}{2}$ uniform valószínűségi mérték) van relatív gyakoriság-modellje, pl. az

$$x : \mathbb{N} \to \Sigma; \ \ x_k = \begin{cases} \{f\}, & \text{ha } k \text{ páros}, \\ \{i\}, & \text{ha } k \text{ páratlan}. \end{cases}$$

[3] A könyv a Bécsi kör gondozásában a Philipp Frank és Moritz Schlick által szerkesztett *Schriften zur wissenschaftlichen Weltauffassung* harmadik köteteként jelent meg.

Anwendung in der Statistik und theoretischen Physik című műveiben fejtette ki, valamint az egyetemi jegyzeteiből 1964-ben posztumusz kiadott *Mathematical Theory of Probability and Statistics*ben. A fejezetet végül egy másik frekventista álláspont, Andrei N. Kolmogorov frekventizmusának ismertetésével zárjuk, aki – amint az a *Grundbegriffe der Wahrscheinlichkeitsrechnung* (1933) című klasszikus munkájából kitűnik – implicit formában szintén a relatív gyakoriság-interpretáció híve volt.

7.1. Kollektívák

A valószínűségelmélet tárgyát von Mises[4] az alábbi módon jellemzi.

> „A valószínűségelmélet tárgya a nagyon gyakran és változatlan körülmények között végzett kísérletek vagy megfigyelések hosszú sorozatai.[5] Megfigyeljük például egy pénzérme vagy egy kockapár ismételt eldobásának kimeneteit; feljegyezzük az újszülöttek nemét egy populációban; meghatározzuk egy céltáblára irányzott egymást követő lövések találati helyeinek koordinátáit; vagy, hogy egy általánosabb példát adjunk, feljegyezzük az 'azonos folyamattal' mért 'azonos fizikai mennyiség' mérési eredményeit. Minden esetben megfigyelések valamely sorozatával van dolgunk; meghatározzuk a lehetséges kimeneteket, és feljegyezzük az aktuálisakat." (1964, 2. o.)

A jellemzésből rögtön kitűnik, hogy a tömegjelenségek reprezentációja valamilyen rendezett módon történik, azaz lehetséges eseménytípusokat (attribútumokat) instanciáló aktuális események sorozatával. A szóba jöhető sorozatok karakterizációja a tömegjelenségek empirikus jegyeinek vizsgálatai révén válik lehetővé. A valószínűségelmélet fogalomkörébe eső jelenségeket von Mises szerint két ilyen jegy jellemez. Az első a relatív frekvencia stabilitása:

> „A valószínűségelmélet szempontjából lényeges, hogy a tapasztalat szerint a kockajátékban, és minden egyéb tömegjelenségben, amelyet említettünk, bizonyos attribútumok relatív gyakorisága egyre stabilabbá válik a megfigyelések számának növekedtével."[6] (1964, 108. o.)

[4]Nem tudjuk megállni, hogy meg ne említsünk egy, az adott interpretációt nem érintő, de von Misestől származó paradoxont, a *születésnapok paradoxonát.* Tudjuk, hogy 366 ember közül legalább kettőnek biztosan egy napra esik a születésnapja. Hány ember esetén lesz ez a valószínűség 99%, azaz hány ember esetén esik legalább kettő születésnapja 99%-os valószínűséggel egy napra?

A megoldás az $1 - \frac{365\times 364\times\cdots\times(365-n+1)}{365^n} = 0,99$ egyenlet megoldása, ami kb. 55. Ez az eredmény azért tűnik paradoxnak, mert a valószínűség 1%-os csökkenéséből adódóan a szükséges emberek számának 366-ról 55-re történő csökkenését túl nagynak találjuk.

[5]V.ö. Wittgenstein (1956): „Nincs olyan különleges tárgy, amely a valószínűségi kijelentések sajátos tárgya lenne." (5.1511)

[6]Sőt, a tapasztalat szerint a relatív gyakoriságnak ez a stabilizálódása viszonylag gyorsan következik be: „Hallgatólagos feltevésünk, hogy a valószínűségelmélet néhány ismert alkalmazási területén (véletlen játékok,

A másik empirikus jegy a véletlenség, amelynek vizsgálata von Mises frekventista valószínűséginterpretációját minden más frekventista interpretációtól megkülönbözteti. A véletlent von Mises a következőképpen ragadja meg. Képzeljünk el egy játékost, aki a kaszinóban rulettezik, és hol a pirosra, hol a feketére tesz. A játékosnak különböző játékstratégiái lehetnek: pl. ha három egymást követő pirosat lát, akkor a következőben feketére tesz; vagy figyeli a szomszéd asztalt, és amikor ott fekete jön ki, akkor a következő körben itt is a feketére fogad; vagy egyszerűen minden prímszámadik menetben a feketére tesz[7] stb. Játékosunk azonban

> „előbb vagy utóbb arra a szomorú következtetésre jut, hogy egyik rendszer sem képes javítani esélyeit a hosszú távú nyerésben, azaz befolyásolni a relatív gyakoriságot, amellyel a különböző színek vagy számok megjelennek abban a sorozatban, amelyet a játékos a játék teljes sorozatából kiválasztott." (1964, 10. o.)

A relatív gyakoriság határértékének ezt a játékstratégiára való invarianciáját nevezi von Mises *a kizárt játékrendszer elvének (Prinzip vom ausgeschlossenen Spielsystem)* vagy *szabálytalansági axiómának.*[8] A tömegjelenségek véletlen jellege tehát a fentihez hasonló esélynövelő játékstratégiák nemlétét jelenti.

A tömegjelenségek két empirikus jegye tehát a relatív gyakoriság stabilizációja és a véletlenség. A két empirikus jeggyel rendelkező rendezett összességeket von Mises *kollektíváknak* nevezi. Az empirikus kollektívák matematikai reprezentációja végtelen sorozatokkal történik, amelyeket von Mises ugyancsak kollektíváknak nevez.

Milyen matematikai kollektívák reprezentálják helyesen az empirikus kollektívákat? A relatív határérték empirikus stabilizációját von Mises matematikailag úgy reprezentálja, hogy a valószínűség fogalmát kizárólag olyan sorozatok esetében tartja alkalmazhatónak, amelyekben a tulajdonságoknak végtelen határértékben van relatív gyakorisága. A másik empirikus jegy, a véletlenség reprezentációját von Mises a *helyszelekció* fogalmának bevezetésével kívánja megoldani. A helyszelekció egy eljárás, amelynek segítségével az eredeti sorozatunkból bizonyos szempontok alapján kiválasztunk egy részsorozatot. (A pontos definíciót lásd alább.) Egy sorozatban egy adott attribútum véletlenszerűen fordul elő (helyszelekciók egy adott halmazára nézve), ha relatív gyakorisága ugyanaz lesz a szelekcióval szűrt valamennyi részsorozatban, mint az eredeti sorozatban. A véletlenség tehát a szelekcióra való érzéketlenséget hivatott kifejezni. A véletlenségnek ez a reprezentációja fejezi ki azt a fizikai tényt, hogy fenti rulettjátékosunk nem képes úgy szelektálni a futamok között, hogy

fizika, biológia, biztosítás stb.) a frekvenciák (a különböző problémák esetében különböző mértékben) viszonylag gyorsan tartanak határértékeik felé ... Ennek a feltevésnek semmi köze a valószínűségi kalkulus axiómáihoz, és nem magyarázható semmilyen elméleti statisztika révén, hiszen az éppen ezen a feltevésen nyugszik." (1964, 108. o.)

[7] Ennek a stratégiának Doob szerint az az előnye, hogy a játékosnak a fogadások között egyre több ideje lesz gondolkodni a valószínűségelméletről.

[8] Amely elvet a termodinamika második főtételéhez, azaz *perpetuum mobile* nem létezésének elvéhez hasonló alapelvnek tart.

ezzel a piros-fekete arányt megváltoztassa. A rulett azért véletlen játék, mert a kimenetek aránya érzéketlen a fizikailag megengedhető helyszelekciókra nézve.

A matematikai kollektívák tehát aszimptotikus relatív gyakorisággal rendelkező véletlen sorozatok. Definíciójuk formálisan a következő:

17. Definíció. Legyen Σ valamilyen tulajdonságok (attribútumok) algebrája, és legyen $x : \mathbb{N} \to \Sigma$ egy sorozat. Egy x sorozatot *kollektívának* nevezünk, ha

(i) minden tulajdonságnak a sorozatban létezik aszimptotikus relatív gyakorisága, vagyis minden $a \in \Sigma$-ra létezik a (7.1)-ben definiált $r_x(a)$ érték;

(ii) és ez a határérték nem változik a megengedhető helyszelekciókra nézve, vagyis ha ϕ megengedhető helyszelekció, akkor

$$r_x(a) \quad = \quad r_{\phi(x)}(a)$$

minden $a \in \Sigma$ tulajdonságra, ahol $\phi(x)$ a helyszelekcióval szűrt részsorozat.[9]

Egy x kollektíva akkor relatív gyakoriság-modellje egy valószínűségi mértéktérnek, ha az 16. definíció értelmében a relatív gyakoriságok képesek reprodukálni a mértéktér valószínűségeit. Ilyenkor a kollektívában egy a tulajdonsághoz tartozó aszimptotikus relatív gyakoriságot az a tulajdonság $p_x(a)$ valószínűségének nevezzük.

A kollektíva fenti definíciójában nem mondtunk semmit arról, hogy mit tekintsünk megengedhető helyszelekciónak. Nyilvánvalóan a kollektíva fogalma érzékenyen függ a megengedhető helyszelekciók körétől (és a tulajdonságok algebrájától). Mielőtt erre a kérdésre térnénk, vizsgáljuk meg előbb azt, hogy hogyan képzeli el von Mises a kapcsolatot a véges empirikus kollektívák és ezek matematikai reprezentációjául használt végtelen kollektívák között. Von Mises kettős szóhasználata nem véletlen; megértéséhez fontos tudatosítani azt az empirikus attitűdöt, amely valószínűségelméletét – a szigorú axiomatizálására való törekvése mellett – megkülönböztette a végül uralomra jutó mértékelméleti megközelítéstől. Von Mises számára a valószínűségelmélet nem az absztrakt matematika egyik ága, hanem ugyanolyan empirikus tudomány, mint a mechanika vagy (sic!) a geometria:

---

[9]Megjegyezzük, hogy a kollektíva fenti definíciója valójában csak a $\Sigma = \mathcal{P}(\{0,1\})$ attribútumterekre áll. Bonyolultabb esetekben a véletlenszerűséget kifejező (ii) követelmény egy kicsit bonyolultabb: Legyen $a, b \in \Sigma$, és legyen x' az a sorozat, amelyet úgy kapunk x-ből, hogy elhagyjuk minden olyan elemét, amely nem tartozik vagy a-ba vagy b-be. Az így kapott x' részsorozatra alkalmazunk egy megengedhető ϕ helyszelekciót, majd a $\phi(x')$ szelektált részsorozatról megköveteljük, hogy

$$r_{\phi(x')}(a) \; : \; r_{\phi(x')}(b) \quad = \quad r_x(a) \; : \; r_x(b),$$

vagyis, hogy a és b relatív gyakorisága határértékének aránya egyezzen meg az eredeti x sorozatbeli arányukkal. Nyilvánvalóan, ha a és b egymás komplementerei – ahogy a $\mathcal{P}(\{0,1\})$ attribútumtérben a 0 és az 1 –, akkor a követelmény ekvivalens a fenti (ii) alatt megfogalmazott követelménnyel; vagyis a véletlen sorozat paradigmatikus esetéhez, a pénzdobáláshoz a fenti definíció elegendő. Mi a továbbiakban csak ilyen, $0-1$-sorozatokkal foglalkozunk.

> „Valószínűségelmélet alatt mi, akárcsak a mechanika vagy geometria alatt, megfigyelt mennyiségek bizonyos tartományának tudományos elméletét értjük. Ha megpróbáljuk leírni a tudományos kutatás ismert formáit, a következőket mondhatjuk: minden egzakt tudomány megfigyeléssel kezdődik, amely kezdetben hétköznapi nyelven nyer megfogalmazást, majd a pontatlan kifejezések egyre finomodnak, míg végül axiomatikus feltevésekkel helyettesítjük őket, amelyek egyúttal az alapfogalmakat is definiálják. Tautologikus, azaz matematikai transzformációk következnek ezután, amelyek a feltevésekből következményekhez vezetnek, amelyeket visszafordítva a köznyelvbe megfigyelésekkel ellenőrizhetünk operacionalista előírásoknak megfelelően.
>
> Így minden elegendőn fejlett matematikai tudományban van egy 'középső' rész, egy tautologikus vagy matematikai rész, amely matematikai levezetésekből áll. Manapság, a valószínűség tanulmányozásánál gyakori tendencia, hogy ezzel a matematikai résszel gondoson és szigorún foglalkozzanak, míg kevés figyelmet szentelnek magának a témának, a valószínűségelméletnek mint tudománynak.
>
> Ez tükröződik abban a tényben is, hogy a 'mértékelméleti megközelítést' általában előnyben részesítik a 'frekvencia-megközelítéshez' képest ... A valószínűségi kalkulusban használt matematikai eszközök leírása azonban csak egy része a történetnek. A tömegeloszlás, a sűrűségeloszlás, az elektromos töltés mindmind additív függvények. Ha a valószínűségben semmi sajátos nincsen, akkor miért definiáljuk valószínűségi eloszlások függetlenségét, és miért nem definiáljuk a tömegeloszlásokét? Miért használunk valószínűségi változót, konvolúciót, láncokat és a valószínűségi kalkulus egyéb sajátos fogalmait?" (1964, 43-44. o.)

Az utolsó két bekezdés utalásai Kolmogorov valószínűségszámítására vonatkozik. A valószínűségelmélet von Mises szerint tehát nem egy speciális mértékelmélet, hanem egy matematizált, ugyanakkor mégis empirikus tudomány. Ezért a nézetéért, amely összhangban állt tágabb empirista tudományfilozófiai nézeteivel – lévén a Berlini kör reprezentáns alakja –, gyakran érte vád:

> „Von Mises definíciója összekeveri az empirikus és teoretikus elemeket, amelyeket általában szétválasztunk a modern axiomatikus elméletekben. Ahhoz hasonlít ez, mintha a geometriai pontot krétapontok végtelenül csökkenő határértékeként definiálnánk."

– írja H. Cramér (1946) von Misesről, aki azután így válaszol Cramér vádjára:

> „Az 'empirikus és teoretikus elemek összekeverése' véleményünk szerint elkerülhetetlen egy matematikai tudományban. Amikor a rugalmasságtanban bevezetjük a feszültség fogalmát, nem szorítkozhatunk pusztán arra a megállapításra, hogy ez egy másodrendű szimmetrikus tenzor. Be kell még vezetnünk a kontinuummechanika alapfeltevéseit is, a Hooke-törvényt, stb., amelyek mindegyike

> empirikus és teoretikus elemek összekeverése. A rugalmasságtan nem tenzoranalízis ... a megfigyeléstől az elméleti fogalmakhoz való átmenet nem teljesen matematizálható. Nem logikai következtetés ez, inkább választás, amely, úgy hisszük, új megfigyelések mellett egyre bizonyosabb lesz." (1964, 45. o.)

A kollektíva, a feszültséghez hasonlóan egy empirikus tudomány elméleti fogalma. Az a tény, hogy a valószínűségelmélet ilyen elméleti fogalmakat használ, nem jelent nagyobb problémát, mint a 'végtelenül vékony vonal' fogalma a geometriában, vagy a sebesség fogalmának használata a mechanikában.

> „A valószínűség fogalmának a valószínűségelméletben ugyanaz a struktúrája, mint bármely olyan terület alapfogalmának, amelyen a valóság leírására és reprezentációjára matematikai analízist használunk. Tekintsük például a sebesség fogalmát a mechanikában. Amíg a sebesség csak mint az s elmozdulás és a t idő hányadosa mérhető, ahol mind s és t véges, nem nulla mennyiség, addig a sebesség a mechanikában mint ennek a hányadosnak a határértéke van definiálva amint $t \to 0$, vagy mint a $\frac{ds}{dt}$ differenciálhányados. Nincs értelme megkérdezni, hogy a differenciálhányados létezik-e a 'valóságban'. Matematikai létének feltevése a mozgás elméletének egyik alapja; igazolását abban leli, hogy segítségével képesek leszünk megfigyelhető mozgásokat leírni és megjósolni." (1964, 1-2. o.)

Hasonlóan, a relatív gyakoriság határértéke, akárcsak a sebesség, elméleti fogalom. Mindkettő bevezetését bizonyos hányadosok értékének stabilizációja indokolja a gyakorlatban. A relatív gyakoriság empirikus stabilizációja azonban még nem jelenti a határérték létezését. Jól ismert matematikai tény, hogy egy végtelen sorozat véges kezdőszeletének ismerete semmilyen információval nem szolgál a sorozat határértékének tekintetében; sőt létezése tekintetében sem. Vagyis egy tetszőlegesen nagy véges minta esetében sem tudjuk kiválasztani azt a kollektívát, amely a mintát hivatott reprezentálni. Véges számú megfigyelés alapján nem jelenthetjük ki, hogy a határérték létezik, és még ha létezik is, nem tudhatjuk, hogy mi ez az érték. A sebesség vagy a sűrűség esetében azonban ugyanez a helyzet, érvel von Mises, mégis ezek a fogalmak jó szolgálatot tesznek a gyakorlati predikciókban. Így a valószínűség relatív gyakoriság-értelmezése semmivel sem áll rosszabbul, mint bármely, a határérték fogalmára épülő elméleti fogalom.

> „A végtelen kollektíva fogalmára épülő elmélet eredményei olyan módon alkalmazhatók megfigyelések véges sorozatára, amely maga logikailag nem definiálható, azonban elegendően pontos a gyakorlatban. Az elmélet viszonya a megfigyeléshez ebben az esetben lényegében ugyanaz, mint az összes többi fizikai tudományban." (von Mises, 1928)

Az analógia azonban sántít. A nem valószínűségi tudományokban egy elmélet általában pontos jóslatokkal rendelkezik egy fizikai mennyiség elméleti értékét illetően. Ez az érték azután a megfelelő (önkényesen vagy a gyakorlat által motiváltan választott) hibahatáron belül összevethető lesz a mérési eredményekkel, így igazolható lesz, vagy legalábbis cáfolható. A valószínűségelmélet esetében azonban az elméleti jóslat és a tapasztalat viszonya bonyolultabb. Hogyan igazolható vagy cáfolható ugyanis egy valószínűségi kijelentés?

Egy tulajdonság valószínűsége nyilvánvalóan nem konfirmálható úgy, hogy egy tetszőleges hibahatárt választva megköveteljük, hogy a valószínűség a vizsgált esetek számának egy adott értékétől kezdve *sohase* térjen el a relatív gyakoriságtól. A relatív gyakoriság empirikus stabilitása nem ennyire szigorú törvény. De talán konfirmálható a nagy számok törvényére hivatkozva. Az valóban nem igaz – hangzik az érv –, hogy egy kollektívában egy véges mintán vett tulajdonság megegyezne a tulajdonság relatív gyakoriságával, azonban a kettő eltérése valószínűség erejéig rögzített. Vagyis megfelelő valószínűségi konfidencia-küszöböt választva, a valószínűségi jóslatok véges mintán ugyanúgy konfirmálhatók vagy falszifikálhatók, ahogyan a nem valószínűségi elmélet jóslatai a nem valószínűségi hibahatár megadása után. Az érv a nagy számok egyik-másik törvényére támaszkodik, amely szerint azonos eloszlású, független valószínűségi változókból képzett számtani átlag a változók számának növekedtével valószínűségi értelemben tart a várható értékhez.

Ez a kiút azonban, mégha működne is, el van zárva von Mises elől. A nagy számok törvénye egy adott valószínűségi változó iterálásából származó átlaggal van kapcsolatban. Ezt a valószínűségi változót interpretálják rendszerint úgy, mint egy szinguláris esemény valószínűségét, amelynek ismétlése adja a megfelelően hosszú sorozatok valószínűségét (a szorzatmértékben). Von Mises azonban épp ezt az interpretációt nem engedheti meg, hiszen számára az egyedi eseményeknek nincs valószínűsége, így nem is rendelhető hozzájuk valószínűségi változó.

Megtehetjük persze, hogy a nagy számok törvényeiben szereplő valószínűséget frekventistán értelmezzük, vagyis mint az illető kollektívában valamely attribútum határértékét. A konfirmáció kérdésére azonban ekkor sem kapunk megnyugtató választ. Ekkor ugyanis a nagy számok törvényeiben szereplő szorzatmértéket is frekventistán leszünk kénytelenek értelmezni, vagyis mint bizonyos tulajdonságú *sorozatok* relatív gyakoriságának határértékét *sorozatok sorozataiban*. Így értelmezve a szorzatmértéket a nagy számok erős törvénye arra a tautológiára zsugorodik, hogy adott aszimptotikus relatív gyakoriságú sorozatok sorozatában az adott aszimptotikus relatív gyakoriságú sorozatok relatív gyakorisága 1.[10] De a gyenge törvény sem alkalmasabb a feladatra; ez ugyanis azt mondja, hogy az olyan sorozatok sorozatainak relatív gyakorisága, amelyekben a sorozat kezdőszeletéből számított relatív gyakoriságok tetszőlegesen kicsit térnek az aszimptotikus relatív gyakoriságoktól, a kezdőszelet hosszával nullához tartanak. De mi köze ennek az állításnak az eredeti kérdésünkhöz, ti., hogy hogyan konfirmálható egy valószínűségi kijelentés? Úgy tűnik tehát, hogy von Mises frekvenciainterpretációja a kideríthetőség tekintetében hagy némi kétséget maga után. Most azonban térjünk át a frekvencia interpretáció megengedhetőségének kérdésére, vagyis arra a kérdésre, hogy a fenti kollektívák mennyire helyes matematika modelljei a

[10] A nagy számok erős törvényének ezt a kiüresedését tekintette W. Stegmüller (1973) a frekvenciainterpretáció legfőbb bajának.

valószínűségelméletnek.

Mindenekelőtt a kollektívák definíciójában szereplő helyszelekciók fogalmát szükséges tisztáznunk. A helyszelekció meghatározásánál Von Mises így fogalmaz: „A végtelen sorozatból egy végtelen részsorozatot úgy választunk ki, hogy a kiválasztandó elem indexénél nem használjuk a tulajdonságbeli különbségeket” (1928). A homályos fogalmazás[11] ellenére az intenció világos: egy tulajdonság relatív határértéke nyilván megváltozik, ha a részsorozatot úgy állítjuk elő, hogy a sorozatból egyszerűen kiválasztjuk az adott tulajdonságú elemeket. Az ilyen szelekciókat tehát zárjuk ki. Néhány példa megengedhető szelekcióra:

(i) Kiválasztani a sorozat minden prím indexű elemét.

(ii) Kiválasztani a sorozat egy elemét, ha előtte három adott tulajdonságú elem áll.

(iii) Kiválasztani a sorozat egy elemét, ha egy *másik kollektíva* azonos indexű eleme egy bizonyos tulajdonságú.

Az (i) és (ii) szelekciókat törvényszerű, a (iii) szelekciót véletlenszerű helyszelekciónak nevezzük. A helyszelekció pontos fogalmával kapcsolatos nehézségekre alább még visszatérünk, előbb azonban térjünk át a kollektívákkal végzett műveletekre. Von Mises (1964) joggal hangsúlyozza:

> „A valószínűségelméletnek nem feladata meghatározni bizonyos események valószínűségének numerikus értékét. Ezek a valószínűségek hasonló szerepet játszanak, mint a kezdőértékek (kezdőkoordináták, kezdősebességek) a mechanikában; ’adva’ vannak, és hogy miként, az nem érdekli az elméletet ... A valószínűségi kalkulus azt tanítja, hogy hogyan számítsuk ki adott kollektívák valószínűségi eloszlásából levezetett kollektívák eloszlását” (14. o.)

Új kollektívák ilyenfajta levezetésére régi kollektívákból három tapasztalatilag motivált művelet áll rendelkezésre: a *keverés*, a *felosztás* és az *összekapcsolás*. A keverés a valószínűségek összeadásához, a felosztás a feltételes valószínűséghez, az összekapcsolás valószínűségek szorzásához szükséges. Ezek a műveletek valamennyien támaszkodnak a helyszelekció fogalmára, de nem függnek attól, hogy mely helyszelekciókat tartunk megengedhetőnek.

Keverés. Legyen Σ és Θ attribútumok két algebrája, és legyen X és Y a (halmaz)algebrák alaphalmaza. Legyen $f : X \to Y$ egy tetszőleges mérhető leképezés, és tegyük fel, hogy f-nek létezik inverze, ami szintén mérhető; vagyis értelmezhető az $f[\ ] : \Sigma \to \Theta$ leképezés az attribútumok között. Legyen x egy kollektíva. Ekkor az f függvény indukálta $y \equiv f[\ ] \circ x : \mathbb{N} \to \Theta$ sorozatról könnyű látni, hogy kollektíva lesz (bármit értsünk is

[11] A megfogalmazás Church, Wald stb. eredményei nyomán (ld. alább) a *Mathematical Theory of Probability and Statistics*ben már jóval világosabb: „Azt mondjuk, hogy egy részsorozatot helyszelekcióval kaptunk, ha a döntés, hogy az n. elemet kiválasszuk-e vagy sem, csak az n számtól függ valamint a megelőző n-1 elemen, x_1, x_2, $\ldots x_{n-1}$-en, és nem függ az n. elemtől vagy valamelyik rákövetkezőtől.” (1964, 9. o.)

megengedhető helyszelekció alatt), valamint, hogy a $p_x \circ f^{-1}$ kompozíció egy p_y valószínűséget generál az y kollektíván. A „keverés" elnevezés azt fejezi ki, hogy az új kollektívában a régi kollektíva attribútumai (az f függvény által) össze vannak keverve. Ha például $x : \mathbb{N} \to \mathcal{P}(\{1,2,3,4,5,6\})$ a szimmetrikus kockával történő dobásokat leíró kollektíva, ahol $p_x(\{i\}) = \frac{1}{6}$ minden i-re, valamint f a páros számokhoz 0-t, a páratlan számokhoz 1-et rendelő függvény, akkor $p_y(\{0\}) = p_y(\{1\}) = \frac{1}{2}$ lesz, vagyis a régi kollektíva valószínűségei egyszerűen összeadódnak. Ugyanakkor a 0 és 1 előfordulás az új kollektívában is véletlen lesz, amennyiben az a régiben az $1 \dots 6$ előfordulása az volt.

Felosztás. Legyen x egy kollektíva, és legyen $a \in \Sigma$ egy attribútum, amely végtelen sokszor fordul elő x-ben, és $p_x(a) \neq 0$. Az a attribútum segítségével az y kollektívát úgy definiáljuk, hogy az x kollektívából elhagyjuk azokat az elemeket, amelyekre a nem áll fenn. Ekkor az új kollektíván kapott p_y valószínűség segítségével a $p_x(\cdot|a) \equiv p_y(\cdot)$ módon definiálható a feltételes valószínűség a régi kollektíván. A szimmetrikus kocka példájánál maradva, és az a attribútumnak a párosságot választva, $p_x(\{i\}|a) = \frac{1}{3}$ lesz, ha i páros, és $p_x(\{i\}|a) = 0$ lesz, ha i páratlan.

Összekapcsolás. Legyen x és y két kollektíva az attribútumok Σ illetve Θ algebráján, és tekintsük a kollektívák $z_k \equiv (x_k, y_k)$ elempárjaiból képzett z sorozatot. A z sorozat nem minden esetben lesz kollektíva. Amennyiben azonban x és y *független* kollektívák, akkor z is kollektíva lesz a $\Sigma \times \Theta$ algebrán. Kollektívák függetlenségét a következő módon definiáljuk. Egy x és y kollektíva *független*, ha az alábbi három lépésben definiált x'' sorozat kollektíva lesz, valamint $p_{x''} = p_x$.

(i) Megengedett helyszelekcióval válasszuk ki y egy y' részsorozatát.

(ii) Válasszunk egy tetszőleges $a \in \Theta$ attribútumot, és képezzük x-nek egy x' részsorozatát úgy, hogy x-nek csak azon k indexű elemeit tartjuk meg, amelyekre $y'_k \in a$.

(iii) Megengedett helyszelekcióval válasszuk ki x' egy x'' részsorozatát.

Nem nehéz ellenőrizni, hogy a definíció szimmetrikus.

A függetlenségnek ez a definíciója jelentősen különbözik attól, ahogy Kolmogorov valószínűségelméletben a függetlenséget kezelik. Ott például egy szabályos érmével egymást követő két dobás illetve a két szabályos érmével való dobás esetéhez egyaránt a szorzatmértéket *posztuláljuk*, ha a dobásokat fizikailag függetlennek tartjuk. Von Misesnél ezzel szemben a szabályos érmével egymást követő két dobás esetében a függetlenség és így a szorzatszabály – a kollektíva, a helyszelekció és a függetlenség fogalma alapján – *bizonyítható.*

Ennek a bizonyításnak a megfordítottja volt az, ami miatt von Mises a kollektívákat *szükségszerűnek* tekintette a valószínűség modellezéséhez: ha egy sorozatban az egymást követő dobások a szorzatszabálynak megfelelően viselkednek, akkor a sorozat invariáns arra a helyszelekcióra nézve, amely a szorzatszabály bizonyításában megjelenik, vagyis a szorzatszabályhoz kollektívák kellenek.

Térjünk most vissza ahhoz a kérdéshez, hogy a kollektíva definíciójában mit jelent a megengedhető helyszelekció. Az első próbálkozások a megengedhető helyszelekció körvonalazására Copelandtől (1932), Reichenbachtól (1932) és Poppertől (1935) származnak, akik egymástól függetlenül érkeztek el az ún. *Bernoulli-szelekció* fogalmához. Informálisan a Bernoulli-szelekció (0-1 sorozatokra) a következőt jelenti. Legyen x egy végtelen 0-1-sorozat. Vegyünk egy véges hosszú 0-kból és 1-esekből álló l hosszúságú 'szót', pl. a 01101-et, és vizsgáljuk meg, hogy a választott szó hol fordul elő x-ben. Ha a szó előfordul a sorozat $x_k, x_{k+1}, \ldots x_{k+l}$ szakaszán, akkor válasszuk ki a sorozat x_{k+l+1}-ik elemét. Ha x végtelenszer tartalmazza a választott szót, akkor a fenti szelekció egy végtelen részsorozatot ad. Az olyan sorozatokat, amelyekben a 0 és 1 relatív gyakoriságának határértéke invariáns minden Bernoulli-szelekcióval szemben *Bernoulli-sorozatnak* nevezzük. A kollektívákat kezdetben a fenti szerzők ezekkel a *Bernoulli-sorozatokkal* azonosították.

A Bernoulli-sorozatok speciális esetei a *normális számok*: azok a Bernoulli-sorozatok, amelyekben az 1-ek és 0-ak valószínűsége a sorozatban $\frac{1}{2}$. A normális számokra vonatkozóan D. G. Champernowne (1933) eredménye fontosnak bizonyult a kollektívákra nézve is. Champernowne ugyanis megmutatta, hogy létezik megkonstruálható normális szám: pl. az $x = 0100011011000\ldots$, amely bináris számok lexikografikus rendezéséből adódik. Ez a számelméleti eredmény közvetve arra is rámutatott, hogy a Bernoulli-sorozatok nem tölthetik be azt a szerepet, amelyet von Mises a kollektíváknak szánt, mivel a konstruálható Bernoulli-sorozatokban maga a konstrukció nyújt lehetőséget olyan helyszelekcióra, amelyre nézve a relatív gyakoriságok nem invariánsak.

A Bernoulli-sorozatok általánosításaként a helyszelekció fogalmára végül a következő meghatározás született. Legyen $x : \mathbb{N} \to \{0,1\}$ egy végtelen sorozat, és $f : \mathbb{N} \to \mathbb{R}$ egy tetszőleges függvény. Az indexek k sorozatából válasszuk ki azt a k_i végtelen részsorozatot (ha van ilyen), amelyre teljesül, hogy $c_k = 1$, ahol $c_k = f(b_k)$, $b_{k+1} = 2b_k + x_k$ és $b_1 = 1$. Ekkor a helyszelekciót a $(\phi x)_k \equiv x_{k_i}$ egyenlettel definiáljuk. A konstrukció mögött a gondolat nyilvánvaló: a megengedhető szelekciók azok, amelyek csak az elem indexénél kisebb indexű elemektől függenek. Ezt a függést a megelőző tagoktól biztosítják a b_k elemek a konstrukcióban.[12]

A definíció azonban ebben a formában, ahogy azt E. Kamke (1932) megmutatta, ellentmondásos. Legyen ugyanis $f(k) = x_{l(k)}$, ahol $l(k)$ a legkisebb pozitív egész, amelyre $2^{l(k)} > k$. Ekkor az x_{k_i} sorozat épp az x_k sorozat 1-es elemeit fogja tartalmazni, vagyis az 1-esek relatív gyakorisága a részsorozatban 1 lesz, bármennyi volt is az eredeti sorozatban. Ha tehát megengedjük a fenti f függvénnyel generált helyszelekciót, akkor kollektívák nem léteznek. Von Mises éppen ettől tartott.

[12] A helyszelekciónak a fentiekkel ekvivalens megfogalmazása a követő: legyen adva 0-1-értékű függvények egy f_1, $f_2(x_1)$, $f_3(x_1, x_2)$, $\ldots f_{k+1}(x_1, x_2 \ldots x_k) \ldots$ végtelen sorozata, ahol az f_{k+1} függvény azt reprezentálja, hogy az x sorozat első $x_1, x_2 \ldots x_k$ elemének függvényében az x_{k+1}-ik elemet kiválasszuk-e vagy sem: ha kiválasztjuk, akkor legyen $f_{k+1} = 1$, ha nem, akkor legyen $f_{k+1} = 0$. A függvénysorozattal x-ből kiválasztott elemek egy $\{x_{k_i}\}$ részsorozatot határoznak meg; a helyszelekciót ismét csak a $(\phi x)_k \equiv x_{k_i}$ egyenlettel definiáljuk.

A fenti definíciót Church (1940) pontosította. A fenti meghatározásban szereplő f függvényekre kikötötte, hogy azok legyenek rekurzívak. Ezzel a kikötéssel nemcsak Kamke ellenvetése vált elkerülhetővé, hanem von Mises azon félelmére is – miszerint kollektívák esetleg egyáltalán nem is léteznek – megnyugtató válasz érkezett. A kollektívák létezése ugyanis nem magától értetődő. Egy végtelen sorozat tudniillik mindig valamilyen rekurzióval adható meg. Az a rekurzió azonban, amellyel egy ilyen sorozatot megadunk, egyben arra is alkalmas, hogy a sorozat elemei között szelektáljunk: vagyis a sorozat nem lesz véletlen. Konstruktíve nem lehet tehát megadni olyan véletlen sorozatot, amely minden helyszelekcióval szemben véletlen lenne.[13] A rekurzív függvények azonban megszámlálhatóan sokan vannak, és így Wald (1937) tétele alapján a helyszelekciók fenti megszámlálható halmazára nézve kontinuum sok véletlen kollektíva létezik.[14]

Az öröm azonban nem tartott sokáig. J. Ville (1939) megmutatta, hogy helyszelekciók bármely megszámlálható halmazához létezik olyan bináris kollektíva, amelyben az 1-esek aszimptotikus frekvenciája p, azonban a kollektíva minden véges kezdőszeletében az 1-esek relatív gyakorisága – véges kivételtől eltekintve – nagyobb vagy egyenlő p-nél. Röviden, a relatív gyakoriság felülről tart határértékéhez. Ez a tulajdonság a (később Khinchin révén bizonyított) iterált logaritmus törvénye fényében meglehetősen atipikus viselkedésnek számított. Ez a törvény ugyanis a relatív frekvenciák oszcillációjára ad küszöböt; ezzel szemben a frekventista elmélet ezekre az oszcillációkra semmilyen korlátot nem ad. Mindezek azt mutatták, hogy von Mises valószínűségelmélete nem azonos az éppen kialakulóban levő mértékelméleti megközelítéssel. Az ellenvetésekre von Mises lakonikusan válaszolt: „Elfogadom a tételt, de nem látom az ellenvetést". Valóban, von Mises elmélete számára nem jelentett a priori előírást, hogy izomorf legyen a kolmogorovi elmélettel.

A frekventista valószínűségelméletre a legsúlyosabb csapást a martingálok megjelenése mérte. A martingálok segítségével Ville megmutatta, hogy a kizárt játékrendszer von Mises-i elvét a kollektívák nem ragadják meg jól, mivel Ville fenti kollektíváihoz lehetséges olyan

---

[13] Reichenbach éppen ezért adta fel a valószínűség definíciójában a véletlenség követelményét. Egy további tipikus korabeli reakció von Mises definíciójának értelmességével szemben E. Tornieré (1933): „Nem hiszem, hogy a próbálkozás, hogy von Mises elméletét tiszta matematikai formába öntsük, sikeresen keresztülvihető, és azt sem hiszem, hogy az ilyen próbálkozásoknak hasznuk volna. Itt nyilvánvalóan azzal a fölöttébb érdekes jelenséggel van dolgunk, hogy egy gyakorlati és teljesen értelmes fogalom – kiválasztás a tulajdonságbeli különbségekre való tekintet nélkül – elvileg zár ki mindenféle tisztán matematikai és axiomatikus megragadást. Mindazonáltal kívánatos volna, hogy ez a kérdés, amely talán alapvető jelentőségű, matematikusok újabb köreinek figyelmét is magára vonja." (320. o.)

[14] Pontosabban a következő igaz: Legyen (Ω, Σ, p) valószínűségi mértéktér, és legyen $K(\Gamma, \Sigma)$ a Σ attribútumtéren definiált, helyszelekciók egy Γ halmazára nézve invariáns kollektívák halmaza, amely kollektívákra igaz továbbá, hogy tetszőleges $a \in \Sigma$ elemre $p(a) = r_x(a)$; röviden, legyen $K(\Gamma, \Sigma)$ a (Ω, Σ, p) mértékteret a 16. definíció értelmében modellező kollektívák halmaza. Ekkor Wald tétele a következőt állítja: Ha Ω és Γ megszámlálható, akkor $K(\Gamma, \Sigma)$ nem megszámlálható.

Még pontosabban szólva, Wald a mértéktértől azt követeli meg, hogy legyen Σ Jordan-mérhető, p pedig Jordan-mérték. A Jordan-mérték választását azzal indokolja, hogy mivel a kollektívák által indukált valószínűségi mérték nem σ-additív, ezért a nem Jordan-mérhető halmazok mértékének, amely halmazok nem közelíthetők véges módon, nincs frekvenciainterpretációja. (Ld. Rédei 2001) stratégiákat gyártani, amely végtelen nyereményhez vezet. Ugyanakkor a kollektívák a von Mises-i értelemben éppen az ilyen nyerő stratégiák létét hivatottak kizárni.

Mint azt sokan hangsúlyozzák, von Mises és Kolmogorov valószínűségelmélete közötti fenti eltérések mindegyike arra a különbségre vezethető vissza, hogy amíg Kolmogorov mértékelmélete σ-additív, addig a frekventista elmélet nem az. Amikor azonban a bizonyításra kerül a sor, rendszerint ennél kevesebbet bizonyítanak. Az alábbiakban mutatunk négy olyan állítást, amely állításokra a szakirodalomban úgy szokás tekinteni, mint amelyek azt *bizonyítják*, hogy a valószínűség von Mises-féle relatív gyakoriság-interpretációja nem megengedhető interpretáció. Amint azonban alább megmutatjuk, valójában egyik állítás sem érinti a von Mises frekvenciainterpretációját, sőt mi több, még általában a frekvenciainterpretációkat sem. Az állítások a következők:

(i) A frekvenciák nem alkotnak σ-additív mértéket a sorozatok terén.

(ii) Az aszimptotikus relatív gyakorisággal rendelkező sorozatok nem alkotnak σ-algebrát.

(iii) Az aszimptotikus relatív gyakorisággal rendelkező sorozatok nem alkotnak algebrát sem.

(iv) A véletlen sorozatok nem alkotnak algebrát.

Nézzük az állításokat sorban. Hogy a frekvenciák nem alkotnak σ-additív mértéket a sorozatok terén, arra ilyen példákat szokás felhozni. Tekintsük a természetes számok $x_k \equiv k$ sorozát. Ebben a sorozatban bármely szám (elemi attribútum) relatív gyakoriságának határértéke 0, ugyanakkor az $\mathbb{N} = \{1\} \cup \{2\} \cup \ldots$ megszámlálható unióra

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} 1_{\mathbb{N}}(x_k) \quad = \quad 1,$$

vagyis a relatív gyakoriságok nem alkotnak σ-additív mértéket.

Ha már a σ-additivitásnál tartunk, hadd említsük meg azt az érvet, amely a szubjektivista de Finettit a σ-additivitás feladására késztették. Amint azt a valószínűség formális tárgyalásánál láttuk, a valószínűséget egy adott Ω alaphalmaznak nem tetszőleges részhalmazára definiáltuk, vagyis nem a teljes $\mathcal{P}(\Omega)$ hatványhalmazra, hanem annak valamely halmazelméletileg zárt részhalmazára. Az ok matematikai jellegű: nem megszámlálható Ω hatványhalmazán nem értelmezhető σ-additív (valószínűségi) mérték. Az additivitás σ-kiterjesztésének a valószínűség értelmezési tartományának a zsugorodása az ára. Az egyik legismertebb példa a $[0,1]$ nem Lebesgue-mérhető halmazai. Definiáljuk $[0,1]$-en az alábbi ekvivalenciarelációt: $x \sim y$ akkor és csak akkor, ha $x - y \in \mathbb{Q}$. Az ekvivalenciareláció az intervallumot $[x] \equiv \{x + r \in [0,1] \mid r \in \mathbb{Q}\}$ alakú ekvivalenciaosztályokra osztja. Legyen E egy olyan halmaz, amelyben minden ekvivalenciaosztályból pontosan egy elemet tartalmaz. (Figyelem: ennél a lépésnél felhasználjuk a kiválasztási axiómát!) Definiáljuk továbbá minden $r \in \mathbb{Q}$ számra az E_r halmazt úgy, hogy E minden eleméhez hozzáadunk

r-et (modulo 1). Az E_r halmazok nyilván megszámlálhatóan sokan vannak, páronként diszjunktak, és uniójuk a teljes $[0,1]$ intervallum. Most tegyük fel, hogy valamely E_r halmaz valószínűségi mértéke p. A Lebesgue-mérték eltolásinvarianciájából adódóan ekkor a többi E_r halmaz valószínűségi mértéke is p lesz. Ha $p=0$, akkor $\sum_r p(E_r)=0$, ha $p \neq 0$, akkor $\sum_r p(E_r)=\infty$. Feltételezve a σ-additivitást azonban $\sum_r p(E_r)=p(\vee_r E_r)=p([0,1])=1$, ami ellentmondás. Vagyis az E halmaz nem Lebesgue-mérhető.

A formalizmus tehát választásra kényszerít: vagy ragaszkodunk a σ-additivitáshoz, de akkor le kell mondanunk arról, hogy minden részhalmazhoz valószínűséget rendeljünk, vagy ragaszkodunk a valószínűség lehető legtágabb értelmezési tartományához, de akkor fel kell adnunk a σ-additivitást. A *main stream* matematikai okokból az előbbi, de Finetti filozófia okokból az utóbbi utat választotta. De Finetti szerint ugyanis semmilyen a priori indokunk nincsen bizonyos eseményeket kizárni a valószínűség logikailag megengedett értelmezési tartományából, vagyis például a fenti E halmazt a priori nem gondolhatjuk kevésbe alkalmasnak arra, hogy valós eseményeket reprezentáljon, mint a $[0,1]$ intervallum bármely más halmazát. Így viszont kénytelenek leszünk beérni σ-additivitás helyett a véges additivitással.

Most azonban térjünk vissza a fenti (ii) ellenpéldához, vagyis ahhoz, hogy az aszimptotikus relatív gyakorisággal rendelkező sorozatok nem alkotnak σ-algebrát. Legyen ugyanis $x:\mathbb{N}\to\Sigma$ tetszőleges sorozat, amelyben egy a attribútum relatív gyakoriságának nincs határértéke. Most legyen $x^{(n)}$ az a sorozat, amelyben az n. elem x_n, a többi elem pedig $\emptyset$. Minden $x^{(n)}$ sorozatban az a tulajdonság relatív gyakoriságának határértéke 0, ugyanakkor az $x^{(n)}$ kollektívák megszámlálható uniójaként előálló x sorozatban az a tulajdonság relatív gyakoriságának nincs határértéke.

A (iii) ellenpélda azt mutatja meg, hogy nemcsak a megszámlálható unióval van a baj: az aszimptotikus relatív gyakorisággal rendelkező sorozatok nem alkotnak algebrát sem. Tekintsük az

$$x \quad = \quad 110011110000111111111111000000000000\ldots$$

sorozatot, amelyet azzal az egyszerű konstrukcióval hozunk létre, hogy a sorozatot két 1-el kezdve addig írunk 0-kat, amíg az 1-ek relatív gyakorisága le nem csökken $\frac{1}{2}$-re, azután pedig addig írunk 1-ket, amíg az 1-ek relatív gyakorisága fel nem nő $\frac{3}{4}$-re; és ezt folytatjuk a végtelenségig. Nyilvánvalóan a sorozatban az 1-esek határértéke a konstrukció miatt nem létezik. Most tekintsük az alábbi négy sorozatot:

$$\begin{aligned}
y &= 101010101010101010101010101010101010\ldots\\
z &= 100110100101101010101010010101010101\ldots\\
y' &= 011001011010010101010101101010101010\ldots\\
z' &= 010101010101010101010101010101010101\ldots
\end{aligned}$$

A sorozatokat úgy hoztuk létre x-ből, hogy mindegyik megegyezik x-nek az egyeseket és nullákat tartalmazó részeinek páros- illetve páratlan sorszámú elemével. Konkrétan:

$$y_k = x_k \text{ akkor és csak akkor, ha } k \text{ páratlan és } x_k = 1\text{; vagy ha } k \text{ páros és } x_k = 0,$$
$$z_k = x_k \text{ akkor és csak akkor, ha } k \text{ páratlan és } x_k = 1\text{; vagy ha } k \text{ páratlan és } x_k = 0,$$
$$y'_k = x_k \text{ akkor és csak akkor, ha } k \text{ páros és } x_k = 1\text{; vagy ha } k \text{ páros és } x_k = 0,$$
$$z'_k = x_k \text{ akkor és csak akkor, ha } k \text{ páros és } x_k = 1\text{; vagy ha } k \text{ páratlan és } x_k = 0.$$

Nyilvánvaló, hogy az egyesek relatív gyakorisága mind a négy sorozatban $\frac{1}{2}$, továbbá az is, hogy

$$x = (y \wedge z) \vee (y' \wedge z'),$$

vagyis a relatív gyakoriság határértéke nem zárt a sorozatokon végzett véges algebrai műveletekre sem.

Végül a (iv) ellenpélda szerint a véletlen sorozatok sem alkotnak algebrát a határérték létezésétől függetlenül. Vegyünk ugyanis egy tetszőleges véletlen 0-1-sorozatot, és cseréljük meg a 0-kat és az 1-ket. Az így kapott sorozat nyilván véletlen lesz, bárhogyan is definiáltuk az eredeti sorozat véletlenségét. A két sorozat (pontonkénti) uniója azonban nem lesz véletlen.[15]

Mint említettük a fenti matematikai nehézségeket a szakirodalomban úgy szokás értelmezni, mint amelyek azt bizonyítják, hogy a valószínűség von Mises-féle relatív gyakoriság-interpretációja nem megengedhető interpretáció. Ez a vélekedés azonban több szempontból is téves: (1) Először is rögtön látnunk kell, hogy az első három ellenpélda egyáltalán nem vonatkozik von Misesre, mivel von Mises a sorozatoktól megköveteli a véletlenséget is. Röviden, az (i)-(iii) sorozatok nem kollektívák.[16] (2) Az a tény, hogy létezik olyan sorozat, amelyben a relatív gyakoriságok nem σ-additívek, még nem jelenti azt, hogy a valószínűségnek ne lehetne relatív gyakoriság-modelljét adni. A modellezés a 16. definíció értelmében nem kíván izomorfiát a két struktúra között, csupán annyit, hogy minden valószínűségi mértéktérnek létezzen relatív gyakoriság-modellje. (3) Ami a (iv) ellenvetést illeti, ismét csak az 16. definícióra kell utalnunk: az, hogy léteznek olyan véletlen sorozatok, amelyek véletlensége nem zárt az algebrai műveletekre, szintén nem jelent ellenpéldát von Mises frekvenciainterpretációjával szemben, hiszen a sorozatok *közötti* műveletek egyáltalán nem jelennek meg a modellben; egy algebra valószínűségi mértékeinek *egy sorozaton belül* kell reprezentálódniuk. (4) Ugyanez a helyzet a (ii) és (iii) esetben is: a sorozatok *közötti* műveleteknek nem kell sem zártaknak, sem σ-zártaknak lenniük. (5) Végül a legfontosabb: az a tény, hogy a frekvenciákra alapozott modell nem izomorf a mértékelméleti modellel, nem jelenti azt, hogy a relatív gyakoriság-modell ne lehetne a valószínűség helyes fizikai

[15] Gyenis Zalán példája.

[16] Hogy az első három állításra adható-e olyan példa is, amelyben a sorozatok kollektívák, vagyis amely sorozatok véletlenek, az – legalábbis a szerző számára – nyitott kérdés.

elmélete, hiszen semmilyen a priori garanciánk nincs arra vonatkozóan, hogy a valószínűség fogalmát a kolmogorovi axiómák ragadják meg helyesen. A mértékelméleti leírás mellett csak matematikai érvek szólnak: az elmúlt nyolcvan évben a kolmogorovi axiómarendszer matematikailag gyümölcsözőnek bizonyult. Hogy melyik modell interpretálható *fizikailag*, az egy ettől független kérdés.

A történeti tény azonban mégis az, hogy a fenti és a hasonló kritikák hatására von Mises elképzeléseit a valószínűség mértékelméleti megfogalmazása végleg háttérbe szorította. A valószínűségszámítás alapjai tisztázásának szentelt 1937-es genfi konferencián – amelyen von Mises maga nem volt jelen – Fréchet csokorba szedte a frekventista megközelítés hátrányait, és ezzel a szakma végleg elfordult von Misestől.

7.2. A „gyakorlatilag biztos" bekövetkezés

Mint azt korábban említettük, a valószínűség formális tárgyalását a mértékelmélet fejlődése tette lehetővé. Az elméletet végleges formába Kolmogorov öntötte a *Grundbegriffe der Wahrscheinlichkeitsrechnung* (1933) című klasszikus munkájában. A művet Kolmogorovnak a valószínűségi konvergenciák és a véletlen folyamatok témakörében kollégájával, Khinchinnel folytatott bő egy évtizedes kutatásai előzték meg. A *Grundbegriffe* a hilberti program jegyében a kibontakozó elmélet axiomatikus felépítésének igényével íródott. A rövid könyvecskéből, amely kitűnő didaktikájával a matematika-tankönyvek iskolapéldája, itt bennünket most csak azok a részek érdekelnek, amelyek a valószínűség és a tapasztalat kapcsolatát tárgyalják. Ezekből a részekből ugyanis kiviláglik, hogy Kolmogorov, kevésbé radikális formában ugyan, mint von Mises, de szintén a relatív gyakoriság-interpretáció híve.

A két elképzelés közötti különbségre Kolmogorov rögtön a bevezetésben utal:

> „A valószínűségszámításnak léteznek más axiomatikus felépítései is, méghozzá éppen olyanok, amelyekben a valószínűség nem alapfogalom, hanem más fogalmakkal kifejezve szerepel.[17] Ekkor azonban másra törekszenek, nevezetesen arra, hogy a valószínűségszámítás tudományát a lehető legszorosabban összekapcsolják a valószínűségfogalom tapasztalati eredetével." (1933/1982, 12. o.)

Kolmogorov óvatosabb elmélet és tapasztalat viszonyát illetően. A kettő viszonyát tisztázó rész a könyv második paragrafusában rögtön az elmélet véges részének axiomatikája után következik.[18]

> „A valószínűségszámítás alkalmazása a tapasztalati valóságra a következő séma szerint történik.

[17] És itt a lábjegyzet von Misesre utal.

[18] Mielőtt Kolmogorov nekifog, a lábjegyzetben még egyszer hivatkozik von Misesre : „A valószínűségszámítás valódi események körére való alkalmazhatóságához szükséges előfeltevések kifejtésében nagymértékben építünk Mises következtetéseire." (14. o.)

1. Felteszik, hogy adott valamilyen korlátlan számú ismétlést megengedő $\mathfrak{S}$ feltételegyüttes.
2. Kiindulnak eseményeknek egy meghatározott köréből, amelyek felléphetnek a $\mathfrak{S}$ feltételek megvalósulása következtében. Az egyes esetekben ezek az események különböző kombinációkban következhetnek, illetve nem következhetnek be. Az Ω halmaz magában foglalja a figyelembe vett események bekövetkezésének és be nem következésének minden lehetséges variációját.
3. Ha a gyakorlatban megvalósuló $\mathfrak{S}$ feltételek realizálódása után egy variáns a (valamilyen feltételekkel meghatározott) A halmazhoz tartozónak bizonyul, akkor azt mondják, hogy bekövetkezik az A esemény.
4. Bizonyos feltételek mellett, amelyeket itt közelebbről nem ismertetünk, feltehetjük, hogy valamely A eseménynek – amelyek a $\mathfrak{S}$ feltételek megvalósulásokat lehet, hogy fellépnek, lehet, hogy nem – megfeleltetnek egy, a következő tulajdonságokkal rendelkező meghatározott $P(A)$ valós számot.

   A. Gyakorlatilag biztosak lehetünk benne, hogy ha a $\mathfrak{S}$ feltételegyüttest nagyszámú n alkalommal megismétlik, és m jelöli azoknak az eseteknek a számát, amelyekben az A eseménye végbement, akkor az $\frac{m}{n}$ hányados kicsit tér el $P(A)$-tól.

   B. Ha $P(A)$ nagyon kicsi, akkor gyakorlatilag biztos, hogy a $\mathfrak{S}$ feltételek egyszeri megvalósításakor az A esemény nem fog bekövetkezni.” (14-15. o.)

Az 1. pont világossá teszi, hogy Kolmogorov a valószínűség fogalmát ismételhető eseményekre, tehát eseménytípusra és nem szinguláris eseményekre alkalmazza. A 2. pont ennek az ismétlődésnek a jellegét körvonalazza. Fontos látni, hogy Kolmogorov – von Misessel ellentétben – nem várja el az eseményektől, hogy azok véletlenszerűen történjenek. Az eltérő attitűdök mögött a két szerző eltérő természetfilozófiai álláspontja áll a véletlen tekintetében. Von Mises meggyőződése szerint a fizika modern fejleményei azt mutatják, hogy a természet leírására nem adható ’algoritmus’, vagyis a véletlenség intrinszikus tulajdonsága az eseményeknek. Ezzel szemben Kolmogorov a valószínűségi kalkulust egyszerű modellnek tartotta, amelynek alkalmazhatósága nem függ a természet véletlenszerű vagy éppen determinisztikus viselkedésétől, hanem csak a rögzített $\mathfrak{S}$ feltételegyüttestől. A 3. feltétel pusztán a terminológiát rögzíti; a valószínűség tapasztalati alkalmazása a 4. pontban történik. A „közelebbről nem ismertetett” feltételek von Plato (1994) szerint a nagy számok törvényeinek legfontosabb feltétele, vagyis a valószínűségi függetlenség tapasztalati alkalmazhatóságára vonatkoznak. A kérdésre Kolmogorov később a függetlenség fogalmának szentelt fejezetben explicite is visszatér:

> „Ennek megfelelően a természettudományos gondolkodás előtt álló egyik legfontosabb feladat az, hogy miután a valószínűség fogalmának a lényegét mint

> kulcskérdést tisztázta, kiderítse és pontosítsa, milyen előfeltevések mellett tekinthetünk adott tapasztalati jelenségeket függetlennek." (22. o.)

Ha az ismétlődő események valószínűségi értelemben függetlennek tekinthetők, akkor (egyéb járulékos feltételek mellett) alkalmazhatóak rájuk a nagy számok különféle törvényei. A nagy számok törvényeit, különösen a Bernoulli-tételnek végtelen szorzatmértéktéren értelmezett karakterisztikus függvényekre kimondott D.5 alakját (ld. 172. oldal) szokásos a következőképpen értelmezni. A karakterisztikus függvények azt mutatják meg, hogy adott események egy végtelen kísérletsorozat adott futamában bekövetkeznek-e vagy sem, ennélfogva a karakterisztikus függvények aritmetikai átlaga a szóban forgó események relatív gyakoriságát modellezi. A Bernoulli-tétel ezek után azt mutatja meg, hogy ez a relatív gyakoriság (a szorzatmérték szerinti) valószínűségi értelemben konvergál az adott esemény valószínűségéhez. Fontos látni, hogy ez a konvergencia valószínűségi konvergencia: a Bernoulli-tétel esetében a valószínűségi változók valószínűségi értelemben konvergálnak a közös várható értékhez, a nagy számok erős törvényeiben pedig „majdnem biztos" értelemben. A valószínűség így nem redukálható maradéktalanul a frekvenciára, mivel a Bernoulli-tétel kiküszöbölhetetlenül tartalmaz egy „másodrendű" valószínűséget is: szorzatvalószínűséget, amelyben a konvergencia meg van fogalmazva.

Ez a valószínűség azonban határértékben 1-hez tart, sőt az erős törvényekben 1. És éppen ez az a tény, amelyben a frekvenciainterpretáció igyekszik megkapaszkodni. Ha a valószínűségtől ebben a speciális esetben, vagyis akkor, amikor értéke közel 1, sikerülne valamiként megszabadulni, akkor – hangzik az érv – az út nyitva állna a valószínűség és a relatív gyakoriság azonosításához a nagy számok törvényein keresztül. Mit is jelent az, hogy egy esemény valószínűsége közel 1; vagy a komplementer esetben, hogy egy esemény valószínűsége közel 0? Itt lép be a gondolatmenetbe Kolmogorov 4. pontjának B. része, amelyet az irodalom *Cournot-szabályként* tart számon: ha egy esemény valószínűsége közel van 1-hez, akkor az esemény *gyakorlati szempontból* biztosan bekövetkezik, ha pedig közel van 0-hoz, akkor az esemény *gyakorlati szempontból* biztosan nem következik be. Vagyis a valószínűséget két kitüntetett esetben *gyakorlatilag* azonosíthatjuk a bekövetkezéssel: ha $P(A) \approx 1$, és ha $P(A) \approx 0$. Az első esetben gyakorlatilag biztosak lehetünk abban, hogy A bekövetkezik, a másodikban pedig, hogy nem következik be. Események konjunkciójára a szabály azonban már nem alkalmazható,[19] vagyis a Cournot-szabály kártyaként tehát csak egyszer játszható ki az érvelésben – ez azonban éppen elég Kolmogorovnak: ha a nagy számok törvényeiben a szorzatvalószínűség közel 1 volta a Cournot-szabály értelmében helyettesíthető azzal, hogy a szóban forgó esemény vagy tényállás gyakorlatilag biztosan bekövetkezik, vagyis a relatív gyakoriság és a valószínűség gyakorlatilag biztosan megegyezik, akkor legalábbis nagyszámú kísérletsorozatban a kettőt *gyakorlatilag* azono-

[19] „A B. elv ... nem jelenti azt, hogy az A esemény elég hosszú kísérletsorozatban sem következik be," és hasonlóan: „Az A. elvből semmiképp sem következik, hogy nagyon nagy számú, n hosszúságú kísérletsorozatra *minden* sorozatban az $\frac{m}{n}$ hányados kicsit fog különbözni $P(A)$-tól" – teszi világossá Kolmogorovnak a szakasz végéhez fűzött két megjegyzése.

síthatjuk. A Bernoulli-tétel a Cournot-szabály és a függetlenség feltételezése mellett tehát *bizonyítja*, hogy a relatív gyakoriságok tartanak a valószínűségekhez, azaz a valószínűség relatív gyakoriság-interpretációja a helyes interpretáció. Ez röviden Kolmogorovnál a valószínűségi „axiómák tapasztalati levezetése.”

Ez az érvelés azonban teljességgel helytelen, és nem azért – ahogy azt sokan vélik – mivel a Cournot-szabály megalapozatlan. A tévedés igazi oka az, hogy a nagy számok törvényei *matematikai tételek*, és így teljes mértékben érzéketlenek az interpretációkra: a tételben szereplő p betű interpretációjától függően más és más lesz a fizikai tartalmuk. Ha a p betűt a relatív gyakoriság-interpretációnak megfelelően frekvenciaként interpretáljuk, akkor a tétel a fentebb von Mises kapcsán már említett empirikus jelentést fogja kapni: az olyan sorozatok sorozatainak relatív gyakorisága, amelyekben a sorozat kezdőszeletéből számított relatív gyakoriságok tetszőlegesen kicsit térnek az aszimptotikus relatív gyakoriságoktól, a kezdőszelet hosszával nullához tartanak. Ha a p betűt a szubjektív interpretációnak megfelelően parciális hitként interpretáljuk, akkor a tétel a következőt fogja jelenteni: ha az empirikus/racionális ágens egy adott eseményben p mértékben hisz, akkor (feltéve a függetlenség által modellezett fizikai feltételek fennállását) határértékben teljes mértékben fog hinni abban, hogy az események relatív gyakorisága egy végtelen sorozatban p. Ha a p betűt a *propensity*-interpretációnak megfelelően valamilyen tendenciaként interpretáljuk, amellyek a fizikai környezet az esemény bekövetkezését előidézi, akkor a tétel a következőt fogja jelenteni: ha egy adott esemény p mértékben hajlamos bekövetkezni, akkor (ismét feltételezve a hajlamok függetlenségét) események egy olyan végtelen sorozatának, amelyben a relatív gyakoriság megegyezik p-vel, 1 lesz a bekövetkezési hajlama. Látható, hogy mindhárom állítás empirikus – igazolásukhoz többek között olyan dolgokat kell tudnunk, hogy hogyan teszteljük végtelen sorozatok sorozatát, hogyan mérjük végtelen sorozatok relatív gyakoriságára vonatkozó hiteinket illetve az ilyenek létrehozására vonatkozó *propensity*-t. Ezek tesztelése jelenthet gondot, mindenesetre a nagy számok törvényei mint matematikai állítások elvben mindhárom értelmezéssel kompatibilisek, vagyis semmilyen módon nem tüntetik ki a valószínűség relatív gyakoriság-interpretációját. Ha igaza van tehát von Platonak abban, hogy Kolmogorov frekventizmusát a nagy számok törvényére, a Cournot-szabályra és a függetlenség posztulásására építette, akkor helytelenül járt el. Az elmélet és a tapasztalat közötti szakadékot éppen olyan kevéssé lehet áthidalni a gyakorlati bizonyosság fogalmával, mint von Misesnél az aszimptotikus relatív gyakoriság segítségével. De akár így, akár úgy – tény az, hogy Kolmogorovot a valószínűség mértékelméleti kanonizációja ellenére élete végéig nyugtalanította a valószínűség tapasztalati alkalmazhatóságának kérdése. Későbbi erőfeszítéseit éppen az határozta meg, hogy a mértékelméleti megfogalmazás mellett érvényt szerezzen a relatív gyakoriságra és a véletlenszerűségre érzékenyebb valószínűségfogalomnak. Ezek a kutatások vezették azután az algoritmikus randomitás és a Kolmogorov-komplexitás megalkotásához.

8. fejezet

A *propensity*-interpretáció

A valószínűség *propensity*-interpretációját Popper vetette fel először 1957-ben a kilencedik Colston-konferencián Bristolban:[1]

> „Minden kísérleti elrendezés, ha elegendően sokszor ismételjük, *hajlamos létrehozni* egy olyan sorozatot, amelynek frekvenciája ezen a bizonyos kísérleti elrendezésen múlik. Ezeket a virtuális frekvenciákat valószínűségnek is nevezhetjük. Mivel azonban ezek a valószínűségek a kísérleti elrendezéstől függnek, ezért úgy is tekinthetünk rájuk, mint *a szóban forgó elrendezés tulajdonságaira*, amelyek a kísérleti elrendezésnek azt a *diszpozícióját vagy propensity-jét jellemzik,* hogy bizonyos karakterisztikus frekvenciákat idézzen elő, ha a kísérletet sokszor megismételjük." (67. o.)

Poppert, aki *A tudományos kutatás logikájában* (1934) még a relatív gyakoriság-interpretáció híve, a kvantumelmélet, azon belül is a kétréses kísérlet értelmezése késztette a pálfordulásra. „A kvantumelmélet" – mint írja

> „meggyőzött, hogy a valószínűségeknek 'fizikailag reálisnak' kell lenniük – vagyis fizikai hajlamoknak, fizikai szituációk absztrakt relációs tulajdonságainak, akárcsak a newtoni erők. És ezek nemcsak abban az értelemben 'reálisak', hogy befolyásolni képesek a kísérleti eredményeket, hanem abban az értelemben is, hogy bizonyos körülmények között (koherencia), interferálni, azaz kölcsönhatni is tudnak egymással." (1959, 28. o.)

A kvantumelmélet interferenciajelenségei tehát arról győzték meg Poppert, hogy a valószínűség fogalmának szinguláris esetekre is alkalmazhatónak kell lennie. Szinguláris valószínűségi kijelentések értelmezésére, mint azt az előző fejezetben a referenciaosztály problémája kapcsán láttuk, a frekvenciainterpretáció nem képes. Egy frekventista persze megteheti, hogy egy egyedi esemény valószínűségét egyszerűen azonosítja az eseménynek egy végtelen

[1] Popper maga nem volt jelen a konferencián; az írást Feyerabend olvasta fel.

vagy elegendően hosszú sorozatban vett relatív gyakoriságával.[2] Ez az azonosítás azonban, amint azt a következő példa mutatja, nem mindig kielégítő:

> „Tegyük fel, hogy van egy cinkelt kockánk, és hogy kísérletek hosszú sorával meggyőződtünk arról, hogy a hatos dobás valószínűsége ezzel a kockával jó közelítésben $\frac{1}{4}$. Most tekintsünk egy b sorozatot, amely ezzel a cinkelt kockával való dobásokból áll, de amely sorozat, mondjuk, tartalmaz néhány (kettő vagy esetleg három) dobást egy homogén és szimmetrikus kockával. Nyilvánvalóan a szabályos kockával végzett néhány dobás esetében azt mondanánk, hogy a hatos valószínűsége $\frac{1}{6}$, nem pedig azt, hogy $\frac{1}{4}$; noha ezek a dobások, feltevésünk szerint, egy olyan dobássorozat tagjai, amelyben a hatos statisztikus frekvenciája $\frac{1}{4}$." (1959, 31-32. o.)

Ebből a dilemmából a frekventista számára a kiutat Popper szerint az jelentheti, ha a fenti inhomogén dobássorozatot valahogyan kizárja azon sorozatok köréből, amelyre a valószínűséget értelmezzük. Ezt megteheti például úgy, ha a sorozatot nem extenzíve, aktuális vagy potenciális elemeivel, hanem intenzíve, a „generáló feltételek halmazán" keresztül adja meg. A valószínűséget tehát a véletlent generáló mechanizmus, vagyis a kockának, az asztalnak, az eldobás körülményeinek stb. együttes fizikai tulajdonságai határozzák meg. Ezek a tulajdonságok nyilvánvalóan mások lesznek a cinkelt és a szimmetrikus kocka esetében, és ez megmagyarázza valószínűségük különbségét.

Ez az ártatlannak tűnő módosítás azonban a frekvenciainterpretációtól átvezet a *propensity*-interpretációra. Nagy különbség ugyanis azt mondani, hogy egy szinguláris esemény valószínűségét az a sorozat határozza meg, amelynek az tagja, vagy azt mondani, hogy a szinguláris esemény valószínűsége az eseményt generáló feltételek tulajdonsága. Ez utóbbi esetben egy szinguláris eseménynek akkor is lehet valószínűsége, ha csak egyszer vagy éppen egyszer sem következik be.

A valószínűség kapcsolatba hozása a generáló feltételekkel persze nem újkeletű. Valószínűségelmélete tapasztalati alkalmazhatóságának vizsgálatánál Kolmogorov is hasonlóan járt el (v.ö. 7.2. fejezet). Ami Poppernél új – és ezt pusztán már a *propensity* szóhasználat is tükrözi – az az, hogy a szóban forgó generáló feltételek nála diszpozicionális értelmet nyernek.[3] A *propensity* Popper szerint a fizikai rendszer hajlama, *diszpozíciója*, hogy bizonyos „szinguláris eseményeket realizáljon" (1959, 28. o.). De mik is azok a diszpozíciók?

[2] Vagy, ahogy Reichenbach (1949, 376-77. o.) tette, a szinguláris eseményre vonatkozó kijelentést az eseménytípusra vonatkozó kijelentés *elliptikus* formájának tekinti.

[3] A *propensity*-interpretációt már 1910-ben megelőlegezte C. S. Peirce, amikor a következőket írta: „Az a kijelentés, hogy 'Ha a kockát eldobjuk, akkor egyharmad a valószínűsége, hogy az eredmény osztható hárommal' ... azt jelenti, hogy a kocka rendelkezik egy bizonyos *would-be*-vel; és azt mondani, hogy a kocka rendelkezik egy bizonyos *would-be*-vel annyi, mint azt állítani, hogy a kocka rendelkezik egy olyan tulajdonsággal, amely egészen analóg az emberi szokásokkal."

8.1. A *propensity* mint diszpozíció

A bevett metafizikai felosztás szerint a tárgyak kétfajta tulajdonsággal rendelkezhetnek: kategorikus és diszpozicionális tulajdonságokkal. Amíg a kategorikus tulajdonságok egy tárgyat abból a szempontból jellemeznek, hogy hogyan állnak a dolgok *aktuálisan* a szóban forgó tárggyal kapcsolatban, addig a diszpozicionális tulajdonságoknak *modális* implikációi vannak. A teáscsészém 20 dkg tömegű és törékeny. A tömeg a csészét aktuálisan jellemzi, a törékenység azonban egy lehetőségre utal: ha a csésze leesne az asztalról, eltörne. A szokásos szemantikai elemzés szerint a diszpozicionális tulajdonságokat épp az különbözteti meg a kategorikus állításoktól, hogy a fentihez hasonló kontrafaktuális állításokat támogatnak. A csésze törékenysége azt *jelenti*, hogy, amennyiben leesne az asztalról, eltörne.

A diszpozicionális tulajdonságok kondicionális elemzésével szemben számos kérdés felvethető. Elsőként: vajon a kategorikus tulajdonságok nem támogatnak-e hasonló módon kontrafaktuális állításokat: a csésze 20 dkg tömegű, ti. ha adott erővel hatnék rá, adott módon gyorsulna. Így tekintve a tömeg is diszpozicionális tulajdonság; vagyis minden tulajdonság diszpozicionális. Másrészt, ha diszpozíciót tulajdonítani egy tárgynak szemantikailag semmi más, mint egy kontrafaktuális kondicionális, akkor úgy tűnik, hogy le kell mondanunk azokról a diszpozíciókról, amelyek fennállása és manifesztációja nem jár együtt, vagyis az olyan – az irodalomban mímelőnek illetve *finkish*-nek keresztelt – diszpozíciók létéről, amelyeket éppen a kontrafaktuális kijelentés antecedense hoz létre vagy tüntet el, és így fennállásuk sem nem szükséges, sem nem elégséges feltétele a kontrafaktuális teljesülésének. Suta példával: Az amúgy nem törékeny csésze valamely rejtélyes oknál fogva mindig éppen akkor válna törékennyé, amikor ledobom; és így a kontrafaktuális anélkül teljesülne, hogy a csésze a törékenység diszpozíciójával rendelkeznék. Az ilyen rejtélyes okok kizárása nem egyszerű feladat a kondicionális elemzés képviselője számára. Megpróbálkozhat azzal, hogy a kontrafaktuálisba beiktat egy *ceteris paribus* kazulát: 'A csésze törékeny' azt jelenti, hogy amennyiben ledobnám, akkor – *ceteris paribus* – eltörne. Mivel azonban a rejtélyes okokról semmi közelebbit nem tudnánk megadni, ez a lépés vagy kiüresítené a kontrafaktuálist: a csésze eltörne, hacsak valamiért nem törne el, vagy körkörössé tenné: a csésze eltörne, hacsak valamiért el nem veszítené törékenységét.

A diszpozíciók kondicionális elemezhetőségének szemantikai kérdésében legtöbbször *ontológiai* álláspontok ütköznek. A kondicionális elemzés képviselői általában *antirealisták* a diszpozíciókkal kapcsolatban, vagyis a diszpozicionális tulajdonságokat kategorikus tulajdonságokra igyekeznek visszavezetni: a törékenység *valójában* nem más, mint a csésze bizonyos anyagszerkezeti tulajdonsága. A visszavezetés röviden úgy történik, hogy a diszpozíciókat első lépésben valamilyen funkcióval, például kauzális szerepükkel azonosítják: törékeny az, amit ha ledobok, eltörik; majd második lépésben megállapítják, hogy az illető kauzális szerepet az *aktuális* világban mely kategorikus tulajdonságok töltik be: pl. bizonyos anyagszerkezeti tulajdonságok. Így a diszpozíciók és kategorikus alapjuk azonossága kontingens azonosság lesz. Egy másik lehetséges világban más anyagszerkezettel rendelkező tárgy is lehetne törékeny. Ezzel az azonosítással az antirealisták a diszpozíciókat megfosztják kauzális erejüktől, és a kauzalitás magyarázatát, a bevett hume-i módon, a világ kategorikus tulajdonságai között tapasztalható korrelációkra bízzák. A törékenységben kifejeződő kauzalitás nem más, mint a megfelelő anyagszerkezeti tulajdonság fennállása és a csésze törését jellemző kategorikus tulajdonságok együttjárása. Röviden az antirealisták szerint sem a szemantikai, sem a kauzalitási problémák nem indokolják a diszpozíciók felvételét a világ „hume-i mozaikjába".

A *diszpozíció-realista* ellentábor ezzel szemben abból a meggyőződésből indul ki, hogy a diszpozíciók reálisan léteznek. A csésze törékenysége épp olyan reálisan létező tulajdonsága a csészének, mint a tömege. Szerintük már a kondicionális elemzés buktatói is arra utalnak, hogy a diszpozíciók kiküszöbölhetetlen ontológiai entitások. Ami a diszpozíciók és a kauzalitás viszonyát illeti, a realisták szerint éppen fordított a helyzet a hume-iánus elképzeléshez képest. Ott a diszpozíciók ráépülnek a világ kategorikus mintázatára, amely mintázat a kauzális viszonyokat is teljes mértékben meghatározza – így a diszpozíciók kauzalitása csak üres frázis. A diszpozíció-realisták ezzel szemben valós kauzális hatékonyságot tulajdonítanak a tárgyaknak, amely nem szuperveniál a tárgyak kategorikus tulajdonságainak korrelációin. A diszpozíciókban kifejeződő kauzalitás akkor is jelen van, ha az nem manifesztálódik aktuálisan. A csésze akkor is törékeny lehet (sőt, valójában csak akkor), ha sohasem esik le, és törik össze. Ráadásul a világban tapasztalható regularitások éppen a tárgyak diszpozíciókban kifejeződő kauzális hatékonyságának segítségével *magyarázhatók*: a csésze azért törik össze, ha leesik, *mert* törékeny.

Most térjünk vissza a *propensity* kérdéséhez! *Propensity*-interpretációjával Popper (1959) a diszpozíció-realisták táborába tartozik:

> „A *propensity*-k lehetőségekként értelmezhetők (vagy lehetőségek mértékeként illetve 'súlyaként'), amelyek tendenciával vagy diszpozícióval vannak ellátva, hogy realizálódjanak." (30. o.)

A *propensity* tehát diszpozíció, ugyanakkor reális fizikai létező:

> „Az elméletnek egy *nem megfigyelhető* fizikai realitással van dolga, amely realitásnak csak néhány felszíni hatása figyelhető meg, és amely hatások lehetővé teszik számunkra az elmélet tesztelését." (35. o.)

Popper szerint tehát a *propensity*, akárcsak a törékenység, a tömeg vagy az erő reális diszpozíció:

> „Az erő – vagy jobban mondva az erőtér – fogalma egy diszpozicionális fizikai entitást vezet be, amelyeket bizonyos egyenletek írnak le (nem pedig metaforák), és amely megfigyelhető gyorsulásokat magyaráz. Hasonlóan a *propensity* vagy *propensity*-mező fogalma is egy szinguláris fizikai kísérleti elrendezés – azaz egy szinguláris fizikai esemény – diszpozicionális tulajdonságát vezeti be, hogy megmagyarázzon bizonyos megfigyelhető frekvenciákat ezen ismétlődő események sorozataiban. Mindkét esetben az új fogalom bevezetése kizárólag a fizikai elméletekben hajtott hasznukra való hivatkozással igazolható." (31. o.)

A fenti idézetekből világossá válik, hogy Popper szerint a *propensity* (i) egy reálisan létező fizikai entitás, (ii) megfigyelhető hatásokban manifesztálódik, és (iii) posztulálását épp ezen hatások magyarázatában játszott szerepe indokolja. Az állításnak azonban mindhárom pontja értelmezést kíván. Jelesül az alábbi három kérdést kell tisztáznunk:

(i) Mely ontológiai entitás rendelkezik *propensity*-vel?

(ii) Mit tekintsünk a *propensity* manifesztációjának?

(iii) Milyen értelemben magyarázza manifesztációit a *propensity*?

Mint azt az alábbiakban látni fogjuk, e három kérdés igen csak megosztotta a *propensity*-tábort, és a kérdésekre adott válaszok nyomán az elképzelések tarka serege jött létre. A következő alfejezetben a fenti kérdések mentén haladva áttekintjük a *propensity*-interpretációk különféle változatait, majd a rákövetkező fejezetben szisztematikusan megvizsgáljuk, hogy mennyiben is tartható a valószínűség *propensity*-értelmezése.

8.2. A *propensity* hordozója, manifesztációja és szerepe

Kezdjük az első kérdésre adott válaszok áttekintésével: Mely ontológiai entitás rendelkezik *propensity*-vel? Popper (1957) szerint a *propensity* hordozója a teljes fizikai szituáció:

> „Az $\frac{1}{4}$ *propensity* nem a cinkelt kocka tulajdonsága. Ezt rögtön látjuk, ha belegondolunk abba, hogy egy nagyon gyenge gravitációs térben a cinkelésnek milyen kicsi hatása lesz – annak valószínűsége, hogy hatos dobjunk, $\frac{1}{4}$-ről $\frac{1}{6}$-ra fog csökkenni ... A tendencia vagy diszpozíció vagy *propensity* ezért magának a kísérleti szituációnak relációs tulajdonsága.” (68. o.)

De mi is jellemez egy kísérleti szituációt? A kockát eldobhatjuk jobb kézzel vagy bal kézzel, laposan vagy pörgetve, tegnap, ma és holnap. Vajon itt ugyanarról a kísérleti szituációról van szó? Popper szerint a *propensity* „a kísérleti elrendezés azon feltételeinek tulajdonsága, amelyeket állandónak szándékozunk tartani” (1959, 37. o.) De melyek is ezek? Az asztal, a kocka és a közvetlen környezet, a dobás hozzávetőleges sebessége és magassága? Mely feltételeket kell rögzíteni, és melyek változhatnak dobásról-dobásra? Végezhetem-e például a dobást egy precízen hangolt hajítógéppel? A kérdéseket nem válaszolja meg Hacking (1965) megoldása sem, aki a *propensity*-t egy véletlengenerátor (*chance setup*) tulajdonságának tekinti, ahol is

> „a *véletlengenerátor* egy berendezés, vagy a világ egy része, amelyen egy vagy több *kísérletet* vagy megfigyelést végezhetünk, amely kísérletek mindegyike egyértelmű rendelkezik valamilyen *kimenettel lehetséges kimenetek egy családjából.*” (13. o.)

Így a kockával való dobások sorozata más-más véletlengenerátort határoz meg attól függően, hogy a hatos dobások, vagy mondjuk az egymást követő azonos paritású dobások *propensity*-jére vagyunk kíváncsiak, mivel a két esetben más lesz a lehetséges kimenetek családja. De a *propensity*-nek a véletlengenerátorokhoz rendelése még nem válaszolja meg az eredeti kérdést, hogy ti. *mi számít a hatos dobás véletlengenerátorának*?

A kérdésre adott radikális válasz Melloré (1971), aki Popperrel ellentétben vitatja, hogy a *propensity* a teljes kísérleti szituáció sajátossága volna. Szerinte „a *propensity*-t nem szabad az érme, a hajító berendezés és a környezet teljes együttesének tulajdonítani." (75. o.) Ez olyasmit jelentene, mint „azt mondani, hogy egy vékony üveg törékeny csak egy kemény padlóval együtt lehet, egy csipet só pedig csak egy vödör vízzel együtt lehet oldható". Más szóval egy diszpozíció akkor is megillet egy entitást, ha a manifesztációjához szükséges környezet nincs jelen. Mellor szerint tehát a *propensity* nem a kísérlet szituáció, hanem a kísérlet szituáció valamely permanens létezőjének tulajdonsága.

> „A *propensity*-t a kockának vagy az érmének tulajdonítjuk, és nem a teljes berendezésnek [setup], amely csak a manifesztálódás során van jelen, mivel a konvenció ezt a permanensebb létezőt választja ki a kísérletben jelen levő többi elem közül." (75. o.)

A *propensity* tehát a kockának mint a kísérleti szituáció permanens létezőjének tulajdonsága. A permanens elem kiválasztása persze egy pontig konvencionális, ahogy az oldékonyságot is tulajdoníthatjuk hol a szilárd anyagnak, hol az oldószernek attól függően, hogy a szituációban háttérként melyik van rögzítve: „a só oldható (vízben)", ugyanakkor „a királyvíz oldja az aranyat".

Végül hadd említsünk meg egy további választ, Fetzerét (1971) és Giere-ét (1976b), amely szerint a *propensity* sem nem az állandónak tartott kísérleti feltételek, sem nem valamely, az egyedi eseményeken átívelő permanens entitás tulajdonsága, hanem a szinguláris kísérleti eseményé, azaz a kocka egyszeri eldobásáé. Giere így ír erről:

> „Amikor valószínűségi értéket tulajdonítunk egy kimenetnek, akkor az közvetlenül az egyedi kísérlethez [trial] tartozik – és nem szükséges egyedi kísérletek halmazára hivatkoznunk." (71. o.)

Most térjünk át a második kérdésre: Mi a kapcsolata a kockával való hatos dobás diszpozíciójának az aktuális dobások kimenetéhez? A diszpozíció-realista egyfelől megkülönbözteti a kocka (vagy a generáló feltételek) *propensity*-jét mint reális tulajdonságot a megvalósult kimenetektől, akárcsak a törékenység diszpozícióját a konkrét leesésektől és törésektől. A *propensity* két módon kapcsolódhat a realizált kimenetekhez: vagy az egyedi kimenetekhez kapcsolódik vagy kimenetek egy hosszú vagy végtelen sorozatához. Ez különbözteti meg a *propensity*-interpretációk két fajtáját: a *hosszútávú* és a *szinguláris propensity*-interpretációt. Gillies (2000a) így írja le a különbséget:

> „A hosszútávú *propensity*-elmélet a *propensity*-t ismételhető feltételekhez köti, és arra való hajlamnak tekinti, hogy a feltételek hosszú távú ismétlésével a valószínűséghez közeli frekvenciák adódjanak. A szinguláris *propensity*-elméletben viszont a *propensity* egy olyan hajlam, amely egy adott egyedi esetben produkál egy bizonyos eredményt." (126. o.)

A kocka (vagy a teljes kísérleti szituáció) *propensity*-je tehát vagy egy adott dobás során egy bizonyos kimenet vagy pedig dobások egy megfelelően hosszú sorozatában egy bizonyos frekvenciamintázat produkálására való hajlam. Elvben mindkét értelemben vett *propensity* a kocka egyéb fizikai tulajdonságaihoz hasonlóan időben változhat. Ha a kocka a dobálások során kopni kezd, akkor tömegéhez hasonlóan a hatos dobás *propensity*-je is változhat, akár a hosszútávú, akár a szinguláris értelemben.

Popper megítélése a kétfajta *propensity*-interpretáció tekintetben nagyon nehéz, mivel egyazon cikkében mindkét értelmezésre utaló kijelentések előfordulnak: „a *propensity* szinguláris események realizálására való hajlam" (1959, 28. o.), és „ …*propensity*, hogy olyan sorozatokat hozzon létre, amelynek frekvenciája megegyezik a valószínűséggel" (1959, 35. o.). Gillies (2000a) a korai Poppert (1957) mégis inkább a hosszútávú, míg a későbbi Poppert (1990) a szinguláris *propensity*-interpretáció képviselői közé sorolja, mivel az előbbinél inkább az ismételhető feltételekre, az utóbbinál pedig az egyedi esetekre való alkalmazhatóságra kerül a hangsúly. A kései Popper szingularista értelmezését vitte tovább Fetzer (1981) és Miller (1994). Ez utóbbi így fogalmaz:

> „A *propensity*-interpretációban a kimenet valószínűsége nem valamilyen relatív gyakoriságnak a mértéke, hanem a dolgok jelenlegi állása hajlamának mértéke arra, hogy a szóban forgó kimenetet realizálják." (182. o.)

Fetzer és Miller álláspontja lényegében abban különbözik, hogy milyen szélesek legyenek azok a feltételek, amelyek „a dolgok jelenlegi állását" jellemzik.

A szingularisták egy további képviselője Mellor, akinek a szingularista elképzelés melletti döntését a diszpozíciókkal szemben kialakított általános álláspontja motiválja. A diszpozíciókat az irodalomban szokás két csoportra osztani. Determinisztikus (*sure-fire*) diszpozícióknak nevezik azokat a diszpozíciókat, amelyek mindig manifesztálódnak, amennyiben a manifesztációhoz szükséges körülmények fennállnak; valószínűségi diszpozícióknak pedig azokat, amelyek még ebben az esetben sem feltétlenül manifesztálódnak. Amíg a tömeg determinisztikus diszpozíció, mivel adott tömegű test, ha adott erővel gyorsítanám, mindig adott módon gyorsulna, addig a hatos dobás *propensity*-je valószínűségi diszpozíció, mivel a dobás eredménye hol hatos lesz, hol nem.

A determinisztikus-valószínűségi megkülönböztetés természetesen nem jelenhet meg egy olyan gondolatmenetben, amely a *propensity* fogalmával éppen a valószínűséget igyekszik értelmezni. Mellor álláspontja azonban ennél radikálisabb. Szerinte valószínűségi diszpozíciók egyáltalán nem léteznek. Egy diszpozíciónak, ha az valóban számot tart nevére, minden egyes szinguláris esetben manifesztálódnia kell, amennyiben a manifesztáció feltételei adottak.

Ez az elképzelés azonban érdekes eredményre vezet a *propensity*-t illetően. Ha ugyanis a hatos dobás *propensity*-je diszpozíció, a fenti terminológiában determinisztikus diszpozíció, *és* a manifesztáció szinguláris, azaz az egyedi dobások során történik, akkor kérdés, hogy *hol* is képes manifesztálódni a hatos dobás diszpozíciója ezekben az egyedi dobásokban? Nyilvánvalóan nem a puszta kimeneten, hiszen az hatféle lehet. A puszta kimenetek még csak egy „tendenciát" mutatnak, és nem a diszpozíciót manifesztálják. Mellor (1971) válasza különös: „a *propensity* manifesztációja [display] a lehetséges kimenetek eloszlása az adott kísérleti futamban."[4] (70. o.)

A *propensity* tehát minden egyes dobás során teljes mértékben manifesztálódik, mégpedig a kimenetek eloszlásfüggvényén. Az eloszlásfüggvényt természetesen nem lehet leolvasni egyetlen kimenetről – ismeri el készséggel Mellor –, ehhez nagyszámú egyedi kísérletek szükségesek valamint különféle statisztikus tesztek. Azonban „fontos hangsúlyozni, hogy a manifesztáció fogalma nem jelent közvetlen megfigyelhetőséget." (81. o.) Az eloszláshoz való speciális episztemikus hozzáférésünk még nem ok arra, hogy ne azt tekintsük a *propensity* manifesztációjának.

Hogy mégis milyen értelemben manifesztálódik a *propensity* az egyedi dobásokon, arra Mellor (1971) kerülőúton, saját szubjektivista elméletén keresztül válaszol.

> „A *propensity* és a perszonalista [szubjektivista] elmélet közötti viszony ez: Az utóbbi szerint egy valószínűségi kijelentés a beszélő 'parciális hitét' fejezi ki abban, aminek valószínűséget tulajdonít, mondjuk, hogy egy *a* érme fejre esik, ha feldobjuk. Az érme *propensity*-jének ismerete a jelenlegi elmélet alapján az, ami megfelelő körülmények között ésszerűvé teszi a dobás eredményében való parciális hitet. Az érme objektív valószínűsége [chance], hogy fejre essen, így az ésszerű racionális hit mértéke." (2. o.)

Az egyedi eseményekhez tehát egy elmélet rendel eloszlásfüggvényt, az az elmélet, amely leginkább igazolni képes racionális hiteinket, azaz a szubjektív valószínűséget. Egy ilyen elméletet a vizsgált rendszer által produkált relatív gyakorságok és a rendszer egyéb fizikai tulajdonságai (szimmetriái) együttesen határoznak meg.[5] A *propensity*-k ennek az elméletnek a posztulátumai; az elmélet közvetítésével ezek igazolják szubjektív valószínűségeinket.

Realista azonban a szintén szingularista Giere (1976a), aki arra a kérdésre, hogy „Metafizikailag tekintve, mik is a *propensity*-k?", így válaszol:

[4] Az elképzelés mögött ismét kvantumelméleti megfontolások állnak. A *propensity*-nek a mérendő tárgyhoz és nem a mérési szituációhoz való rendelése Mellor szerint párhuzamos azzal a kvantumelméleti lépéssel, hogy a hullámfüggvényt nem a teljes kísérleti szituációhoz, hanem magához a részecskéhez rendeljük. A mérési szituációt operátorokkal jellemezzük, a mérési kimenetek valószínűségét pedig ezekből az operátorokból és a részecske hullámfüggvényéből számoljuk. Így tekintve *propensity*-je tehát a *részecskének* van, hogy a különböző kísérleti szituációkban különböző súllyal különböző kimeneteket produkáljon.

[5] A meghatározás mikéntjéhez Mellor sok mindent felhasznál, többek között Lewis legjobb rendszer analízisét, valamint *Principal Principle*-jét, valamint saját konnektivitási-elvét. Ezek a részletek azonban itt nem játszanak szerepet.

> „Súlyok stochasztikus kísérletek fizikailag lehetséges állapotain – súlyok, amelyek valószínűségi eloszlást generálnak az állapotok halmazain." (332. o.)

Vagyis a *propensity*-kijelentések igazságfeltételei a lehetséges világok fölötti súlyokkal adhatók meg. Ezek után annak a valószínűségi kijelentésnek, hogy „$p(E) = r$" a helyes interpretációja így hangzik: „A véletlengenerátor [chance setup] azon *propensity*-jének az ereje r, hogy az E kimenetet hozza létre." (1973, 471. o.) Giere szerint tehát a *propensity* egy tendencia, kauzális erő, frekvenciákra nem redukálható reális diszpozíció.

A hosszútávú *propensity*-interpretációt Popper után leginkább Hacking és Gillies képviselte. Hacking (1965) szerint a *propensity* „az érme diszpozíciós tulajdonsága, hogy mi vagy mi lenne a hosszútávú frekvencia" (10. o.); Gillies (2000b) szerint pedig

> „$P(A|S) = p$ azt jelenti, hogy létezik egy p *propensity* arra, hogy S-et [az ismételhető feltételeket] sokszor ismételve az A esemény p-hez közeli relatív frekvenciával forduljon elő." (828. o.)

Vagyis a hosszútávú *propensity*-értelmezések szerint a kocka vagy a kísérleti együttes *propensity*-jének nem az egyes kimenetek produkálására van hajlama, hanem hosszútávú frekvenciák létrehozására. Ezek a frekvenciák vagy megegyeznek a *propensity* értékével, vagy csak egy kicsit térnek el tőle, interpretációtól függően. De még ha megegyeznek is vele, a frekvencia nem maga a *propensity*, csupán annak manifesztációja, amely megfelelő statisztikus tesztek segítségével konfirmálja a *propensity*-hipotézist. A *propensity* itt is, mint fentebb, determinisztikus diszpozíció: a manifesztáció feltételei mellett a diszpozíció „mindig" tüzel, azaz létrehozza a megfelelő frekvenciamintázatot.

Végül térjünk át a harmadik kérdésre: Milyen értelemben magyarázza manifesztációit a *propensity*? A kérdésre ilyen válaszokat kapunk:

> Popper (1959): „A *propensity*-ket azért vezetjük be, hogy segítsenek megmagyarázni és megjósolni bizonyos sorozatok statisztikus tulajdonságait; és *ez minden funkciójuk.*" (30. o.)
>
> Fetzer (1973): „a *propensity*-interpretáció elméleti alapot nyújt ahhoz, hogy ezekről a mintázatokról a *rendszer kezdőfeltételei* révén adjunk számot, mivel a bekövetkezések aktuális frekvenciái az őket generáló diszpozicionális tendenciák segítségével magyarázhatjuk." (10. o.)
>
> Giere (1973): „a *propensity*-interpretáció mögötti intuitív ötlet az, hogy a stochasztikus rendszerhez társított valószínűségi eloszlások olyan kauzális tendenciák eloszlása, amely nem redukálható relatív frekvenciákra, legyenek azok aktuálisak vagy lehetségesek." (327. o.)

Mint az idézetekből látható, a *propensity* bevezetésének közös indítéka a statisztikus regularitások magyarázatának vágya – vagyis az a remény, hogy a *propensity* diszpozicionális tulajdonságának posztulálásával magyarázhatóvá válnak az észlelt statisztikus tulajdonságok.[6]

Miben áll egy ilyen magyarázat? A hosszútávú *propensity*-értelmezés hívének itt könnyebb a dolga, mivel a *propensity* közvetlenül a frekvenciákkal áll kapcsolatban. A magyarázat valójában a *propensity* és a megfelelő frekvenciaminta közötti oksági viszonyra hivatkozik. A kocka vagy a teljes kísérleti berendezés a dobálások során azért produkálja a megfelelő frekvenciát, *mert* rendelkezik azzal a kauzális erővel, hogy ilyen frekvenciákat hozzon létre. A magyarázat szándékoltan hasonló ahhoz, ahogy az egyéb diszpozíciók kauzális hatását használjuk arra, hogy velük egy regularitást megmagyarázzunk: az üveg azért törik el, ha ledobom, *mert* törékeny; a test azért gyorsul adott erőhatásra adott mértékben, *mert* adott tömeggel, tehetetlenséggel rendelkezik. A magyarázat szerkezete tehát egyszerű: a *propensity*-nak kauzális hatást tulajdonítunk frekvenciák produkálására, amely aztán a *propensity* jelenléte mellett magyarázza a frekvenciákat.

A szingularista *propensity*-értelmezés hívének mindezekhez még egy plusz lépést is meg kell tennie: összekapcsolnia a szinguláris *propensity*-t a hosszútávú *propensity*-vel. Ez legtöbbször a nagy számok valamelyik törvényére hivatkozva történik. Tekintsük a pénzfeldobás esetét. Legyen p az érme vagy a teljes berendezés azon szinguláris *propensity*-jének értéke, hogy az érme eldobására fejet kapunk. Mivel a *propensity* az érmétől vagy a teljes elrendezés állandónak tartott tulajdonságaitól függ, ezért *feltehető*, hogy értéke minden individuális dobás során független és azonos. A *propensity*-k függetlenségének hipotézise az erők szuperpozíciójának hipotéziséhez hasonlít a mechanikában. Ezen empirikus feltevések után modellezzük az egyedi dobásokat a $\mathcal{P}(\{0,1\})$ halmazalgebrával, az érme *propensity*-jét pedig az algebrán értelmezett p valószínűségi mértékkel. A függetlenség miatt az n dobás *propensity*-je, hogy csupa különböző számú fejeket és írásokat kapjunk, a $\mathcal{P}(\{0,1\}^n)$ eseménytéren értelmezett szorzatmértékből számolható.

A szingularista ezen a ponton hívja segítségül a nagy számok valamelyik törvényét. A gyenge törvény „*propensity*-interpretációja" szerint az n dobásból álló sorozat azon *propensity*-je, hogy a sorozatban a fejek relatív gyakorisága megegyezzen p-vel, az $n \to \infty$ határértékben 1-hez tart. Az erős törvény szerint annak *propensity*-je 1, hogy az $n \to \infty$ határértékben az n dobásból álló sorozatban a fejek relatív gyakorisága megegyezzen p-vel. Az első esetben tehát egy 1 értékű, a második esetben pedig egy 1-hez tartó *propensity*-t kapunk. De honnan tudjuk, hogy aminek a *propensity*-je 1 vagy határértékben 1-hez tart, az bekövetkezik? Az előbbi esetben nem tudhatjuk, de az utóbbiban igen – állítja Popper (1983), aki miután leszámol a nagy számok gyenge törvényének használhatóságával, az erős törvényre vonatkozóan így érvel:

„Más a helyzet, ha a valószínűség [*propensity*] pontosan megegyezik 1-gyel (vagy

[6] Giere – Popperhez hasonlóan – a *propensity* fogalmának további hasznát abban látja, hogy segítségével lehetőség nyílik a kvantummechanikai valószínűségfogalom magyarázatára. Mint írja: „a *propensity* paradigmatikus példái a kvantumjelenségek". Giere a kvantumelmélet koppenhágai- és a Ballentine-féle statisztikus értelmezés mellett ajánlja a maga *propensity*-értelmezését, amely nézete szerint a kvantumelmélet egyetlen olyan interpretációja, amely komolyan veszi a rejtett paraméterek nem létezését.

> 0-val a nullmértékű esetben). Kétségkívül, a 'valószínűség'-nek ebben az esetben is valami olyat kell jelentenie, ami kapcsolatban áll a relatív gyakorisággal. De a kapcsolatnak nem kell pontosnak lennie – nincs határérték axióma és nincs véletlenségi axióma [von Mises értelmében], mivel ezeknek amúgy is csak a nulla valószínűségű (mértékű) eseteket leszámítva kell érvényesnek lennie, így elhagyhatók. Így pusztán annyit kell feltételeznünk, hogy a nulla valószínűség (vagy nulla mérték) a véletlen eseményeknél *olyan valószínűséget jelent, amely elhanyagolható mintha lehetetlen volna.*"

Vagyis Popper szerint *propensity* és a relatív gyakorság között hidat az a feltevés alapozza meg, hogy aminek a *propensity*-je 1, az biztosan bekövetkezik. Mint emlékszünk, ez nem más, mint a Cournot-szabály, amelyre már Kolmogorov is hivatkozott, amikor elméletének tapasztalati alapjait vizsgálta.[7]

Röviden tehát a szingularista *propensity*-értelmezés a *propensity*-k létének és függetlenségének empirikus hipotézise után a nagy számok erős törvényére hivatkozva magyarázza meg, hogy a szinguláris *propensity* hosszú távon miért egyezik meg a frekvenciával.[8]

Mielőtt alaposabban megvizsgáljuk a fenti három kérdésre együtt és külön-külön adott válaszok konzisztenciáját, összefoglalásképpen soroljuk fel a kérdésekre adható lehetséges válaszokat és ezek összefüggéseit. Arra kérdésre, hogy mi a hatos dobás *propensity*-jének a hordozója, lényegében három lehetséges válasz van: 1. az ismételhető kísérleti elrendezés, 2. az egyedi dobások, 3. a kocka. Ontológiailag ez három kategóriának felel meg: 1. egy eseménytípusnak, 2. egy szinguláris eseménynek, 3. egy fizikai tárgynak. A második kérdésre, hogy ti. mit tekintsünk a *propensity* manifesztációjának, két válasz lehetséges: 1. az egyedi kimeneteket, 2. egy frekvenciamintázatot. A harmadik kérdésben a magyarázatok vagy 1. egy lépésben összekapcsolják a *propensity*-t és a frekvenciákat, a *propensity*-nak kauzális hatákonyságot tulajdonítva, vagy 2. a frekvenciákat a szinguláris *propensity*-kből vezetik le valamilyen módon, pl. a nagy számok valamelyik törvényére hivatkozva.

A kérdésre adott válaszok nem függetlenek egymástól. Aki a *propensity* manifesztációját a frekvenciákban keresi, az a *propensity*-t vagy az ismételhető kísérleti elrendezés tulajdonságának tekinti vagy magának a kocka tulajdonságának. Ami a magyarázatot il-

[7] Poppernek erre a feltevésére von Mises reakciója a következő volt: ha két speciális értékre, ti. a 0-ra és az 1-re feltesszük, hogy ott a *propensity* megegyezik a relatív frekvenciával, akkor miért nem azonosítjuk a *propensity*-t a többi értékre is a relatív gyakorisággal, vagyis miért nem képviseljük kezdettől fogva a frekvenciainterpretációt?

[8] Giere (1973) egyenesen arra vállalkozik, hogy a nagy számok törvényére hivatkozva kimutassa, a valószínűség mint *propensity* miért szükségképpen diszpozíció, vagyis miért nem helyezhető el egy hume-i világban. Egy hume-i világban ugyanis mind a determinisztikus, mind a valószínűségi természettörvények az aktuális világ partikuláris eseményeire épülnek, ez utóbbiak a frekvenciákra. Mivel azonban a nagy számok törvényei szerint a valószínűségek csak (a szorzatmértékben vett) valószínűség erejéig azonosíthatók a frekvenciákkal, vagyis *de facto* nem azonosíthatók velük, így – érvel Giere – a valószínűségnek szükségképpen valami többnek kell lennie, mint amennyi a világ hume-i mintázatából kiderül – azaz, valami sajátosan modálisnak.

leti, a *propensity* kauzális hajlama ebben az esetben korlátozódhat az egyedi kimenetekre, vagy kiterjedhet a teljes frekvenciamintázatra. Aki szerint ellenben a *propensity* az egyedi kimenetekben manifesztálódik, az a *propensity*-t az egyedi dobások diszpozíciójának tartja, és a frekvenciák magyarázatánál szükségképpen használja a nagy számok törvényét vagy valami hasonlót. Az egyetlen keresztpárosítás Melloré, aki szerint a *propensity* hordozója a kocka, a manifesztációra mégis az egyedi kimenetek eloszlásában kerül sor. Nála, mint láttuk, a magyarázat is komplex: a *propensity* egy, az ésszerű parciális hitet igazoló diszpozíció.

Ezen történeti összefoglaló után az alábbiakban szisztematikusan megvizsgáljuk, hogy vajon összességében tartható-e a valószínűség *propensity*-interpretációja.

8.3. Tartható-e a *propensity*-interpretáció?

Térjünk át tehát annak a kérdésnek a vizsgálatára, hogy a *propensity* fogalmának bevezetése valóban olyan eszközt ad-e a kezünkbe, amelynek segítségével a valószínűség természetét értelmezni tudjuk. Kezdjük ismét az első kérdésre, a *propensity* hordozójának kérdésére adott válaszok elemzésével.

Ha a *propensity*-t egy ismételhető kísérleti elrendezéshez rendeljük, akkor meg kell mondanunk, hogy mit is értünk egy ilyen kísérleti elrendezés alatt. A kísérleti szituációt ebben az esetben nem jellemezhetjük az összes létező tulajdonsága révén, beleértve téridőbeli koordinátáit is, mivel egy ilyen leírásnak megfelelő esemény szinguláris lesz, csak egyszer fog megtörténni. A tulajdonságok között tehát válogatnunk kell. Popper elképzelése, mint láttuk, az volt, hogy csak az állandónak tekintett tulajdonságokat vegyük bele a kísérleti szituáció karakterizációjába. Ez a gondolat húzódik meg részben azon elképzelés mögött is, hogy a *propensity*-t a szituáció permanens elemének, azaz a kocka tulajdonságának tekintsük. Ezzel az elképzeléssel szemben azonban rögtön felmerül a kérdés, hogy vajon időben változó kísérleti körülmények miért ne karakterizálhatnának egy kísérleti elrendezést. Ha a kockát dobásonként egy arasszal magasabbról hajítom el: ez az elrendezés vajon nem rendelkezik *propensity*-vel? Miért nem változhatnának a kísérleti berendezés sajátosságai időben, ha egyszer változhatnak térben? Az a tulajdonság, hogy a dobások magassága araszonként nő, ugyanúgy jellemezheti a kísérletet, mint az, hogy az asztal balra lejt. Így állandóság helyett talán helyesebb volna szabályozottságot mondani. Ha azonban az állandóságról átkerül a hangsúly a szabályozottságra, akkor nincs megállás: bármely szabállyal leírható dobássorozat kísérleti szituációnak fog számítani; jelesül azok az „inhomogén” dobássorozatok is, amelyek miatt Popper a *propensity* fogalmat nélkülözhetetlennek tartotta. Egy ilyen cinkelt és szabályos kockából összevegyített dobássorozat – a homogén dobássorozatokhoz hasonlóan – éppúgy rendelkezhet hajlammal bizonyos frekvenciamintázatok realizációjára. Az állandóságot Popper és követői csak azért tekinthették a karakterizáció szükséges elemének, mert a *propensity*-re szinguláris értelemben tekintettek, vagyis olyan hajlamként, amely futamonként jellemzi a kísérletet. Ha a kísérleti elrendezés tulajdonságai változnak, akkor nyilván nincs okunk azt gondolni, hogy az egymást követő kísérleti futamokban a *propensity* értéke ugyanaz marad. Mint alább megmutatjuk azonban, ezt feltenni akkor sem lesz több okunk, ha a kísérleti elrendezés fizikai tulajdonságai állandóak.

Az állandóság tehát nem szükséges elem. A *propensity*-t hordozó kísérleti elrendezés jellemzéséhez elegendő valamilyen fizikai leírás. Mindazonáltal egy ilyen fizikai leírás léte még nem elegendő ahhoz, hogy a leírás egy *chance setup*-ot határozzon meg. Ha a kockát minden alkalommal a hatossal felfelé egyszerűen leteszem az asztalra, akkor ez a kísérleti elrendezés csak triviális értelemben rendelkezik azzal a hajlammal, hogy kimenetként hatost adjon. Hasonlóan nem számítana hazárdjátéknak a kockázás akkor sem, ha a kockadobás kimenetei nem függenének ilyen érzékenyen a kezdeti kísérleti feltételektől, vagyis ha egy átlagos képességű kockajátékosnak nem jelentene nagyobb nehézséget hatost dobni, mint, mondjuk, a szemetet a szemétkosárba hajítani. Vagyis ahhoz, hogy valami *chance-setup*-nak minősüljön, a kísérleti elrendezés specifikációján túl még az is kell, hogy a *nem* specifikált fizikai jellemzők *egyenletes eloszlást* mutassanak. A túl alacsonyról eldobott kockadobásokat éppen ilyen megfontolásból nem tekintjük „jó” kockadobásnak. Hogy egy adott kísérleti szituációban ezeknek a „rejtett paramétereknek” az eloszlása egyenletes-e, az természetesen empirikus kérdés.

Végül megjegyezzük, hogy ha a *propensity*-t nem az ismételhető, hanem az egyszeri kísérleti elrendezés hajlamának tartjuk, hogy egy bizonyos kimenetet realizáljon, akkor a kísérleti elrendezés karakterizációja nem jelent problémát: egy szinguláris kísérleti elrendezést összes tulajdonságának adott időbeli értéke jellemez, legyenek ezek a tulajdonságok állandók vagy változók. Ez azonban – mint azt rövidesen látjuk – nem jelenti azt, hogy a *propensity*-nek az egyedi kísérleti elrendezéshez rendelése jó döntésnek számítana.

Most térjünk át a második kérdésre adott válaszok értékelésére: mit tekintsünk a *propensity* manifesztációjának? Kezdjük a hosszútávú *propensity*-értelmezéssel.

A hosszútávú *propensity*-értelmezés szerint a *propensity* egy diszpozíciós tulajdonság, amely adott frekvenciamintázatban manifesztálódik. A szabályos kockával való dobálás hosszú távon egy hatod arányban hatost ad. Először is fontos tisztán látnunk, hogy a *propensity* diszpozíciós karakterének hangsúlyozása nem emeli őt ki a többi fizikai mennyiség közül. Ha akarjuk, éppenséggel fogalmazhatunk úgy, hogy egy test tömege az a diszpozíciós tulajdonsága, hogy ha adott erővel hatunk rá, akkor adott módon gyorsul. A testnek ezt a diszpozícióját ezek után rendelhetjük a magához a testhez, vagy a testhez és a mérőberendezéshez együttesen. Ez természetesen megfelel a *propensity*-interpretáció eredeti intenciójának, amely szerint a *propensity* éppen olyan „rendes” fizikai mennyiség, mint a tömeg vagy az erő. A fizikai tulajdonságokat tehát elemezhetjük szemantikailag tetszés szerint akár kategorikus, akár diszpozicionális predikátumokkal, azaz mondjuk kontrafaktuálisan elemezhető predikátumokkal. Vagy tovább mehetünk, és ontológiai különbséget is vonhatunk a kategorikus és diszpozicionális tulajdonságok között: a lényeg, hogy a *propensity* és a tömeg minden esetben ugyanarra az oldalra fog esni.

Ekkor azonban feleslegessé válik annak a szempontnak a hangsúlyozása, ami a *propen-*

sity-interpretációt elválasztja a frekvenciainterpretációtól: az tudniillik, hogy a valószínűség nem maga a relatív frekvencia, hanem a frekvenciákat létrehozó kísérleti elrendezés diszpozíciója. Ez olyan, mintha azt állítanánk, hogy a csésze tömege nem a 20 dkg maga, hanem a csésze vagy a csésze és a mérleg együttes diszpozíciója arra, hogy ha a csészét megmérem, akkor mérési eredményül 20 dkg-ot kapok. Mi a különbség a két állítás között? Mi mást értenénk azon, hogy egy csésze 20 dkg, mint azt, hogy a mérlegre téve ennyit mutat? Mennyivel állítunk többet a diszpozíciós leírással annál, mint ami a szokásos kifejezésben is benne foglaltatik? Hasonlóképpen ha azt mondjuk, hogy a valószínűség a kísérleti elrendezés hajlama arra, hogy egy bizonyos frekvenciát hozzon létre, akkor semmivel sem mondunk többet annál, mint amikor magát a frekvenciát közöljük. Hiszen honnan is tudnánk, hogy a frekvencia minek is a frekvenciája, ha nem határoztuk meg előre azokat a kísérleti feltételeket, amelyek az adott frekvenciát elválasztják az azonos számértékeket viselő egyéb frekvenciáktól? Amint ugyanis nagy különbség azt mondani, hogy egy *valamilyen* mérés eredménye 20, és azt mondani, hogy egy *tömegmérés* eredménye 20 dkg, úgy nagy különbség azt mondani, hogy egy eseménysorozatban kiválogatott események relatív frekvenciája egy hatod, és azt mondani, hogy a kockadobások esetében a hatos dobások frekvenciája egy hatod. Frekvencia-interpretációnak nyilvánvalóan csak az utóbbi minősül, az előbbi pusztán egy aritmetikai közlés.

Poppert nyilván az zavarhatta a frekvenciainterpretációban, hogy az mindenféle *ad hoc* eseményekből képzett kompendiumoknak valószínűséget kénytelen tulajdonítani, ugyanakkor úgy vélte, hogy a *propensity*-értelmezés képes a valószínűségeket szabályozottan egy *chance setup* által generált homogén eseménytípusra, pontosabban magára a *chance setupra* vonatkoztatni. Ha ugyanis összevegyítek mindenféle szedett-vedett fizikai eseményt – érvelhet a *propensity*-hívő –, egy hatos dobást kockával, egy puskalövést agyaggalambra, egy lottóötöst stb., és az eseményeket csak a „talált-nem talált” szempontjából jellemzem, akkor frekventistaként kénytelen leszek a sorozatban a találatokból számolt relatív gyakoriságokat valószínűségeknek interpretálni, holott világos, hogy az előző felsorolás mögött nem adható meg egy olyan *chance setup*, amely a talált-nem talált szóban forgó frekvenciáját hajlamosítaná. Ez pedig arra mutat rá, hogy a fenti frekvenciák nem is *igazi* valószínűségek.

Ez azonban tévedés. A fenti felsorolás, a hatos dobást, a puskalövést stb. ugyanolyan kísérleti elrendezést karakterizál, mint egy ismétlődő kockadobás; és ha tetszik, ennek a kísérleti elrendezésnek is tulajdoníthatok valamilyen találati frekvenciát realizáló hajlamot. Ez a kísérleti elrendezés természetesen összetettebb és ennél fogva szokatlanabb, mint a szokásos véletlengenerátorok. De ez még nem jelenti azt, hogy az általa létrehozott események frekvenciája kevésbé kötődne a kísérleti elrendezéshez, mint egy „homogén” kísérleti elrendezésben. Az pedig, hogy a különböző fajtájú események között a „talált-nem talált” szempontjából ekvivalenciaosztályokat hozok létre, semmiben sem különbözik attól, ahogy a kockadobásokat osztályozhatom a „páros-páratlan” szempontjából. Amint ez az eset is mutatja, a frekvenciák megadása tehát nem lehetséges a fizikai szituáció pontos körvonalazása nélkül. De ha a frekvenciák megadása egyúttal a fizikai szituáció megadását is magában foglalja, akkor mi újat is nyújt a *propensity*-interpretáció a frekvenciainterpretációhoz képest? A tulajdonságok diszpozicionális karakterének hangsúlyozása legfeljebb a frekvenciainterpretáció bennfoglalt feltételeinek kiemelését jelentheti.

Most azonban térjünk át a szinguláris értelmezésre. A szinguláris interpretációt a következő kérdés motiválja: Mit jelent a hosszútávú *propensity*? Hogyan lehetséges – akárcsak metafizikailag is –, hogy egy diszpozíció aktív legyen hosszú távon anélkül, hogy az egyedi esetekben aktív volna? Hogyan képes hosszú távú stabil frekvenciákat létrehozni, ha nem az egyedi esetek „kényszerítése" révén? Vagyis hogyan lehetséges hosszútávú *propensity*, hogy az egyben ne volna szinguláris is?

A szingularista elképzelés szerint tehát a frekvenciamintázat nem egy hosszútávú hajlam manifesztációja, hanem szinguláris *propensity*-k valamilyen iterációjának eredménye. A szinguláris *propensity* az általános vagy egyedi kísérleti elrendezés azon hajlama, hogy egy individuális kísérleti futamban egy adott kimenet realizálódjon. A hatos dobás *propensity*-ja a kocka azon hajlama, hogy *most* hatost dobjak. De mit is jelent ez azon túl, hogy ebben a kísérletben az adott kimeneti esemény néha bekövetkezik, néha nem? Vagy blikkfangosabban: a kísérleti elrendezésnek hajlama van arra, hogy néha realizáljon egy kimenetet, néha pedig ne. A *propensity*-elmélet híve azonban ennél többet akar mondani. Azt, hogy ezeknek a lehetséges eseteknek mértéke, súlya van, amelyek ráadásul additívek is. De vajon mi garantálja ezt? Miért is kellene a *propensity*-knek kielégíteni a kolmogorovi axiómákat?

Vessünk egy pillantást a többi interpretációra. Láttuk, hogy mind a szubjektív, mind a frekvenciainterpretáció esetében baj volt az interpretációk megengedhetőségével: egyik interpretáció valószínűségfogalma sem volt σ-additív. A *propensity*-interpretáció esetében azonban ennél sokkal rosszabb a helyzet: egyáltalán nincs kritériumunk annak eldöntésére, hogy a *propensity*-k matematikai értelemben valószínűségek-e vagy sem. Vagyis nincs egy másik struktúránk, mint a hitek, amelyekről a fogadások operacionalizációja után feltételezzük, vagy a frekvenciák, amelyekről belátjuk, hogy additív struktúrát mutatnak. A *propensity*-k additivitása puszta kikötés, és ezt az a priori kikötést nehéz volna mással indokolni, mint azzal a tautológiával, hogy ahhoz, hogy a *propensity*-k a valószínűség megengedhető interpretációját nyújtsák, maguknak is valószínűségeknek kell lenniük matematikai értelemben (ld. Hitchcock 2002).

Ha már a kolmogorovitásnál tartunk, érdemes azon is eltöprengeni, hogy ha a *propensity*-értelmezés híveinek jelentős része a *propensity* bevezetését a kvantumelmélet valószínűségi értelmezésének nehézségével indokolja, és a *par excellence propensity*-knek a kvantumvalószínűségeket tekintik, akkor miért nem rögtön egy nem-additív valószínűségi mezőt vezetnek be a *propensity*-kre? Miért kellenek egyáltalán a kolmogorovi axiómák?[9]

Röviden szólva, abból a tényből, hogy egy kísérleti elrendezésben egy kimenet néha létrejön, nem következik az az erős metafizikai posztulátum, hogy az elrendezésnek egy a valószínűség axiómáit kielégítő hajlama volna. De talán a posztulátumot majd a frekvenciák

[9]Mivel a kérdéses szempontból semmivel sem jobbak, ezért itt most eltekintünk azoknak a *propensity*-elméleteknek az ismertetésétől, amelyek egy nem kolmogorovi struktúrára épülnek. (Ld. pl. Suppes 1973)

igazolják! Térjünk át a harmadik kérdésre: milyen értelemben magyarázza a *propensity* a manifesztálódó frekvenciákat?

Mint láttuk, a hosszútávú *propensity*-értelmezés magyarázó sémája egyszerű: a frekvencia azért jön létre, mert a kísérleti elrendezésnek van egy frekvenciát létrehozó hajlama. A szinguláris *propensity*-interpretáció egyik motivációja éppen az, hogy az efféle tautologikus magyarázatokat elkerülje. A feladata tehát szinguláris *propensity* és a manifesztálódó frekvenciák viszonyának tisztázása. Mint fentebb láttuk, ennek egyik módja a nagy számok segítségül hívása. A gondolatmenet lépései a következők: (i) Posztulálunk egy adott kísérleti elrendezéshez tartozó szinguláris *propensity*-t. (ii) Az adott kísérleti elrendezés futamait nagyon sokszor megismételjük, és feltételezzük, hogy az egymást követő *propensity*-k azonosak és függetlenek. (iii) A szinguláris *propensity*-kat egy valószínűségi mértékkel, a hosszútávú *propensity*-t, a függetlenségnek megfelelően, egy szorzatmértékkel azonosítjuk. (iv) A nagy számok valamelyik törvényére hivatkozva belátjuk, hogy a hosszútávú *propensity* valószínűségi értelemben tart a frekvenciákhoz. Így végül beláthatjuk, hangzik az érv, hogy a hosszútávú *propensity* posztulálása felesleges, mivel a nagy számok törvényének segítségével a szinguláris *propensity*-ből levezethető.

Ez az érvelés azonban ezer sebből vérzik. Haladjunk a fenti premisszák szerint. Első lépésben tehát posztulálunk egy szinguláris *propensity*-t, amellyel egy kísérleti elrendezés egy adott kimenet produkál. Metafizikailag ez a szinguláris hajlam Giere-nél mint elágazó lehetséges világok súlyozott halmaza jelenik meg, máshol mint tovább nem elemezhető, primitív kauzális tendencia. Mint fentebb már megvizsgáltuk, annak feltevése, hogy ezek a súlyok egy additív normált mértéket követnek, súlyos metafizikai posztulátum.

De vegyük a (ii) premisszát! Mivel igazoljuk azt, hogy azonos kísérleti elrendezésben a szinguláris *propensity*-k azonosak és függetlenek? Láttuk, hogy a kísérleti elrendezés állandóságának megkövetelése mögött az az elképzelés húzódik, hogy így talán a *propensity*-k azonossága is biztosítható. De miért volna ez így? Még ha a kísérleti szituációt karakterizáló jegyek időben változatlanok is, akkor is miért kellene a szinguláris *propensity*-nek éppen ezektől az állandó tulajdonságoktól függenie, és miért nem éppen a leírásban nem specifikált változó körülményektől? Vagyis miért nem függhet a hatos dobás *propensity*-je mondjuk éppen az eldobás sebességétől, ami egyik dobásról a másikra változik? Vagy még tovább menve: miért kell *egyáltalán* bármilyen egyéb tulajdonságtól függenie a *propensity*-nek? Amíg semmi egyebet nem tudunk erről a *propensity*-ről, mint hogy a kimenetek bekövetkezésére való hajlam, addig természetesen a fenti kérdésekre nem tudunk válaszolni. De akkor azt sem mondhatjuk, hogy a *propensity*-k azonosak volnának.

A függetlenséggel hasonló a helyzet. Az egymást követő kísérleti futamokban a *propensity*-k függetlenségének feltételét vagy éppen tagadását semmi sem motiválja. Az erők mechanikai függetlenségének hipotézisére hivatkozni itt fölöttébb megtévesztő. Az ugyanis, hogy mondjuk a Coulomb-erőre igaz a szuperpozíció, kísérletileg könnyen alátámasztható: az eredő erő okozta gyorsulás a külön-külön vett erők gyorsulásainak vektori összege. A *propensity*-k függetlenségét azonban semmilyen módon nem tudjuk empirikusan verifikálni.

A *propensity*-interpretáció képviselői ezen a ponton rendszerint a kvantumelméletre hivatkoznak, mondván, a kvantumelmélet *no-go*-tételei megmutatták, hogy a mikrorendszereket jellemző hullámfüggvénynek nincs rejtett paraméteres modellje, vagyis a rendszer teljes állapotleírásából sem következik, hogy egy adott mérésben melyik kimenet fog megvalósulni. Az azonban, hogy nincs objektív kritériuma annak, hogy melyik kimenet valósul meg, azt jelenti, hogy a rendszer egy objektív hajlammal rendelkezik arra nézve, hogy az bizonyos kimeneteket realizáljon – és ez a *propensity*.

Hogy a kvantumelmélet no-go-tételei implikálják-e az objektív indeterminizmust, azt most nem tárgyaljuk (ld. Szabó 2002). Azonban még ha az objektív indeterminizmus igaz is, a szinguláris *propensity* bevezetése akkor sem indokolt. Ha egy kísérleti elrendezés két egymást követő futamában nem léteznek olyan rejtett paraméterek, amelyek egyértelműen meghatároznák, hogy az első futamban miért ezt a kimenetet kaptuk, és a másodikban miért azt – hát nem léteznek. De az még nem jelenti azt, hogy a rendszernek volna valamilyen objektív hajlama azon a triviális értelmen túl, hogy egyszer ez történt, egyszer pedig amaz. Hát még, hogy ez a hajlam a két esetben azonos és független volna!

A nagy számokra épülő fenti levezetésnek az már csak külön pikantériája, hogy ha az első három premisszát elfogadnánk is, vagyis léteznének a kolmogorovi axiómákat kielégítők, az egymást követő kísérleti futamokban azonos és egymástól független szinguláris *propensity*-k, a kívánt eredmény akkor sem volna elérhető. A nagy számok törvényei ugyanis csak annyit állítanak, hogy a frekvenciák és a szinguláris *propensity valószínűség erejéig* megegyezik vagy egymáshoz tart. De hogyan is interpretáljuk *propensity*-ként ez a végtelenszeres szorzattérben vett valószínűségi mértéket? Miféle eseménynek a szinguláris *propensity*-jéről van itt szó? Minek a hajlamáról mire vonatkozóan?

Jól látható hát, hogy a szinguláris *propensity*-elképzelés teljességgel tarthatatlan. Mielőtt azonban végleg felhagynánk fele, befejezésképpen hadd tegyünk pár megjegyzést a *propensity*-irodalomban tárgyalt némely problémát illetően.

1. *A propensity mint diszpozíció.* Mindenekelőtt szeretnénk hangsúlyozni, hogy a *propensity*-interpretációval szembeni megfogalmazott fenti kritikánkban sehol sem vontuk kétségbe a diszpozíciók létét *általánosságban*, érveink kizárólag a *propensity* mint speciális diszpozíció ellen irányultak. A diszpozíciókkal szembeni általános érvek természetesen a *propensity*-interpretáció számára is kihívást jelentenek, sőt, némelyik itt válik különösen élessé. Tekintsük elsőként a realizmus-antirealizmus kérdést. Mint láttuk, a diszpozícióknak, és így a *propensity*-nek is ontológiailag két értelmezése van: a realista és az antirealista. A realista álláspont szerint a *propensity*-k mint diszpozíciók nem épülnek rá az aktuális világ egyéb kategorikus tulajdonságaira, az antirealista álláspont szerint azonban igen. Mindkét álláspontnak megvannak maga problémái, amelyek a *propensity* kapcsán fokozottan jelentkeznek.

A realista makacsul ragaszkodhat ahhoz, hogy a *propensity* feltételezése nem mond ellent a hume-i szupervencia empirista elvének, lévén, hogy a szinguláris *propensity*-k is a világ lokális mintázatába tartoznak: a kockadobásnak mind a hatos kimenete, mind a p

propensity-je az aktuális világ lokális tulajdonsága. A világ egyéb tényei pedig a kategorikus tulajdonságokra *és* a *propensity*-kre együttesen épülnek rá. A kérdés azonban éppen az, hogy a világnak mely tényei azok, amelyek ontológiai leírásához a *propensity*-kre is szükségünk volna. Félő, hogy magukon a *propensity*-ken kívül nincsenek ilyen tények.

Másfelől ha a *propensity*-kkel szemben antirealisták vagyunk, vagyis azt valljuk, hogy az aktuális világban egyéb kategorikus tulajdonságok is betölthetik a *propensity* funkcióját, akkor a kérdés az lesz, hogy mi is az a funkció, amit ezek az egyéb tulajdonságok betöltenek, hacsak nem az, hogy matematikai valószínűségként viselkednek. Mint láttuk, a valószínűségi kalkulus respektálása önmagában is kétséges, de még ha teljesül is, ezt leszámítva nem tudjuk megadni a *propensity*-nek egyetlen olyan jellemvonását sem, amelyet az illető tulajdonságnak mutatnia kellene ahhoz, hogy őt a *propensity* aktuális betöltőjének nevezzünk.

A jobb megértés végett hasonlítsuk össze a *propensity* fogalmát a fájdalom fogalmával az elmefilozófiában. Ha a fájdalomról mint mentális tulajdonságról azt valljuk, hogy az csupán egy bizonyos agyállapot, akkor ennek igazolásához úgy kezdünk hozzá, hogy először a fájdalmat valamiként *meghatározzuk*, pl. megadjuk a viselkedésben betöltött oksági szerepét – hogy például elkerülő magatartást vált ki – illetve a többi mentális állapotok között betöltött oksági szerepét. Majd miután a fájdalmat eképpen meghatároztuk, megmutatjuk, hogy ezt a szerepet bizonyos agyállapotok képesek betölteni. A *propensity* esetében azonban éppen ez a meghatározás az, ami hiányzik. Minek is kellene teljesülnie egy kategorikus tulajdonságra nézve ahhoz, hogy az a *propensity*-szerepet betöltse?

2. *Finkish propensity.* A valószínűség és a lehetőség fogalmának mély fogalmi összetartozása egy további problémát generál a *propensity* diszpocionális, vagy legalábbis a kontrafaktuális elemzése számára (ld. Eagle 2004). Amint azt a diszpozíciókat taglaló résznél láttuk, a *finkish* illetve mímelő diszpozíciók léte kérdésessé teszik a diszpozíciók kontrafaktuális elemzését. A *finkish propensity* kontrafaktuális elemzése azonban a valószínűség és a lehetőség szoros fogalmi kapcsolata miatt egyenesen lehetetlen. Ha ugyanis egy esemény valószínűsége nem nulla, akkor az eseményt lehetségesnek tartjuk. Mármost ha egy nem nulla valószínűségű, azaz *propensity*-jű esemény bekövetkezését a *propensity* manifesztációjának feltételei megakadályozhatják, akkor az esemény lehetetlenné válik, bár *propensity*-je nem nulla. Az érv a többi diszpozíciókra nem vonatkozik, mivel azok nem ápolnak ilyen szoros kapcsolatot a lehetőség fogalmával.

Erre a *propensity*-hívő még mindig mondhatja, hogy a *propensity* kontrafaktuális elemzése elhibázott vállalkozás, és Levi (1980) valóban ezt is mondja: „minden olyan kísérlet, amely diszpozicionális predikátumokat kontrafaktuálisok segítségével kíván elemezni, a szekeret fogja a ló elé" (248. o.). A kontrafaktuálisokat épp a diszpozíciók léte és természete igazolja. Ekkor azonban meg kellene mondani, hogy miben is áll a diszpozíciók természete, ha nem a kontrafaktuálisokban. Mint láttuk, éppen ez az, amire a *propensity*-interpretáció képtelen.

3. *Humphreys-paradoxon.* Végül hadd tegyünk egy megjegyzést egy, az irodalomban a

propensity-interpretáció ellen felhozott kritikával kapcsolatban (ld. Szabó 2002). Szokás a *propensity*-elméletet avval a kritikával elutasítani, hogy a valószínűségként felfogott *propensity* nem fejez ki *kauzális* hajlamot, és mivel a *propensity* egyfajta parciális kauzális tendencia, ezért a valószínűségnek nem lehetséges *propensity*-interpretációja. Mint az alábbiakban megmutatjuk, ez az érv nem helytálló.

Ha a *propensity* parciális kauzális hajlam, és a valószínűség *propensity*, akkor minden valószínűségi összefüggést kauzálisan is értelmezni kell tudnunk, jelesül a feltételes valószínűségeket is. Hogy ez nem lehetséges, arra a standardnak tekintett ellenpélda Humphreys-től (1985) származik: Legyen a az az esemény, hogy egy kísérleti elrendezésben az elektron áthaladt egy féligáteresztő tükrön, b pedig az az esemény, hogy ezek után egy detektorba csapódik. A $p(b|a)$ feltételes valószínűség jelentése a *propensity*-interpretáció szerint az elektronnak az a hajlama, hogy amennyiben áthaladt a tükrön, megérkezzen a detektorba. Mi a *propensity*-értelmezése azonban a $p(a|b)$ fordított feltételes valószínűségnek? Amennyiben a valószínűségek nem nullák, ez utóbbi valószínűség kifejezhető a speciális Bayes-tétel segítségével a

$$p(a|b) \quad = \quad \frac{p(b|a)p(a)}{p(b|a)p(a) + p(b|a^{\perp})p(a^{\perp})}$$

formában, vagyis $p(a|b)$ a valószínűségi kalkulusban értelmezhető kifejezés. A kauzális *propensity*-interpretáció a $p(a|b)$-t azonban kénytelen úgy értelmezni – hangzik az érv –, mint az elektronnak azon hajlamát, hogy amennyiben megérkezett a detektorba, *előtte* átmenjen a tükrön. Vagyis a Bayes-tétel értelmezése feltételezi a retrokauzalitást.[10] A *propensity*-t tehát nem lehet anélkül azonosítani a kauzális szereppel, hogy ne ütköznénk ellentmondásba a valószínűségi kalkulussal.

Az érv azonban azon a feltételezésen múlik, hogy amennyiben a $p(b)$ valószínűséget úgy értelmezzük, mint az elektronnak azon kauzális hajlamát, hogy becsapódjon a detektorba, akkor a $p(b|a)$ feltételes valószínűséget úgy kell értelmeznünk, mint az elektronnak azon kauzális hajlamát, hogy becsapódjon a detektorba, *feltéve*, hogy átment a tükrön. A feltételes valószínűségnek ez az értelmezése azonban teljesen indokolatlan. A $p(b|a)$ feltételes valószínűséget egyszerűen a $\frac{p(b \wedge a)}{p(a)}$ hányados definiálja, ami, még ha a valószínűségeket *propensity*-nek értelmezzük is, akkor is csak úgy érthető, mint az elektron azon kauzális hajlamainak a *hányadosa*, hogy egyfelől becsapódjon a detektorba, *és* átmenjen a tükrön, másfelől egyszerűen átmenjen a tükrön. Ennek pedig egyáltalán nem kell megegyeznie az érvben használt feltételes kauzális hajlammal.[11] Vagyis a Humphreys-paradoxon nem jelent ellenérvet a valószínűségnek kauzális hajlamként való értelmezésével szemben.

[10] A Humphreys-paradoxonnak léteznek olyan feloldási kísérletei, amelyek a $p(a|b)$ valószínűséget a részecske az a és a b esemény bekövetkezése *előtti* hajlamának tekinti, hogy amennyiben megérkezett a detektorba, előtte átmenjen a tükrön. Ez a közös ok típusú megoldás mindazonáltal sok szempontból nem meggyőző.

[11] A feltételes valószínűség tipikus félreértéseiről ld. (Szabó 2002, 87. o.) valamint (Hajek 2003).

Összefoglalva tehát a valószínűség *propensity*-interpretációja jócskán elmarad a többi értelmezés mögött, mivel a *propensity* természetét illetően egy sereg alapvető kérdésre nem képes választ adni. Giere (1976a) igyekszik elhitetni velünk, hogy a *propensity*-interpretáció elleni kifogások mondhatni esztétikaiak:

> „A *propensity*-interpretáció elleni fő kifogás nem az, hogy homályos vagy hogy üres, hanem hogy metafizikailag túl extravagáns. Nemcsak azt állítja, hogy a természetben vannak fizikai lehetőségek, hanem azt is, hogy a természet tartalmaz innát tendenciákat ezen lehetőségek felé, tendenciákat, amelyeknek logikai struktúrája valószínűségi jellegű." (348. o.)

A problémák gyökerénél azonban nem a *propensity*-értelmezés metafizikai extravaganciája áll, hanem az az orvosolhatatlan hiányosság, hogy a *propensity*-interpretáció egészen egyszerűen elmulasztja meghatározni akár fizikailag, akár metafizikailag a *propensity* fogalmát, amellyel azután a valószínűséget azonosítaná.[12] Úgy beszél egy fizikainak posztulált mennyiség mértékéről, hogy – a triviális hosszútávú értelmezést leszámítva – nem mondja meg, hogy mi is ez a mennyiség. Mindezek alapján a *propensity*-interpretációt illetően kénytelenek vagyunk egyetérteni Kyburg (1974) szarkasztikus megfogalmazásával, amely szerint „a *propensity*-elmélet többek között azért vonzó, mivel megengedi – egyenesen *propensity*-vel hívja elő – a vad metafizikai spekulációkat egy olyan kontextusban, amelyben az ember felszabadítva érzi magát mindenféle kényszer alól, hogy átgondolt metafizikai érvekkel szolgáljon". (365. o.)

Arra a kérdésre tehát, hogy miért is van a hatos dobásnak adott valószínűsége, a *propensity*-interpretáció válasza semmivel sem különb, mint Moliére *Képzelt betegének* az ópium altató hatásának okát firtató kérdésére adott válasza: azért, mert a kocka „aleatorikus erővel" rendelkezik.

[12]Ráadásul a fizikai és metafizikai meghatározás az interpretációkban rendszeresen összekeveredik. Nem állíthatjuk ugyanis egyszerre azt is, hogy a *propensity* egy fizikai hipotézis, és ugyanakkor azt is, hogy az a valószínűség fogalmának szemantikai vagy metafizikai elemzése. Ha a *propensity* egy fizikai hipotézis, mint Popper állította, akkor a *propensity* és a valószínűség azonossága kontingens tény, ha ellenben a *propensity* a valószínűség fogalmának szemantikai elemzése, akkor azt kell megmutatni, hogy minden lehetséges világban a valószínűséget reprezentáló akármicsodák *propensity*-k. Mint láttuk, egyik feladat sem biztatóbb a másiknál.

Összegzés

Könyvünkben arra a kérdésre kerestük a választ, hogy mit is értünk valószínűség alatt. Amint azt a 3. fejezet bevezetőjében kifejtettük, a kérdés megválaszolása egy olyan fizikai elmélet felállítását jelenti, amely empirikus fogalmakban definiálja a valószínűséget, és amely valószínűségterminusa igazodik a valószínűség fogalmának előzetes köznyelvi használatához. Ezzel a magas igénnyel közeledve a valószínűség paradigmatikus interpretációi felé, meg kell állapítanunk, hogy azok nem állták ki a próbát. Nézzük őket sorban!

A *klasszikus interpretáció* arra a kérdésre, hogy mit jelent az, hogy szabályos kockával dobva a hatos valószínűsége egy hatod, az „egyenlően lehetséges" fogalmára hivatkozva válaszol: mivel szimmetrikus kocka esetén mindegyik oldal előfordulása egyenlően lehetséges, és az esetek közül nekünk csak az egyik kedvez, ezért a kedvező esetek és az egyenlően lehetséges esetek számának aránya egy hatod lesz. Az „egyenlően lehetséges" fogalma azonban az a pont is egyben, ahol az interpretáció csődöt mond: az „egyenlően lehetséges" ugyanis sem objektív, sem szubjektív értelemben nem értelmes. Ha „egyenlően lehetséges" alatt ugyanis azt értjük, hogy a dobás mindegyik kimenetelének megegyezik a „modális súlya", akkor ezzel átlépjük a verifikálhatóság határát, és a valószínűséget egy fizikailag nem értelmezett fogalomra vezetjük vissza. Ha az „egyenlően lehetséges"-t úgy értelmezzük, hogy a hat kimenet összes fizikai tulajdonsága megegyezik, akkor nyilván többet állítunk, mint amennyit szeretnénk, hiszen ebben az esetben a Leibniz-elv értelmében nem is tudjuk majd megkülönböztetni a kimeneteket. Ha pedig válogatni kezdünk a tulajdonságok között mondván, hogy csak a releváns tulajdonságoknak kell megegyezniük, akkor a releváns tulajdonságok kiválasztásánál nem tudunk majd másra támaszkodni, mint arra, hogy mely tulajdonságok megegyezése szolgáltat azonos relatív gyakoriságokat. Ha az „egyenlően lehetséges"-t szubjektíven értjük, vagyis úgy mint ami az empirikus ágens számára megkülönböztethetetlen, akkor szintén vállalnunk kell a következményt, hogy az így kapott valószínűségfogalom esetleg nem kvadrál majd a relatív gyakorisággal. Mi több, amint a Mounty Hall-paradoxon mutatja, bonyolultabb esetekben az empirikus ágens számára a kombinatorikai lehetőségek számbavétele is komoly feladatnak bizonyul.

A klasszikus interpretációval kapcsolatban rámutattunk továbbá arra is, hogy az irodalomban keringő matematikai ellenpéldák, úgy mint a Bertrand-paradoxon nem jelentenek cáfolatot az interpretáció tarthatóságával szemben, mivel a kérdés itt nem az, hogy az „egyenlően lehetséges" fogalma értelmezhető-e matematikailag, hanem, hogy fizikailag értelmezhető-e.

De nem értelmezhető-e az „egyenlően lehetséges” logikailag? A *logikai interpretációt* éppen az a törekvés hívta életre, hogy szigorúbb alapokra helyezze az „egyenlően lehetséges” jelentését. A logikai interpretáció szerint a hatos dobás valószínűsége azért egy hatod, mert az a kijelentés, hogy a kockát eldobtuk, egy hatod mértékben konfirmálja a hatos dobást kifejező kijelentést egy mindkét kijelentést tartalmazó közös nyelvben. Az „egyenlően lehetséges” helyébe tehát a „konfirmáció” varázsszava lép. Ezzel a lépéssel azonban csak tovább nőnek a nehézségeink. Mit jelent ugyanis az, hogy egy eseményre vonatkozó kijelentés adott mértékben konfirmál egy másik eseményre vonatkozó kijelentést? Hogyan szerzünk tudomást a konfirmáció mértékéről? Itt két válasz lehetséges. Az első szerint a konfirmáció fizikai viszony, amely megfigyeléssel megállapítható. De mi más alapján is dönthetnénk erről a viszonyról, mint ismét csak a megfigyelt relatív gyakoriságok alapján. Ekkor azonban a konfirmáció közbülső fogalma felesleges lesz, a valószínűség értelmezését kezdhetjük nyugodtan a relatív gyakoriságokkal is. A másik válasz Carnapé: a konfirmáció egy, a következtetéshez hasonló logikai viszony két kijelentés között. Ha viszont a valószínűséget a konfirmációnak erre a logikailag értett fogalmára vezetjük vissza, akkor egyben arról is lemondunk, hogy fizikailag interpretáljuk a valószínűséget. Így hát a logikai interpretáció sem állja ki a próbát.

Akkor talán fogadjuk el, hogy a valószínűség nem a priori viszony, hanem események bekövetkezésébe vetett parciális hit. Ez a *szubjektív interpretáció.* Az, hogy a hatos dobás valószínűsége egy hatod, azt jelenti, hogy egy hatod mértékben hiszünk a kijelentés igazságában. Az pedig, hogy ilyen mértékben hiszünk benne, abban jut kifejezésre, hogy 1 : 5 arányban vagyunk hajlandóak fogadni a hatos bekövetkezésére. Itt ismét csak arra kényszerülünk, hogy különbséget tegyünk a racionális és az empirikus ágens között. Ha a fenti mondat a racionális ágensre vonatkozik – ahogy az igen gyakran történik –, akkor ez visszatérést jelent a logikai interpretációhoz – legalábbis mindaddig, ameddig a racionalitást nem karakterizáljuk empirikus fogalmakban. Ha a racionalitást *Dutch book*-mentességgel vagy a várható haszon maximalizálásával jellemezzük, akkor meg kell mutatnunk, hogy ezek a jellemzők valóban fennállnak az empirikus döntési szituációkban. Az Allais-döntés illetve a hasonló pszichológiai kísérletek azonban egyáltalán nem igazolják ezt a várakozásunkat.

Megtehetjük persze, hogy a valószínűséget nem a racionális, hanem közvetlenül az empirikus ágens hitének mértékével azonosítjuk. Ekkor természetesen búcsút kell mondanunk mindazoknak az elegáns matematikai eredményeknek (Ramsey–de Finetti-tétel, stb.), amelyek a szubjektív interpretáció fő vonzerejét adják. A valószínűség az empirikus ágens hitének mértéke, akár racionális, *Dutch book*-mentes ez a hit, akár nem. Egy probléma azonban még ekkor is fellép: honnan tudjuk, hogy a fogadásokban manifesztálódó mértékek valóban a szubjektum intenzionális hitállapotait tükrözik? A kísérleti pszichológia rámutatva a fogadásnak a hitre visszaható torzító erejével mindenesetre e nézet ellen szól. Azonosítsuk tehát egyszerűen a valószínűséget az aktuális fogadási arányokkal? Ha így járunk el, akkor messze sodródunk az eredeti ambiciózus tervektől.

De talán a valószínűség objektív fogalom – relatív gyakoriság vagy *propensity.* Kezdjük az utóbbival. A hatos dobásnak azért egy hatod a valószínűsége – vallja a *propensity-interpretáció* –, mert a kocka fizikai környezetével együtt rendelkezik azzal a hajlammal, hogy egy elegendően hosszú dobássorozatban a dobások közel egy hatoda hatos legyen. De honnan tudnunk arról, hogy a kocka rendelkezik ezzel a hajlammal, ha nem éppen onnan, hogy „egy elegendően hosszú dobássorozatban a dobások közel egy hatoda hatos"? Miért kellene bevezetnünk egy új diszpozícionális tulajdonságot annak érdekében, hogy magyarázni tudjunk egy adott jelenséget, ha ennek a tulajdonságnak semmi egyéb empirikus ismertetőjegye nincsen, mint az adott jelenség? És akkor még nem beszéltünk a *propensity*-interpretációknak arról a metafizikailag igen terhelt típusáról, amely a hajlamot nem egy frekvenciamintázat létrehozására vonatkoztatja, hanem egy egyedi eseményére. A szinguláris *propensity*-interpretációk a nagy számok törvényeit segítségül hívó minden alátámasztási kísérlet ellenére nyilvánvalóan tarthatatlanok.

Egy interpretáció maradt hátra: a *relatív gyakoriság-interpretáció*. Amint látjuk, az összes többi interpretáció is efelé az interpretáció felé gravitál: az „egyenlő lehetőség", a konfirmáció, a hit mértéke, ha empirikusan értjük, végül mind relatív gyakoriságokra fut vissza. Mi több, a relatív gyakoriság-interpretáció az uralkodó nézet természettudományos körökben is. A frekventista interpretáció szerint a hatos dobás egy hatod valószínűsége semmi mást nem jelent, mint hogy a hatos relatív gyakorisága a kockadobások egy elegendő hosszú sorozatában közel egy hatod lesz. Az ördög azonban a részletekben lakik. Mit értsünk ugyanis azon, hogy „a hatos relatív gyakorisága a kockadobások egy elegendő hosszú sorozatában közel egy hatod"? Az aktuális dobások egy véges sorozatáról vagy egy ideális végtelen dobássorozatról van szó? Az előbbi nyilván túl szigorúan operacionalizálja a valószínűség fogalmát, az utóbbi pedig verifikációtranszcendens: egy aktuális, véges dobássorozat nem határozza meg annak a végtelen sorozatnak a relatív gyakoriságait, sőt még ezek létezését sem, amelynek kezdőszelete. Ráadásul, ha a sorozattól a véletlenszerűséget is megköveteljük von Mises módjára, akkor a randomitás meghatározásával külön nehézségeink támadnak.

És akkor még nem is említettük az egyes interpretációknak a Kolmogorov-féle valószínűségszámításhoz való viszonyának kérdését. Modellje-e egy adott elmélet a standard valószínűségszámításnak? Ha nem, melyik az a tulajdonság (pl. σ-additivitás), amelyik nem teljesül? Mi az oka annak, hogy az adott tulajdonság az elméletben nem teljesül, stb.?

Ha nem sikerül magyarázni a valószínűséget egyetlen interpretációval, hátha sikerül egyszerre többel. Az egyes interpretációk hibáit elkerülendő néhány szerző az interpretációkból egyfajta *patchwork*-öt varr. Jó példa erre Lewis valószínűséginterpretációja, amely kerülő úton, az ún. *legjobb rendszer analízisen* keresztül jut el a valószínűséghez. A gondolatmenet középpontjában a *hume-i szuperveniencia* fogalma áll: „A hume-i szuperveniencia... az a tan, hogy a világot lokális, partikuláris tények hatalmas mozaikja alkotja, egyik kis dolog a másik után... Amink van, az tulajdonságok elrendeződése. És ez minden. Nem létezik különbség tulajdonságok elrendeződésében álló különbség nélkül. Minden egyéb ezen szuperveniál." (Lewis 1986, ix-x. o.) Ha minden egyéb, akkor a valószínűség is. Kérdés, hogy mi az a lokális minta, amelyen a valószínűség szuperveniál. Lewis objektivista, így csak két jelölt merül fel: a szimmetria és a relatív gyakoriság. Az elsőt Lewis az indifferencia elvének gyengéi miatt elutasítja. A meglepő azonban az, hogy a második jelölttel is ezt teszi, jóllehet a frekvencia kézenfekvő jelölt volna egy hume-iánusnak. A frekvenciával szemben felhozott ellenérvek inkább ellenérzések; ahogy Lewis fogalmaz: az objektív valószínűség fogalmát „szeretném valami általánosabb alatt látni." Ez az általános pedig Lewis (1994) legjobb rendszer analízise. Nézzük a természettörvények legjobb rendszer analízisét egyelőre a valószínűség fogalma nélkül:

> „Vegyük az összes deduktív rendszert, amelynek tételei igazak. Némely egyszerűbb lesz, rendszerezettebb, mint a többi. Mások erősebbek, informatívabbak a többinél. Ezek az értékek konkurálnak: egy nem informatív rendszer nagyon egyszerű lehet, egy vegyes információkból összehordott rendezetlen kompendium nagyon informatív. A legjobb rendszer az, amelyik a legjobban egyensúlyoz egyszerűség és erő között megőrizve az igazságot... Egy szabályszerűség akkor és csak akkor törvény, ha a legjobb rendszer tétele." (478. o.)

És most tegyük bele az analízisbe az objektív valószínűség fogalmát is.

> „Módosítjuk a legjobb rendszer analízist, hogy egy csomagban adjanak számot az objektív valószínűségekről, és az őket irányító törvényekről. Tekintsünk deduktív rendszereket, amelyek nemcsak arra vonatkoznak, ami történik, hanem különböző szituációkban különböző kimenetelek valószínűségére is, például különféle izotópok bomlási valószínűségeire. Követeljük meg ezeknek a rendszereknek az igazságát arra vonatkozóan, amit a történésekre mondanak. De a valószínűségekre vonatkozóan nem követelhetjük meg az igazságukat, mert még meg kell mondanunk, hogy mit jelent a valószínűség; rendszereink még nincsenek teljesen interpretálva... Mint előbb, néhány rendszer egyszerűbb lesz, mint a többi. Majdnem mint előbb, néhány erősebb lesz a többinél: néhány megmondja, hogy mi fog történni, vagy hogy mik lesznek az objektív valószínűségek bizonyos szituációkban, míg mások hallgatnak a kimenetekről is, és a valószínűségekről is. Továbbá, néhány jobban illeszkedik a történések aktuális menetéhez, mint mások. Vagyis, a történések menetének valószínűsége nagyobb lesz néhány rendszer szerint, mint mások szerint... Az egyszerűség, az erő és az illeszkedés egyensúlyoznak. A legjobb rendszer az, ahol e három egyensúlya a legjobb. Mint előbb, a törvények azok a regularitások, amelyek a legjobb rendszer tételei. De most néhány törvény valószínűségi. Így hát a valószínűséget eképpen elemezzük: valószínűség az, amit a legjobb rendszer valószínűségi törvényei annak mondanak." (480. o.)

Hogy mi tehát a valószínűség, azt a törvények teljes rendszere mondja meg figyelembe véve az egyszerűség, az erő és az illeszkedés konkuráló szempontjait. A valószínűség legjobb rendszer analízise a klasszikus és logikai interpretáció szimmetriája, és a frekventista interpretáció relatív gyakorisága között egyensúlyoz. Ezzel azonban mindkét interpretáció hibáit a nyakába veszi.

Amint látható tehát, sem a standard interpretációk, sem pedig azok kombinációi nem hoztak megnyugtató választ a valószínűség empirikus alkalmazhatóságát és fogalmi természetét kutató kérdésre. Mindegyik interpretáció megragad valamit a valószínűség sokarcú fogalmából, ugyanakkor egyik sem képes teljes mértékben számot adni a fogalom minden aspektusáról.

A kudarcok dacára az élet megy tovább, és a valószínűség értelmezésére egyre újabb kísérletek születnek. Mivel az öt standard interpretáció kitölti azt a logikai teret, amelyet a valószínűségi interpretációk egyáltalán elfoglalhatnak, azért az új interpretációs próbálkozások mind valamelyik interpretációhoz kapcsolódnak – azok szellemében érvelnek, azok problémáira kívánnak megoldást találni. Hogy csak egy-egy példát említsünk: a klasszikus és logikai interpretációra mennek vissza azok a próbálkozások, amelyek a *maximális entrópia elvére* hivatkozva igyekeznek meghatározni egy eseménytéren a természetes valószínűségi mértéket (Barta and Jones, 2000). A szubjektív interpretáció szellemében fogantak Schervish, Seidenfeld és Kadane (2000) kutatásai, amelyek az inkoherencia, vagyis a kolmogorovi elmélettől való eltérés mértékét tesztelik. A kolmogorov komplexitás (Kolmogorov 1965), és a randomitás egyéb elméletei (Martin-Löf, 1966) nyilvánvalóan von Mises frekventista elméletének továbbfejlesztései, ahogyan (Suppes, 1973) a *propensity*-interpretáció nem-klasszikus általánosításása.

Amíg tehát a valószínűségszámítás axiomatikája lezárt fejezetnek számít a matematikatörténetben, addig a valószínűség konceptuális elemzése továbbra is a modern analitikus filozófia egyik legizgalmasabb és legkihívóbb kérdése marad.

Függelékek

A. A Boole-algebra

Az alábbiakban bevezetjük a legfontosabb hálóelméleti fogalmakat, és definiáljuk a Boole-algebra fogalmát.

18. Definíció. Egy $(S, \leq)$ párt *részben rendezett halmaznak nevezzük*, ha S egy nem üres halmaz, $\leq$ pedig egy reláció S-en az alábbi tulajdonságokkal. Minden a, b és $c \in S$-re

(i) $a \leq a$ (reflexív);

(ii) ha $a \leq b$ és $b \leq a$, akkor $a = b$ (antiszimmetrikus);

(iii) ha $a \leq b$ és $b \leq c$, akkor $a \leq c$ (tranzitív).

Részben rendezett halmazra jó példa egy halmaz hatványhalmaza a $\subseteq$-relációval vagy a természetes számok halmaza az oszthatósági relációval. A valós számok halmaza a $\leqslant$-relációval teljesen (lineárisan) rendezett halmazt alkot, vagyis olyan halmazt, ahol a fentieken kívül még az is teljesül, hogy bármely két $a, b \in S$-re vagy $a \leq b$ vagy $b \leq a$.

Legyen T az S egy részhalmaza. S egy $sup(T)$ elemét T *legkisebb felső korlátjának* nevezünk, ha minden $a \in T$-re $a \leq sup(T)$, és minden $b \in S$-re, amelyekre $a \leq b$ minden $a \in T$-re, fennáll, hogy $sup(T) \leq b$. Hasonlóképpen, S egy $inf(T)$ elemét T *legnagyobb alsó korlátjának* nevezünk, ha minden $a \in T$-re $inf(T) \leq a$, és minden $c \in S$-re, amelyekre $c \leq a$ minden $a \in T$-re, fennáll, hogy $c \leq inf(T)$. Mivel a $\leq$ reláció antiszimmetrikus, $sup(T)$ és $inf(T)$, ha létezik, akkor egyértelműen létezik. A $sup(T)$ és az $inf(T)$ segítségével a háló fogalma a következőképpen definiálható.

19. Definíció. Egy részben rendezett L halmazt *hálónak* nevezünk, ha bármely két a és b eleméhez létezik $sup\{a, b\}$ és $inf\{a, b\}$.

A háló fogalmát azonban egy másik úton is bevezethetjük, amely nem a részben rendezésre, hanem absztrakt műveletekre támaszkodik.

20. Definíció. Egy L halmazt *hálónak* nevezünk, ha adva van rajta egy $\wedge : L \times L \to L$ metszet illetve egy $\vee : L \times L \to L$ unió művelet úgy, hogy minden a, b és $c \in L$-re az alábbi tulajdonságok fennállnak:

$$\left.\begin{array}{r} a \wedge b = b \wedge a \\ a \vee b = b \vee a \end{array}\right\} \quad \text{kommutativitás}$$

$$\left.\begin{array}{r} (a \vee b) \vee c = a \vee (b \vee c) \\ (a \wedge b) \wedge c = a \wedge (b \wedge c) \end{array}\right\} \quad \text{asszociativitás}$$

$$\left.\begin{array}{r} a \wedge a = a \\ a \vee a = a \end{array}\right\} \quad \text{idempotencia}$$

$$\left.\begin{array}{r} a \wedge (a \vee b) = a \\ a \vee (a \wedge b) = a \end{array}\right\} \quad \text{elnyelési tulajdonság}$$

A háló két definíciója az alábbi módon tehető ekvivalenssé. Az első definícióról a másodikra a

$$\begin{array}{rcl} a \wedge b & \equiv & inf\{a, b\} \\ a \vee b & \equiv & sup\{a, b\} \end{array}$$

azonosításokkal térhetünk át, míg visszafelé a

$$a \leq b \text{ akkor és csak akkor, ha } a \wedge b = a$$

kikötéssel.

Egy hálót *teljesnek* nevezünk, ha nemcsak minden elempárhoz, hanem L minden T részhalmazához létezik $sup(T)$ és $inf(T)$. Ha $sup(T)$ és $inf(T)$ létezik minden megszámlálható részhalmazhoz, akkor a hálót *σ-teljesnek* mondjuk. Egy hálót *null- és egységelemesnek* nevezünk, ha létezik olyan $0, 1 \in L$ elem, hogy minden $a \in L$-re teljesül, hogy

$$0 \leq a \leq 1$$

Ez a tulajdonság véges hálón mindig teljesül.

Vegyük észre, hogy a háló definiáló tulajdonságai között nem szerepel a disztributivitás; s valóban ez a tulajdonság az elágazási pont a klasszikus és a nem-klasszikus eseményterek között.[13] Egy háló akkor *disztributív*, ha minden a, b és $c \in L$ elemre teljesülnek az alábbi tulajdonságok:

$$\begin{array}{rcl} (a \wedge b) \vee c & = & (a \vee c) \wedge (b \vee c) \\ (a \vee b) \wedge c & = & (a \wedge c) \vee (b \wedge c) \end{array}$$

A komplementer esemény megragadásához bevezetjük a következő műveletet.

[13] A nem-klasszikus valószínűségelmélethez lásd (Rédei, Summers, 2007).

21. Definíció. Egy null- és egységelemes L hálón értelmezett $^\perp : L \to L$ műveletet *ortokomplementációnak* nevezünk, ha az alábbi műveletek minden $a, b \in L$ elemre teljesül teljesülnek:

$$
\begin{array}{rl}
(a^\perp)^\perp = a & \text{involúció} \\
a \leq b \text{ akkor és csak akkor, ha } b^\perp \leq a^\perp & \text{kontrapozíció} \\
a \wedge a^\perp = 0 & \text{ellentmondás elve} \\
a \vee a^\perp = 1 & \text{kizárt harmadik elve}
\end{array}
$$

Egy ortokomplementációval ellátott hálót *ortokomplementáris hálónak* nevezünk. Ortokomplementáris hálón mindig teljesülnek a *de Morgan-azonosságok*:

$$
\begin{aligned}
(a \wedge b)^\perp &= a^\perp \vee b^\perp \\
(a \vee b)^\perp &= a^\perp \wedge b^\perp
\end{aligned}
$$

Ezzel elérkeztünk a Boole-algebra definíciójához.

22. Definíció. Egy $\Sigma = (L, \wedge, \vee, ^\perp, 0, 1)$ disztributív, null- és egységelemes, ortokomplementáris hálót *Boole-algebrának* nevezünk.

A Boole-algebrák absztrakt struktúrája könnyen lefordítható a halmazok nyelvére. Egy Ω halmaz $\mathcal{P}(\Omega)$ *hatványhalmazának* nevezzük Ω összes részhalmazának halmazát; $\mathcal{P}(\Omega)$ egy részhalmazát pedig *halmazalgebrának*, ha az tartalmazza magát Ω-t, valamint minden a elem $\overline{a} \equiv \Omega \setminus a$ komplementerét, és zárt véges unióra nézve. Stone reprezentációs tétele alapján azonban minden absztrakt Boole-algebra izomorf valamilyen halmazalgebrával. Így minden Boole-algebra halmazalgebrára, a rajta értelmezett absztrakt műveletek $(\wedge, \vee, ^\perp)$ pedig az ismert halmazelméleti metszet, unió és kompementer $(\cap, \cup, ^-)$ műveletekre redukálhatók.

Ha egy Boole-algebrában a $\wedge$ és $\vee$ műveletek zártak minden megszámlálhatóan végtelen részhalmazra, akkor a Σ-t *Boole-σ-algebrának* vagy röviden *σ-algebrának* nevezzük. Másfelől, egy Ω halmaz részhalmazainak olyan Σ halmazalgebráját, amely zárt megszámlálhatóan végtelen részhalmazra vett halmazelméleti metszet és unió műveletekre nézve, szintén σ-algebrának nevezzük. Megjegyezzük, hogy Stone tétele nem terjeszthető ki σ-algebrákra, így az absztrakt σ-algebrák nem képezhetők le mind halmazelméleti σ-algebrákra.

A Boole-algebrák osztályozására szolgálnak az alábbi fontos definíciók:

23. Definíció. (i) Egy $b \in L$ hálóelemről azt mondjuk, hogy *lefedi* $a \in L$-t, ha $a < b$ (vagyis $a \leq b$, de $a \neq b$), és minden $c \in L$-re, amelyre $a \leq c \leq b$, fennáll, hogy $c = a$ vagy $c = b$. Más szóval egy elem akkor fed le egy másikat, ha rögtön utána következik a rendezésben.

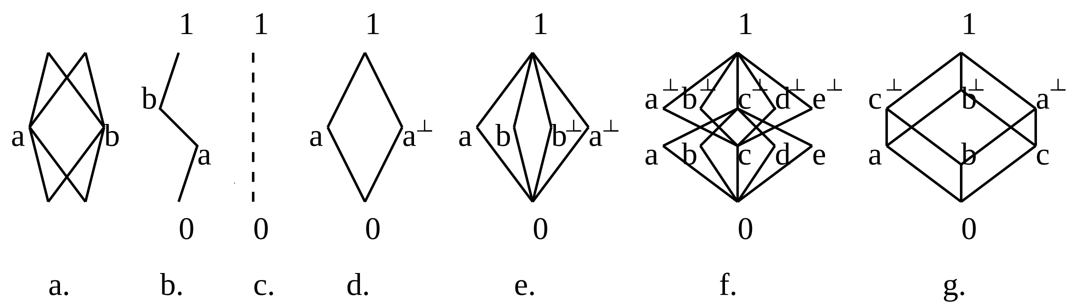
(ii) A 0-t lefedő elemeket a háló *atomjainak*, a háló azon elemeit, amelyeket az 1 lefed, *koatomoknak* nevezünk.

(iii) Egy L hálót *atomosnak* nevezünk, ha minden nem nulla $b \in L$-re létezik egy a atom úgy, hogy $a \leq b$, vagyis ha minden elem alatt létezik atom.

(iv) Egy L hálót *atomisztikusnak* nevezünk, ha minden nem nulla $b \in L$ elemre $b = \vee_i a_i$, ahol a_i atom minden i-re. Más szóval, ha a háló minden nem nulla eleme előáll mint atomok uniója.

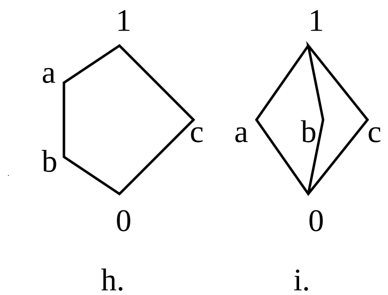
(v) Végül L elemeinek egy $\{b_i\}_{i=1}^n$ részhalmazát a háló egy *partíciójának* nevezzük, ha $b_i \wedge b_j = 0$ minden $i \neq j$-re és $\vee_{i=1}^n b_i = 1$.

A hálók szemléltetésének kényelmes módja a Hasse-diagram (8.1. ábra). A háló elemeit a diagramok csúcsai jelölik, a részben rendezést az élek. Ha két elemet él köt össze, akkor a diagramon feljebb elhelyezkedő elem lefedi a lentebbit. Egy adott elemtől lefelé az élek mentén haladva megkapjuk az elemnél kisebb elemeket, felfelé haladva pedig a nagyobbakat. Az így kapott rendezésben a közvetlenül a 0 fölött álló elemek az atomok, a közvetlenül az 1 alatt állók a koatomok. Két adott elemnél kisebb elemek közül a legnagyobb a két elem metszete, a két elemnél nagyobb elemek közül a legkisebb pedig a két elem uniója. A háló definíciójánál fogva mind a metszet elem, mind az únió elem egyértelműen létezik. Az $^\perp$ ortokomplementációt a diagramon külön jelöljük.



8.1. ábra. Hasse-diagramok

Tekintsük a 8.1. ábra diagramjait. *a.* diagramja részben rendezett halmaz, de nem háló, mert a $sup\{a, b\}$ és az $inf\{a, b\}$ nem létezik. A *b.* diagram null- és egységelemes, atomos, de nem atomisztikus háló, mivel b nem áll elő atomok uniójaként; ugyanakkor nem Boole-algebra, mivel nincs (és nem is lehet) rajta ortokomplementáció. A *c.* diagram egy folytonos háló (ezt jelöli a szaggatott vonal), elemei a valós számok a $[0, 1]$ intervallomon, a rendezés pedig a szokásos $\leqslant$-reláció. A diagram egy nem atomos (és így nem is atomisztikus) Boole-algebra. A *d.* diagram két atomos (és atomisztikus) Boole-algebra; az *e.* diagram az ún. kínai lámpa, amely egy nem disztibutív, ortokomplementáris háló. Az *f.* diagram a G_{12}-nek nevezett szintén nem disztributív háló; végül a *g.* diagram pedig három atomos (és atomisztikus) Boole-algebra. A Stone-tétel értelmében a *c.* Boole-algebra a $\mathcal{P}(\{a\})$, a *d.* Boole-algebra a $\mathcal{P}(\{a, b\})$, a *g.* Boole-algebra pedig a $\mathcal{P}(\{a, b, c\})$ halmazalgebrával izomorf.



8.2. ábra. Az N_5 és M_5 hálók Hasse-diagramja

Hálók disztributivitásának (és egyben a nem-klasszikus eseménystruktúrák felbukkanásának) ellenőrzése a Hasse-diagramok alapján az alábbi szemléletes módon lehetséges. Egy L háló akkor és csak akkor disztributív, ha a 8.2. ábrán látható N_5 és M_5 hálók egyike sem ágyazható be L-be (ha L-nek nincs N_5-tel vagy M_5-tel izomorf részhálója). Így például a fenti *e.* és *f.* diagramok azért nem disztributív hálók, mert az *e.* diagramba az M_5, az *f.* diagramba pedig az N_5 háló ágyazható be.

B. Példák valószínűségi mértékterekre

Az alábbiakban valószínűségi mértékterekre illetve valószínűségi változókra adunk meg példákat. A példák esetében a valószínűségszámítási kézikönyveknek azt a gyakorlatát követjük, hogy valószínűségi mértékterekre különféle fizikai szituációk segítségével hivatkozunk. Így például a $\Sigma = \mathcal{P}(\{f, i\})$ hatványhalmazon értelmezett uniform mértékre mint a szabályos pénzérmével való dobást reprezentáló mértéktérre fogunk hivatkozni. Ezt a kézikönyvekben didaktikai szempontból helyénvaló lépést egy, a valószínűség jelentésével foglalkozó munkában csak annyiban alkalmazhatjuk, amennyiben világos marad, hogy a fenti példák fizikai címkéjük ellenére interpretálatlan matematikai példák (ezért az idézőjel az elnevezésüknél): sem az eseménytér elemeiről nem tudjuk, hogy azok fizikailag mit jelentenek (pl. szinguláris események vagy eseménytípusok-e), sem a valószínűségről nem tudjuk, hogy az hogyan is van értelmezve. Az alábbi példákban tehát mind az „eseménytér”, mind a „valószínűség” kifejezés jelentése pusztán annyi, amennyi a valószínűségszámítás axiómáiból következik. Így például az *1.* példában az a kifejezés, hogy „az érme szabályos” pusztán annyit jelent, hogy a fej és az írás valószínűsége egyaránt $\frac{1}{2}$. Hogy mindez azután milyen viszonyban áll a kocka fizikai szimmetriáival, a fej relatív előfordulásával hosszú dobássorozatban, a fej bekövetkezésében való racionális hit mértékével stb. – mindez már a valószínűség interpretációs kérdéseihez tartozik.

1. „Szabályos pénzérme”. Egy szabályos pénzérmét feldobunk, és egy forintot kapunk, ha a dobás eredmény fej. Az eseménytér négy elemből áll: a „fej dobás” eseményéből, az „írás dobás” eseményéből, valamint a lehetetlen és a biztos eseményből. Az eseményteret tehát az alábbi módon reprezentálhatjuk. Legyen $\Omega = \{f, i\}$ kételemű halmaz, és legyen $\Sigma = \mathcal{P}(\{f, i\}) = \{\emptyset, \{f\}, \{i\}, \Omega\}$ az Ω hatványhalmaza. Legyen továbbá $p(\{f\}) = p(\{i\}) = \frac{1}{2}$, azaz p legyen uniform valószínűségi mérték. Még egyszer hangsúlyozzuk, hogy a „szabályos" kifejezés pusztán annyit jelent, hogy a p függvény értéke a fejre és az írásra is egyaránt $\frac{1}{2}$. A fej dobásért járó jutalmat a következő diszkrét (Bernoulli-) valószínűségi változó reprezentálja: $g : \{f, i\} \to \{0, 1\}$; $g(f) = 1$, $g(i) = 0$. Ekkor g várható értéke $Exp(g) = 1 \times \frac{1}{2} + 0 \times \frac{1}{2} = \frac{1}{2}$; szórása pedig $Var(g) = \sqrt{\frac{1}{2} - \frac{1}{4}} = \frac{1}{2}$ lesz.

2. „Két szabályos mágneses pénzérme". Most két szabályos érmével dobjunk. Tegyük fel, hogy valamilyen fizikai okból (pl. az érmék mágnesezve vannak) az azonos oldalú kimenetek, tehát a fej-fej és az írás-írás kétszer nagyobb valószínűségűek, mint a különböző oldalú, fej-írás illetve írás-fej kimenetek. (Hogy ez miben manifesztálódik: frekvenciában, hajlamban stb., az ismét csak a valószínűség interpretációja után válaszolható meg.) Jutalmul annyi forintot kapunk, ahány fejet dobtunk.

Az elemi események tere ekkor az $\Omega = \{(f, f), (i, f), (f, i), (i, i)\}$ négyelemű rendezett párokból álló halmaz, az összetett események halmaza pedig a $\Sigma = \mathcal{P}(\Omega)$ hatványhalmaz, amely a következő elemekből áll:

$$\Omega$$

$$\{(f,f),(i,f),(f,i)\}\,;\ \{(f,f),(i,f),(i,i)\}\,;\ \{(f,f),(f,i),(i,i)\}\,;\ \{(i,f),(f,i),(i,i)\}$$

$$\{(f,f),(i,f)\}\,;\ \{(f,f),(f,i)\}\,;\ \{(f,f),(i,i)\}\,;\ \{(i,f),(f,i)\}\,;\ \{(i,f),(i,i)\}\,;\ \{(f,i),(i,i)\}$$

$$\{(f,f)\}\,;\ \{(i,f)\}\,;\ \{(f,i)\}\,;\ \{(i,i)\}$$

$$\emptyset$$

A valószínűségi mérték Σ-n a következő: $p(\{(f, f)\}) = p(\{(i, i)\}) = \frac{1}{3}$, és $p(\{(i, f)\}) = p(\{(f, i)\}) = \frac{1}{6}$. Az $\{(f, f), (i, f), (i, i)\}$ összetett esemény valószínűsége ekkor például $\frac{5}{6}$.

A fejekért járó jutalmat a $g : \{(f, f), (i, f), (f, i), (i, i)\} \to \mathbb{N}$; $g((f, f)) = 2$, $g((f, i)) = g((i, f)) = 1$, $g((i, i)) = 0$ valószínűségi változó modellezi. Ekkor g várható értéke $Exp(g) = 2 \times \frac{1}{3} + 1 \times \frac{1}{6} + 1 \times \frac{1}{6} + 0 \times \frac{1}{3} = 1$; szórása pedig $Var(g) = \sqrt{\frac{5}{3} - 1} = \sqrt{\frac{2}{3}}$.

3. „Cinkelt kocka". Legyen most az eseménytér a kockadobások eseménytere, és tegyük fel, hogy a kocka úgy van cinkelve, hogy a hatoson kívül minden más oldal előfordulásának a valószínűsége 0. A valószínűségi változó mérje egyszerűen a kockadobás kimenetét. Vagyis az alaphalmaz $\Omega = \{1, 2, 3, 4, 5, 6\}$ hatelemű halmaz, $\Sigma = \mathcal{P}(\Omega)$ az Ω 64 elemű hatványhalmaza, és a mérték pontra koncentrált (Dirac-)mérték: $p(\{6\}) = 1$, $p(\{i \neq 6\}) = 0$.

A valószínűségi változó $f : \{1, 2, 3, 4, 5, 6\} \to \mathbb{R}$; $f(i) = i$, amelyre $Exp(f) = 1 \times 6 = 6$ és $Var(f) = \sqrt{36 - 36} = 0$, vagyis f szórásmentes.

4. „Egy 0 és 1 közötti szám véletlen választása" egy σ-algebrára példa. Legyen Ω a $[0, 1]$ zárt intervallum, legyen $\Sigma = \mathcal{L}([0, 1])$ a $[0, 1]$ intervallum Lebesgue-mérhető halmazainak σ-algebrája, a valószínűségi mérték pedig legyen a Lebesgue-mérték megszorítva $[0, 1]$-re.

Mérje a valószínűségi változónk az origótól való távolságot: $f : [0,1] \to \mathbb{R}$; $f(x) = x$. Ekkor $Exp(f) = \displaystyle\int_0^1 x\,dx = \frac{1}{2}$; $Var(f) = \sqrt{\frac{1}{3} - \frac{1}{4}} = \sqrt{\frac{1}{12}}$.

5. „Dobássorozat pénzérmével”. Legyen most $\Omega = \{0,1\}^{\mathbb{N}}$ a 0-kból és 1-kből álló végtelen hosszú sorozatok tere. Jelöljük $x(n)$-nel az x sorozat n hosszúságú kezdőszeletét. Legyen továbbá w egy véges hosszú szintén 0-kból és 1-kből álló sorozat, azaz egy bináris *szó* vagy *sztring*, és jelöljük w hosszát $|w|$-vel, $|w_0|$-val illetve $|w_1|$-gyel pedig a w-ben szereplő 0-k illetve 1-esek számát. Minden ilyen w meghatározza Ω-nak egy $[w]$ részhalmazát: azokat a végtelen sorozatokat, amelyeknek kezdőszelete éppen w, vagyis $[w] = \{x \in \Omega \mid x(|w|) = w\}$. Az ilyen részhalmazokat *cilinder halmazoknak* nevezzük. Legyen Σ a cilinder halmazokat tartalmazó legszűkebb σ-algebra.

Ezek után definiáljunk egy valószínűségi mértéket a cilinderhalmazokon a következő természetes módon: $p([w]) = q^{|w_1|}(1-q)^{|w_0|}$, ahol $q \in [0,1]$. Ez a cilindereken definiált mérték egyértelműen kiterjeszthető Σ-ra, és így (Ω, Σ, p) egy valószínűségi mértéktér lesz.

Ismeretes, hogy minden x végtelen bináris sorozat azonosítható a $[0,1]$ intervallum valamely r_x elemével az alábbi bináris kifejtés alapján:

$$r_x \equiv \sum_{n=1}^{\infty} \frac{1}{2^n}\, x_n.$$

Igaz továbbá az is, hogy ezzel az azonosítással a $[0,1]$ intervallumon értelmezett Lebesgue-mérték a cilinder halmazokon éppen azt a fenti mértéket fogja meghatározni, amelyre $q = \frac{1}{2}$.

Legyen tehát $q = \frac{1}{2}$, és legyen a valószínűségi változó $f : \Omega \to \mathbb{R}$ az alábbi módon értelmezve: $f(x) = 2^n$, ahol n a sorozatnak az a helye, ahol először áll 1-es. Ekkor könnyen látható, hogy $Exp(f) = \sum_{n=1}^{\infty} \frac{1}{2^n}\, 2^n = \infty$.

Interpretáljuk ezek után Ω-t úgy, mint egy pénzérmével való végtelen hosszú dobássorozat kimeneteit: jelentsék a sorozat egyesei azt, hogy az érmével fejet dobtunk, nullái pedig, hogy az érmével írást dobtunk. Ezek után az eseménytér atomi eseményei a végtelen hosszú dobássorozatok lesznek. Például az az atomi esemény, hogy minden dobásra fejet kaptunk, az azonosan 1 sorozatnak felel meg. Az összetett események között lesznek a véges hosszú dobássorozatok, például azt az összetett eseményt, hogy három dobásból az első dobásra fejet, a másodikra és a harmadikra írást kaptunk, az [100] cilinderhalmaz reprezentálja.

Ismét csak tudatosítanunk kell, hogy az így interpretált mérhető téren bevezetett $p([w]) = q^{|w_1|}(1-q)^{|w_0|}$ valószínűségi mérték még nincs interpretálva. Vagyis hiába választjuk q-t $\frac{1}{2}$-nek, és hívjuk az (Ω, Σ, p) valószínűségi mértékteret egy szabályos érmével való dobássorozatot reprezentáló struktúrának, egyelőre nem tudjuk, hogy mit értsünk itt „szabályos” alatt, minthogy nem tudjuk, hogy mit értsünk valószínűségen.

Értelmezni tudjuk azonban a példában szereplő $f : \Omega \to \mathbb{R}$ valószínűségi változót, amelyet az alábbi módon definiáltunk: $f(x) = 2^n$, ahol n a sorozatnak az a helye, ahol először áll 1-es. Ez a valószínűségi változó a *pétervári paradoxon* szerint reprezentálhatja például egy bankkal folytatott játékban azt az összeget, amelyet a bankos egy érmével való dobássorozat alapján kifizet: ha elsőre fejet dob, akkor 2 aranyat, ha csak a másodikra, akkor 4-et, és így tovább a kettőhatványoknak megfelelően. Bernoullinak arra a kérdésére azonban, hogy mennyi legyen egy ilyen játék ára, azaz várható értéke, a szokásos $Exp(f) = \sum_{n=1}^{\infty} \frac{1}{2^n} 2^n = \infty$ válasz egyelőre pusztán egy matematikai kijelentés, mivel a várható értékhez szükséges valószínűség még nincs értelmezve.

6. „Véletlenszerűen dobálunk egy céltáblára”. Végül utolsó példánkban az alaphalmaz legyen a sík egységkörlapja: $\Omega = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 \leqslant 1\}$, Σ pedig Ω Lebesgue-mérhető halmazainak σ-algebrája, vagyis $\Sigma = \{a \cap \Omega \,|\, a \in \mathcal{L}(\mathbb{R}^2)\}$. p legyen az egységkörre normált Lebesgue-mérték. Interpretáljuk ezt a mérhető teret úgy, mint egy egységsugarú céltáblára való dobálás eseményterét. Az atomi események közé olyan események tartoznak, mint „eltalálni a céltábla középpontját”, az összetett események közé pedig olyanok, mint „az alsó félkörbe dobni”. Mérje ezek után a valószínűségi változónk az origótól való távolságot: $f : \Omega \to \mathbb{R}$; $f\big((x, y)\big) = \sqrt{x^2 + y^2} \equiv r$. Ekkor f várható értéke: $Exp(f) = \frac{1}{\pi} \displaystyle\int_0^{2\pi} \int_0^1 r\, r dr d\phi = \frac{2}{3}$, szórása pedig $Var(f) = \sqrt{\frac{1}{2} - \frac{4}{9}} = \sqrt{\frac{5}{18}}$. A valószínűségi mértéktér ismét csak félig interpretált: a p valószínűségnek egyelőre nincs jelentése. Így aztán semmi sem jogosít fel bennünket arra, hogy a Lebesgue-mértékből származtatott mértéket „a véletlenszerű dobálás modelljének” tekintsük.

C. *Dutch book*-tételek

Az alábbiakban három *Dutch book*-tételt bizonyítunk, amelyek központi szerepet játszanak a bayesianizmusban. A bizonyítások egyszerűsége végett mindhárom esetben feltesszük, hogy a hálók atomisztikusak. Hangsúlyozzuk azonban, hogy a tételek e nélkül a megszorítás nélkül is érvényesek.

Az első tétel a hagyományos Ramsey–de Finetti-tétel.

5. Tétel. Egy $q : \Sigma \to \mathbb{R}$ kvóciensfüggvény akkor és csak akkor konzisztens, vagyis nem adható ellene *Dutch book*, ha valószínűségi mérték.

Bizonyítás. Először megmutatjuk, hogy ha a kvóciensfüggvény sérti a valószínűségi mértékre előírtakat, akkor nem konzisztens. Kezdjük a

$$q(\Omega) = 1$$

kikötéssel. Tegyük fel X kvóciensfüggvényéről, hogy $q(\Omega) = q \neq 1$. Ekkor tétfüggvénynek az $S(\Omega) = \mathrm{sgn}(q-1)$, $S(b \neq \Omega) = 0$ hozzárendelést választva Σ-nak bármely Ω-t tartalmazó F részhalmazára minden interpretációban az össznyeremény $G^I(F) = -|q| < 0$ lesz, vagyis Y minden alkalommal $|q|$ összeget nyer X-től.

Most tegyük fel, hogy a kvóciensfüggvény sérti a σ-additivitást:

$$q(\vee_i a_i) = \sum_i q(a_i), \text{ ha } a_i \wedge a_j = \emptyset \text{ minden } i \neq j\text{-re.}$$

Vagyis tegyük fel, hogy $r \equiv q(\vee_i a_i) - \sum_i q(a_i) \neq 0$ az események valamely diszjunkt, megszámlálható $\{a_i\}$ halmazára. Legyen $F = \{\vee_i a_i, a_1, a_2 \ldots\}$. Ha az Y játékos tétfüggvénynek a

$$\begin{aligned} S(\vee_i a_i) &= \operatorname{sgn}(r) \\ S(a_i) &= -\operatorname{sgn}(r) \\ S(b) &= 0 \end{aligned}$$

hozzárendelést választja, ahol b valamely minden a_i-től és $\vee_i a_i$-től is különböző eleme Σ-nak, akkor könnyen látható, hogy minden interpretációban $G^I(F) = -|r| < 0$; vagyis – bármi történik – Y nyer $|r|$ összeget.

Fordított irányban azt kell megmutatnunk, hogy ha q valószínűségi mérték, akkor Y bárhogyan választ egy $F \subseteq \Sigma$ halmazt, azon nem tud a q kvóciensfüggvény ellen *Dutch book*ot adni.

Először is használjuk ki, hogy hálóink atomisztikusak. Atomos hálókon az interpretációk kölcsönösen egyértelműen megfeleltethetők az atomoknak: minden interpretáció csak egy atomon vesz fel 1-et, a többi atomhoz 0-t rendel. Jelöljük az I interpretációnak megfelelő atomot a^I-vel. Ezek után tekintsük Σ atomi partícióját, $\{a^I\}$-t, ahol is a partíció minden eleme kölcsönösen megfelel egy interpretációnak. Nyilvánvalóan minden I-re $q(a^I) \geqslant 0$, valamint

$$\sum_I q(a^I) = 1. \tag{C.1}$$

Az is nyilvánvaló, hogy

$$\sum_I q(a^I)(I(a) - q(a)) = 0 \tag{C.2}$$

minden $a \in F$-re, mivel az olyan interpretációk járuléka, amelyekre $I(a) = 1$, egyenlő $q(a)(1 - q(a))$-val, az olyan interpretációké pedig, amelyekre $I(a) = 0$, egyenlő $(1 - q(a))(-q(a))$-val; ezek pedig éppen kiejtik egymást.[14] A $q(a^I)$-k segítségével képezzük a nyeremények konvex kombinációját:

$$\begin{aligned} \sum_I q(a^I) G^I(F) &= \sum_I q(a^I) \left(\sum_{a \in F} (I(a) - q(a))\, S(a) \right) \\ &= \sum_{a \in F} \left(\sum_I q(a^I)\, (I(a) - q(a)) \right) S(a) \end{aligned}$$

[14] Itt használjuk ki, hogy a háló atomisztikus.

A konvex kombináció azonban (C.2) miatt 0, és így léteznie kell egy olyan I interpretációnak, amelyre $G^I(F) \geqslant 0$, vagyis q ellen nem adható *Dutch book* F-en. Mivel F tetszőleges lehet, ezért q konzisztens. □

A második *Dutch book*-tétel azt mutatja meg, hogy a feltételesen konzisztens kvóciensfüggvény miért lesz valószínűségi mérték.

6. Tétel. Egy $p : \Sigma \to \mathbb{R}$ valószínűségi mérték és egy $q(\cdot|\cdot) : \Sigma \times \Sigma \to \mathbb{R}$ feltételes kvóciensfüggvény akkor és csak akkor feltételesen konzisztens, ha minden $a, b \in \Sigma$ elemre ($p(b) \neq 0$) fennáll, hogy

$$q(a|b) \quad = \quad \frac{p(a \wedge b)}{p(b)}. \tag{C.3}$$

Bizonyítás. Kezdjük annak bizonyításával, hogy ha (C.3) nem teljesül, akkor p és $q(\cdot|\cdot)$ ellen adható *Dutch book.* Tegyük fel, hogy $r \equiv q(a|b) - \frac{p(a \wedge b)}{p(b)} \neq 0$ valamely $a, b \in \Sigma$ elemre, ahol $p(b) \neq 0$. Válassza ekkor az Y játékos F-nek az $F_1 = \{b^\perp, a \wedge b\}$ és az $F_2 = \{(a, b)\}$ halmaz unióját, tétfüggvénynek pedig az

$$\begin{aligned} S(b^\perp) &= -\operatorname{sgn}(r)\, q(a|b) \\ S(a \wedge b) &= -\operatorname{sgn}(r) \\ S(a|b) &= \operatorname{sgn}(r) \end{aligned}$$

hozzárendelést. Könnyen látható, hogy ezzel a választással az X játékos $G^I(F)$ nyereménye minden I interpretációban $-|r|\, p(b) < 0$, azaz p és q ellen adható (feltételes) *Dutch book.*

Visszafelé, azt kell megmutatnunk, hogy ha $q(\cdot|\cdot)$ olyan, hogy (C.3) teljesül minden $a, b \in \Sigma$ párra ($p(b) \neq 0$), akkor nem adható ellene *Dutch book.* A bizonyításhoz ismét tekintsük az $\{a^I\}$ atomi partíciót, és a $p(a^I)$-k segítségével ismét képezzük a nyeremények konvex kombinációját:

$$\begin{aligned} \sum_I p(a^I) G^I(F) \quad &= \quad \sum_I p(a^I) \left(\sum_{a \in F_1} (I(a) - p(a))\, S(a) \right) \\ &\quad + \sum_I p(a^I) \left(\sum_{(b,c) \in F_2} \big(I(b \wedge c) + (I(c^\perp) - 1)\, q(b|c)\big)\, S(b|c) \right) \\ &= \quad \sum_{a \in F_1} \left(\sum_I p(a^I)\, (I(a) - p(a)) \right) S(a) \\ &\quad + \sum_{(b,c) \in F_2} \left(\sum_I p(a^I) \big(I(b \wedge c) + (I(c^\perp) - 1)\, q(b|c)\big) \right) S(b|c) \end{aligned}$$

Könnyen meggyőződhetünk róla, hogy nemcsak (C.2) áll fenn ($q(a^I)$ helyébe $p(a^I)$-t írva), hanem a

$$\sum_I p(a^I)\big(I(b\wedge c) + (I(c^\perp) - 1)\, q(b|c)\big)$$

összeg is 0 minden $(b,c)\in F_2$ párra, mivel azon tagok járuléka, amelyekre $I(c)=0$, egyenlő 0-val; az $I(c)=1$ interpretációkon belül pedig az $I(b)=1$ és az $I(b)=0$ interpretációkhoz tartozó járulékok (C.3) miatt éppen kiejtik egymást, az első $p(b\wedge c)\,(1-q(b|c))$, a másik $p(b^\perp\wedge c)\,(-q(b|c))$ lévén. Így a konvex kombináció is 0. Ezért léteznie kell egy olyan I interpretációnak, amelyre $G^I(F)\geqslant 0$, vagyis p és $q(\cdot|\cdot)$ ellen nem adható feltételes *Dutch book* F-en. Mivel F tetszőleges lehet, ezért p és $q(\cdot|\cdot)$ feltételesen konzisztens, vagyis $q(\cdot|\cdot)$ feltételes valószínűség. □

Végül a harmadik, diakron *Dutch book*-tétel a bayesiánus kondicionálás alapja.

8. Tétel. p és p' valószínűségi mértékek akkor és csak akkor diakron konzisztensek (a 6.3. alfejezetben bevezetett diakron értelemben), ha minden $a\in\Sigma$ elemre fennáll, hogy

$$p'(a) \quad = \quad p(a|b). \tag{C.4}$$

Bizonyítás. Kezdjük ismét annak bizonyításával, hogy ha (C.4) nem teljesül, akkor p és p' ellen adható diakron *Dutch book*. Tegyük fel, hogy $r\equiv p'(a)-p(a|b)\neq 0$ valamely $a\in\Sigma$ elemre. Válassza ekkor az Y játékos H-nak az $F=\{a\wedge b, b, b^\perp\}$ és $F'=\{a\}$ halmazok unióját. Legyen a tétfüggvény továbbá

$$\begin{aligned} S(a\wedge b) &= -\operatorname{sgn}(r)\\ S(b) &= \operatorname{sgn}(r)\, r\\ S(b^\perp) &= -\operatorname{sgn}(r)\, p(a|b)\\ S'(a) &= \operatorname{sgn}(r) \end{aligned}$$

hozzárendelést. Könnyen látható, hogy ezzel a választással az X játékos $G^I(H)$ nyereménye minden I interpretációban $-|r|\,p(b)<0$, azaz p és p' ellen adható diakron *Dutch book*.

Visszafelé, azt kell megmutatnunk, hogy ha p és p' olyan, hogy (C.4) teljesül minden $a\in\Sigma$ eseményre, akkor nem adható ellene diakron *Dutch book*. A bizonyításhoz ismét tekintsük

Σ atomi partícióját, és vegyük a nyeremények konvex kombinációját a $p(a^I)$-k segítségével:

$$
\begin{aligned}
\sum_I p(a^I) G^I(F) &= \sum_I p(a^I) \left(\sum_{a \in F} (I(a) - p(a))\, S(a) \right) \\
&\quad + \sum_I p(a^I) \left(\sum_{a' \in F'} \big(I(a') - p'(a')\big) S'(a')\, I(b) \right) \\
&= \sum_{a \in F} \left(\sum_I p(a^I)\, (I(a) - p(a)) \right) S(a) \\
&\quad + \sum_{a' \in F'} \left(\sum_I p(a^I) \big(I(a') - p'(a')\big)\ I(b) \right) S'(a')
\end{aligned}
$$

Mivel p valószínűségi mérték, azért az $S(a)$-kat szorzó összeg (C.2) miatt nulla. A

$$\sum_I p(a^I) \big(I(a') - p'(a')\big)\ I(b)$$

összeg pedig $p(a' \wedge b) - p'(a')p(b)$-vel egyenlő, ami szintén 0, ha (C.4) fennáll. Így a konvex kombináció ismét csak 0, vagyis léteznie kell egy olyan I interpretációnak, amelyre $G^I(H) \geqslant 0$, azaz p és p' ellen nem adható diakron *Dutch book* H-en. Mivel H tetszőleges lehet, ezért p és p' diakron konzisztens. □

D. Bayesiánus tételek

A 2.3. fejezetben definiáltuk a feltételes valószínűség fogalmát. Az alábbiakban csokorba szedjük azokat a legfontosabb tételeket, amelyek e fogalomból következnek. Ha a és b események, akkor közöttük fennállnak az alábbi összefüggések:

$$
\begin{aligned}
p(a \vee b) &= p(a) + p(b) - p(a \wedge b) && \text{összegszabály} \\
p(a \wedge b) &= p(a|b)p(b) = p(b|a)p(a) && \text{szorzatszabály} \\
p(b) &= p(b \wedge a) + p(b \wedge a^{\perp}) && \text{partíciótétel} \\
p(b) &= p(b|a)p(a) + p(b|a^{\perp})p(a^{\perp}) && \text{teljes valószínűség tétele} \\
p(a|b) &= \frac{p(b|a)p(a)}{p(b|a)p(a) + p(b|a^{\perp})p(a^{\perp})} && \text{általános Bayes-tétel}
\end{aligned}
$$

Általános esetben az összeg- és a szorzatszabály a Boole-algebra $\{a_1, a_2 \dots a_l\}$ elemeire a következő alakot ölti:

$$\left.\begin{aligned} p(a_1 \vee a_2 \cdots \vee a_l) = \sum_{i=1}^{l} p(a_i) - \sum_{1=i<j}^{l} p(a_i \wedge a_j) \\ + \sum_{1=i<j<k}^{l} p(a_i \wedge a_j \wedge a_k) \cdots \pm p(a_1 \wedge a_2 \cdots \wedge a_l) \end{aligned}\right\} \quad \text{összegszabály}$$

$$\left.\begin{aligned} p(a_1 \wedge a_2 \cdots \wedge a_l) = p(a_1)p(a_2|a_1)p(a_3|a_1 \wedge a_2) \\ \dots p(a_l|a_1 \wedge a_2 \cdots \wedge a_{l-1}) \end{aligned}\right\} \quad \text{szorzatszabály}$$

A másik három összefüggés pedig a Boole-algebra egy b eleme és egy $\{a_i\}_{i=1}^{n}$ partíciója között a következőképpen fest:

$$p(b) = \sum_{i=1}^{n} p(b \wedge a_i) \qquad \text{partíciótétel}$$

$$p(b) = \sum_{i=1}^{n} p(b|a_i)p(a_i) \qquad \text{teljes valószínűség tétele}$$

$$p(a_j|b) = \frac{p(b|a_j)p(a_j)}{\sum_{i=1}^{n} p(b|a_i)p(a_i)} \qquad \text{általános Bayes-tétel}$$

A tételek bizonyítása elemi. A teljes valószínűség tétele a partíciótélelnek és a szorzatszabálynak közvetlen következménye; az általános Bayes-tétel pedig a Bayes-tételnek és a teljes valószínűség tételének.

A 2.3. illetve a 3.5. alfejezet gondolatmenetét folytatva a függetlenség illetve a feltételes függetlenség fogalmát valószínűségi változókra is kiterjeszthetjük. Egy $f : \Omega \to \mathbb{R}$ és egy $g : \Omega \to \mathbb{R}$ valószínűségi változót *függetlennek* nevezünk, ha az $f^{-1}(\mathcal{B}')$ és $g^{-1}(\mathcal{B}')$ halmazok függetlenek minden $\mathcal{B}' \subseteq \mathcal{B}(\mathbb{R})$-re. Kettőről több valószínűségi változóra áttérve, a $g_1, g_2 \cdots : \Omega \to \mathbb{R}$ valószínűségi változókat *páronként függetlennek* mondjuk, ha bármely kettől kiválasztva, azok függetlenek, és *teljesen függetlennek*, ha a $g_1^{-1}(\mathcal{B}'), g_2^{-1}(\mathcal{B}') \dots$ halmazok minden $\mathcal{B}' \subseteq \mathcal{B}(\mathbb{R})$-re függetlenek.

Együttes sűrűségfüggvénnyel rendelkező valószínűségi változók esetében a függetlenséget egyszerűen tudjuk kifejezni. Ehhez azonban be kell vezetnünk az együttes eloszlás, az együttes eloszlásfüggvény és az együttes sűrűségfüggvény fogalmát. Az (Ω, Σ, p) valószínűségi mértéktéren egy $g = (g_1, g_2, \dots g_n) : \Omega \to \mathbb{R}^n$ leképezést *valószínűségi vektorváltozónak* nevezünk, ha minden $a \in \mathcal{B}(\mathbb{R}^n)$ elemre $g^{-1}(a) \in \Sigma$. A vektorváltozó eloszlása a $p \circ g^{-1} : \mathcal{B}(\mathbb{R}^n) \to [0,1]$ leképezés, amelyet másképpen a $g_1, g_2, \dots g_n$ valószínűségi változók *együttes eloszlásának* is nevezünk. A g vektorváltozó *eloszlásfüggvényének*, vagy másrészt a

$g_1, g_2, \dots g_n$ valószínűségi változók *együttes eloszlásfüggvényének* az

$$F_g : \mathbb{R}^n \to \mathbb{R};\ F_g(x_1, x_2 \dots x_n) \equiv p \circ g^{-1}\big((-\infty, x_1) \times (-\infty, x_2) \times \dots \times (-\infty, x_n)\big)$$

függvényt nevezzük. Ha az F_g eloszlásfüggvény előáll az

$$F_g(x_1, x_2 \dots x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_n} f_g(x'_1, x'_2 \dots x'_n)\, dx_1\, dx'_2 \dots dx'_n$$

integrálként, ahol $f_g : \mathbb{R}^n \to \mathbb{R}$ egy nem negatív Lebesgue-integrálható függvény, akkor f_g-t a g vektorváltozó *sűrűségfüggvényének* vagy a $g_1, g_2, \dots g_n$ valószínűségi változók *együttes sűrűségfüggvényének* nevezzük.

Ezek után egy f_g együttes sűrűségfüggvénnyel rendelkező $g_1, g_2, \dots g_n$ valószínűségi változók esetében a teljes függetlenség egyszerűen annyit jelent, hogy az együttes sűrűségfüggvény szorzat alakban áll elő, azaz

$$f_g(x_1, x_2 \dots x_n) = f_{g_1}(x_1) f_{g_2}(x_2) \dots f_{g_n}(x_n)$$

minden $(x_1, x_2 \dots x_n) \in \mathbb{R}^n$-re.

Most térjünk át a feltételes függetlenség fogalmára. A $g_1, g_2, \dots g_n : \Omega \to \mathbb{R}$ valószínűségi változókat egy $h : \Omega \to \mathbb{R}$ valószínűségi változóra nézve *feltételesen függetlennek* nevezzük, ha bármely $b \in \mathcal{B}(\mathbb{R})$ elemre és $\mathcal{B} \subseteq \mathcal{B}(\mathbb{R})$ halmazra a $g_1^{-1}(\mathcal{B})$, $g_2^{-1}(\mathcal{B}) \dots g_n^{-1}(\mathcal{B})$ halmazok függetlenek a $p_{h^{-1}(b)} \equiv p(\,\cdot\,|h^{-1}(b))$ feltételes valószínűségi mértékre nézve. Ha $g_1, g_2, \dots g_n$ valamint h egy együttes $f_{g,h}$ sűrűségfüggvénnyel rendelkeznek, akkor a feltételes függetlenség szintén egyszerűbb formát ölt. Vezessük be a g_i valószínűségi változónak a h valószínűségi változóra vett $f_{g_i|h} : \mathbb{R}^2 \to \mathbb{R}$ *feltételes sűrűségfüggvényét* az alábbi módon:

$$f_{g_i|h}(x \,|\, x') \equiv \frac{f_{g_i,h}(x, x')}{f_h(x')},$$

ahol $f_{g_i,h}$ az g_i és a h valószínűségi változók együttes sűrűségfüggvénye. A feltételes sűrűségfüggvény minden olyan $x' \in \mathbb{R}$ helyen értelmezve van, ahol $f_h(x') \neq 0$. Hasonlóan bevezethetjük a $g_1, g_2, \dots g_n$ valószínűségi változóknak a h valószínűségi változóra vett $f_{g|h} : \mathbb{R}^{n+1} \to \mathbb{R}$ *együttes feltételes sűrűségfüggvényét* is:

$$f_{g|h}(x_1, x_2 \dots x_n \,|\, x_{n+1}) \equiv \frac{f_{g,h}(x_1, x_2 \dots x_n, x_{n+1})}{f_h(x_{n+1})},$$

ahol $f_{g,h}$ a $g_1, g_2, \dots g_n$ és h valószínűségi változók együttes sűrűségfüggvénye. Ezek után a $g_1, g_2, \dots g_n$ valószínűségi változók a h valószínűségi változóra feltételesen függetlenek, ha minden olyan $\mathbb{R}^{n+1}$-beli helyen, ahol $f_h(x_{n+1}) \neq 0$

$$f_{g|h}(x_1, x_2 \dots x_n \,|x_{n+1}) = f_{g_1|h}(x_1 \,|x_{n+1}) f_{g_2|h}(x_2 \,|x_{n+1}) \dots f_{g_n|h}(x_n \,|x_{n+1}),$$

vagyis ha az együttes feltételes sűrűségfüggvény előáll mint a feltételes sűrűségfüggvények szorzata.

Független és azonos eloszlású valószínűségi változókra fontos példa a szorzat valószínűségi mértéktereken értelmezett karakterisztikus függvények. Legyen (Ω, Σ, p) valószínűségi mértéktér, és képezzük Σ elemeiből a *mérhető téglákat*, azaz olyan végtelen $\times_{i\in\mathbb{N}} a_i$ szorzatokat ($a_i \in \Sigma$ minden i-re), ahol a szorzatban csak véges számszor fordul elő Ω-tól különböző elem. Jelöljük $\Sigma^{\mathbb{N}}$-nel a mérhető téglák által generált σ-algebrát. Ekkor a $(\Omega^{\mathbb{N}}, \Sigma^{\mathbb{N}})$ mérhető teret (Ω, Σ) végtelenszeres *szorzatterének* nevezzük. Ezen a mérhető téren egyértelműen létezik egy olyan $p^{\mathbb{N}}$ valószínűségi mérték, hogy

$$p(\times_{i\in\mathbb{N}} a_i) \quad = \quad \prod_i p(a_i),$$

azaz a téglák mértéke megegyezik oldalaik mértékének szorzatával. Az ezzel a valószínűségi mértékkel ellátott $(\Omega^{\mathbb{N}}, \Sigma^{\mathbb{N}}, p^{\mathbb{N}})$ valószínűségi mértékteret (Ω, Σ, p) (végtelenszeres) *szorzatmértékterének* nevezzük.

$\Sigma^{\mathbb{N}}$ végtelen sok teljesen független elemmel rendelkezik. Ilyenek például azok az $\tilde{a}_j$ szorzatelemek, amelyekre

$$\tilde{a}_j = \times_{i\in\mathbb{N}} b_i, \quad \text{ahol} \ \ b_i = \begin{cases} a, & \text{ha } i = j \\ \Omega, & \text{ha } i \neq j. \end{cases}$$

Most értelmezzük minden $a \in \Sigma$ eseményre az alábbi $\chi_j^a : \Omega^{\mathbb{N}} \to \{0,1\}$ valószínűségi változókat:

$$\chi_j^a(x_1, x_2 \dots) = \begin{cases} 1, & \text{ha } x_j \in a \\ 0, & \text{ha } x_j \notin a. \end{cases}$$

A χ_j^a valószínűségi változót az $\tilde{a}_j$ esemény *karakterisztikus függvényének* nevezzük. A karakterisztikus függvények (teljesen) függetlenek, mivel $\tilde{a}_j$-k függetlenek, és azonos (Bernoulli-) eloszlásúak, mivel minden j-re

$$\begin{aligned} (p^{\mathbb{N}} \circ \chi_j^{a\,-1})(\{1\}) \quad &= \quad p(a) \\ (p^{\mathbb{N}} \circ \chi_j^{a\,-1})(\{0\}) \quad &= \quad p(a^{\perp}). \end{aligned}$$

Mivel függetlenek és azonos eloszlásúak, továbbá mindegyik karakterisztikus függvény $Exp(\chi_j^a)$ várható értéke $p(a)$, $Var(\chi_j^a)$ szórása pedig $\sqrt{p(a)(1-p(a))}$, ezért alkalmazható rá a Bernoulli-tétel (1. tétel a 32. oldalon), vagyis igaz lesz, hogy a

$$\overline{\chi}_n^a \quad \equiv \quad \frac{1}{n}\sum_{j=1}^{n} \chi_j^a$$

átlag valószínűségi változó valószínűségi értelemben tartani fog a közös várható értékhez, vagyis

$$\lim_{n\to\infty} p^{\mathbb{N}}(|\overline{\chi}_n^a - p(a)| \leqslant \varepsilon) = 1. \tag{D.5}$$

A függetlenségnél gyengébb, de nagyon fontos fogalom a felcserélhetőség fogalma. Az (Ω, Σ, p) valószínűségi mértéktéren a $g_1, g_2, \dots g_n$ valószínűségi változókat *felcserélhetőnek* mondjuk, ha az $S(n)$ permutációcsoport minden π elemére a $g = (g_1, g_2, \dots g_n)$ és a $g_\pi = (g_{\pi_1}, g_{\pi_2}, \dots g_{\pi_n})$ vektorváltozók eloszlásfüggvénye megegyezik, azaz $p \circ g_\pi^{-1} = p \circ g^{-1}$. A definíció valószínűségi változók végtelen sorozatára is kiterjeszthető, ha megköveteljük, hogy a permutációinvariancia a sorozat minden kezdőszeletére teljesüljön.

Ha a valószínűségi változók függetlenek és azonos eloszlásúak (iid = *independent and identically distributed*), akkor nyilván felcserélhetők is. Erre jó példa fenti karakterisztikus függvények $\chi_1^a, \chi_2^a, \dots$ végtelen sorozata. Visszafelé azonban a felcserélhetőségből csak az azonos eloszlás következik, a függetlenség nem. Ezt illusztrálja Pólya híres urnamodellje.

Pólya-urna. Egy urnában p piros és k kék golyó van. Az urnából véletlenszerűen húzunk egy golyót, feljegyezzük a színét, majd a kihúzott golyót egy vele megegyező színű golyóval együtt visszahelyezzük az urnába. A húzásokat *ad infinitum* folytatjuk. Legyen g_i az i. húzást modellező valószínűségi változó, amelynek értéke legyen 1, amennyiben piros golyót húzunk, és legyen 0, amennyiben kéket. Könnyű megmutatni, hogy annak valószínűsége, hogy n húzásból m pirosat kapunk, nem függ a piros golyók elhelyezkedésétől a sorozatban, csak az m számtól; más szóval a valószínűségi változók minden $g_1, g_2, \dots g_n$ véges sorozata felcserélhető. Annak valószínűsége például, hogy az urnából három húzásból két pirosat kapunk a piros golyók sorrendjétől függetlenül $\frac{k\,p\,(p+1)}{(p+k)(p+k+1)(p+k+2)}$, mivel

$$\begin{aligned}
p(g_1 = 1, g_2 = 1, g_3 = 0) &= \frac{p}{p+k}\,\frac{p+1}{p+k+1}\,\frac{k}{p+k+2}, \\
p(g_1 = 1, g_2 = 0, g_3 = 1) &= \frac{p}{p+k}\,\frac{k}{p+k+1}\,\frac{p+1}{p+k+2}, \\
p(g_1 = 0, g_2 = 1, g_3 = 1) &= \frac{k}{p+k}\,\frac{p}{p+k+1}\,\frac{p+1}{p+k+2}.
\end{aligned}$$

Igaz továbbá az is, hogy az g_i valószínűségi változók minden i-re azonos eloszlásúak, vagyis

$$p \circ g_i^{-1}(\{x\}) = \begin{cases} \frac{p}{p+k}, & \text{ha } x = 1, \\ \frac{k}{p+k}, & \text{ha } x = 0. \end{cases}$$

Az első két húzás esetében ez könnyen látható:

$$\begin{aligned}
p(g_2=1) &= p(g_1=1,g_2=1)+p(g_1=0,g_2=1)= \\
&\frac{p}{p+k}\,\frac{p+1}{p+k+1}+\frac{k}{p+k}\,\frac{p}{p+k+1}=\frac{p(p+1)+kp}{(p+k)(p+k+1)} \\
&=\frac{p}{p+k}=p(g_1=1) \\
p(g_2=0) &= p(g_1=1,g_2=0)+p(g_1=0,g_2=0)= \\
&\frac{p}{p+k}\,\frac{k}{p+k+1}+\frac{k}{p+k}\,\frac{k+1}{p+k+1}=\frac{pk+k(k+1)}{(p+k)(p+k+1)} \\
&=\frac{k}{p+k}=p(g_1=0)
\end{aligned}$$

Ugyanakkor a valószínűségi változók nem függetlenek, hiszen például

$$\frac{p}{p+k}\,\frac{p+1}{p+k+1}=p(g_1=1,g_2=1) \quad \neq \quad p(g_1=1)\,p(g_2=1)=\frac{p}{p+k}\,\frac{p}{p+k}$$

Vagyis a felcserélhetőség erősebb tulajdonság, mint az azonos eloszlás, de gyengébb, mint a függetlenség *és* az azonos eloszlás együtt.

A felcserélhetőség azonban szoros kapcsolatban áll a feltételes függetlenség fogalmával: feltételesen független valószínűségi változók keveréke felcserélhető valószínűségi változókat eredményez. Ahhoz, hogy a keverék fogalmát értelmezni tudjuk, szükségünk lesz a teljes valószínűség tételének valószínűségi változókra alkalmazott folytonos alakjára. Amint azt az imént láttuk, a teljes valószínűség tétele egy $\{a_j\}_{j=1}^m$ partícióra és egy b algebraelemre vonatkozóan az alábbi összefüggés:

$$p(b) \quad = \quad \sum_{j=1}^{m} p(b|a_j)p(a_j).$$

Legyen most $b=\wedge_{i=1}^n c_i$, és tegyük fel, hogy a c_i események az $\{a_j\}_{j=1}^m$ partíció minden elemére nézve feltételesen *teljesen* függetlenek. Így minden $j=1\ldots m$-re teljesül, hogy

$$p(\wedge_{i=1}^n c_i|a_j) \quad = \quad \prod_{i=1}^{n} p(c_i|a_j),$$

és ezért a teljes valószínűség tétele a következő alakot ölti:

$$p(\wedge_{i=1}^n c_i) \quad = \quad \sum_{j=1}^{m}\left(\prod_{i=1}^{n} p(c_i|a_j)\right)p(a_j).$$

Most tegyük fel, hogy a c_i események illetve az $\{a_j\}_{j=1}^m$ partíció valamely g_i illetve h valószínűségi változó inverzképei, azaz léteznek $\{x_i\}_{i=1}^n$ és $\{x'_j\}_{j=1}^m$ valós számok, hogy $g_i^{-1}(\{x_i\}) = c_i$ és $h^{-1}(\{x'_j\}) = a_j$. Ekkor a fenti egyenletet az alábbi alakba írhatjuk:

$$p(g_1 = x_1, g_2 = x_2 \dots g_n = x_n) \;\; = \;\; \sum_{j=1}^m \left(\prod_{i=1}^n p(g_i = x_i \,|\, h = x'_j) \right) p(h = x'_j).$$

Végül pedig ha feltesszük, hogy h folytonos valószínűségi változó, valamint $g_1, g_2 \dots g_n$ és h együttes sűrűségfüggvénnyel rendelkezik, akkor a jobb oldali összeg helyett integrált írva megkapjuk a teljes valószínűség tételének valószínűségi változókra alkalmazott folytonos alakját.

$$p(g_1 = x_1, g_2 = x_2 \dots g_n = x_n) \;\; = \;\; \int_{\mathbb{R}} \prod_{i=1}^n f_{g_i|h}(x_i \,|\, x')\, f_h(x')\, dx', \tag{D.6}$$

ahol f_h a h valószínűségi változó sűrűségfüggvénye, $f_{g_i|h}$ pedig a g_i valószínűségi változónak a h valószínűségi változóra vett feltételes sűrűségfüggvénye.

Most térjünk vissza újra a felcserélhetőség és a feltételes függetlenség kapcsolatához. Legyen (Ω, Σ, p) valószínűségi mértéktér, és legyen $h : \Omega \to [0,1]$ egyenletes eloszlású valószínűségi változó, azaz legyen h sűrűségfüggvénye a következő:

$$f_h(x) \;\; \equiv \;\; \begin{cases} 1, & \text{ha } x \in [0,1], \\ 0, & \text{ha } x \notin [0,1]. \end{cases}$$

Legyenek továbbá $g_1, g_2, \dots g_n : \Omega \to \{0,1\}$ feltételesen független Bernoulli-változók a h valószínűségi változóra nézve. Ha $g_1, g_2, \dots g_n$-nek valamint h-nak van együttes sűrűségfüggvénye, akkor ez azt jelenti, hogy az $f_{g|h}$ együttes feltételes sűrűségfüggvény szorzat alakot ölt, azaz

$$f_{g|h}(\varepsilon_1, \varepsilon_2 \dots \varepsilon_n \,|x_{n+1}) \;\; = \;\; f_{g_1|h}(\varepsilon_1 \,|x_{n+1}) f_{g_2|h}(\varepsilon_2 \,|x_{n+1}) \dots f_{g_n|h}(\varepsilon_n \,|x_{n+1}) \tag{D.7}$$

minden $\varepsilon_i \in \{0,1\}$ és $x_{n+1} \in [0,1]$-re. Végül definiáljuk a feltételes sűrűségfüggvényeket minden i-re az alábbi módon:

$$f_{g_i|h}(\varepsilon_i \,|x_{n+1}) \;\; \equiv \;\; \begin{cases} x_{n+1}, & \text{ha } \varepsilon_i = 1, \\ 1 - x_{n+1}, & \text{ha } \varepsilon_i = 0. \end{cases} \tag{D.8}$$

A teljes valószínűség tételének (D.6) alakját használva a $g_1, g_2, \dots g_n$ valószínűségi változók eloszlása ekkor a következő lesz:

$$p(g_1 = \varepsilon_1, g_2 = \varepsilon_2, \dots g_n = \varepsilon_n) \;\; = \;\; \int_0^1 x^k (1-x)^{n-k}\, dx = \frac{k!(n-k)!}{(n+1)!},$$

ahol $k = \sum_{i=1}^{n} \varepsilon_i$. Mivel az eloszlás ($n$-nen kívül) csak k-tól, azaz az ε_i-k összegétől függ, sorrendjüktől azonban nem, ezért a $g_1, g_2, \dots g_n$ valószínűségi változók felcserélhetők. Feltételesen független Bernoulli-eloszlású valószínűségi változóknak keveréke tehát felcserélhető Bernoulli-valószínűségi változókat eredményez.

A (D.9) összefüggés fontos alkalmazása Laplace rákövetkezési szabálya.

3. Tétel. (Laplace rákövetkezési szabálya.) Legyen (Ω, Σ, p) valószínűségi mértéktér, és legyenek $g_1, g_2, \dots g_{n+1}$ feltételesen független Bernoulli-változók a h egyenletes eloszlású valószínűségi változóra nézve, ahol a $g_1, g_2, \dots g_{n+1}$ és h valószínűségi változóra vonatkozó feltételes sűrűségfüggvények a (D.7)-(D.8) egyenletek szerint vannak definiálva. Ekkor

$$p(g_{n+1} = 1 \mid g_1 + \dots + g_n = s) = \frac{s+1}{n+2}. \tag{D.9}$$

Bizonyítás. A (D.9) összefüggés alapján a

$$p(g_{n+1} = 1 \mid g_1 + \dots + g_n = s) = \frac{p(g_{n+1} = 1,\, g_1 + \dots + g_n = s)}{p(g_1 + \dots + g_n = s)}$$

feltételes valószínűség számlálója illetve nevezője:

$$\begin{aligned} p(g_{n+1} = 1,\, g_1 + \dots + g_n = s) &= \binom{n}{s} \int_0^1 x^{s+1}(1-x)^{n-s}\, dx = \binom{n}{s} \frac{(s+1)!(n-s)!}{(n+2)!}, \\ p(g_1 + \dots + g_n = s) &= \binom{n}{s} \int_0^1 x^{s}(1-x)^{n-s}\, dx = \binom{n}{s} \frac{s!(n-s)!}{(n+1)!}, \end{aligned}$$

amely két kifejezés hányadosa $\frac{s+1}{n+2}$. $\square$

A felcserélhető és feltételesen független eloszlások közötti fenti összefüggésnek azonban az ellenkezője is igaz, amint azt az alábbi de Finettitől (1937) származó tétel állítja:

10. Tétel. (De Finetti reprezentációs tétele.) Legyen (Ω, Σ, p) valószínűségi mértéktér, és legyen $g_1, g_2, \dots : \Omega \to \{0, 1\}$ felcserélhető Bernoulli-változók egy végtelen sorozata. Ekkor a valószínűségi változók minden véges n elemű halmazának eloszlása előáll mint valamilyen $h : \Omega \to [0, 1]$ valószínűségi változóra nézve feltételesen független, a (D.7)-(D.8) egyenleteknek megfelelő Bernoulli-változók keveréke, vagyis

$$\begin{aligned} p(g_1 = \varepsilon_1, g_2 = \varepsilon_2, \dots g_n = \varepsilon_n) &= \int_{\mathbb{R}} \prod_{i=1}^{n} f_{g_i|h}(\varepsilon_i \,|\, x)\, dF_h(x) \\ &= \int_0^1 x^k (1-x)^{n-k}\, dF_h(x), \end{aligned}$$

ahol az F_h eloszlásfüggvény a „relatív gyakoriságok” határértéke:

$$F_h(x) = \lim_{n \to \infty} p\left(\frac{1}{n} \sum_{i=1}^{n} \varepsilon_i \leqslant x \right).$$

A reprezentációs tétel csak felcserélhető valószínűségi változók *végtelen* sorozatára igaz. Arra, hogy véges esetben nem igaz, álljon itt a következő egyszerű ellenpélda (Diaconis 1977), amely a nem visszatevéses húzást modellezi egy olyan urnából, amely egy fekete és fehér golyót tartalmaz. Legyen g_1 és g_2 felcserélhető Bernoulli-változók az alábbi együttes eloszlással:

$$\begin{aligned} p(g_1 = 0, g_2 = 0) &= p(g_1 = 1, g_2 = 1) = 0, \\ p(g_1 = 0, g_2 = 1) &= p(g_1 = 1, g_2 = 0) = \frac{1}{2}, \end{aligned}$$

Ha most létezne olyan $\mu(x)$ mérték, amelyre a

$$0 = p(g_1 = 1, g_2 = 1) = \int_0^1 x^2 \, d\mu(x),$$

egyenlet fennállna, akkor μ teljes mértékben a 0-ra lenne koncentrálva, vagyis

$$0 = p(g_1 = 0, g_2 = 0) = \int_0^1 (1-x)^2 \, d\mu(x),$$

nem teljesülhetne.

De Finetti tétele nemcsak Bernoulli-változókra igaz. A tétel általánosabb megfogalmazásaihoz lásd például (Chow-Teicher, 1988, 226-228. o.).

Irodalom

Allais, M. (1979). "The so-called Allais Paradox and rational decisions under uncertainty,", in: Allais and Hagen (ed.) *Expected Utility Hypothesis and the Allais Paradox*, Dordrecht: Kluwer, 2009.

Arnauld, A., P. Nicole (1662). *La logique, ou l'art de penser*; Clair P., F. Girbal (eds.), Paris, 1965.

Arrow, K. (1974). *Essays in the Theory of Risk-Bearing*, Amsterdam: North-Holland.

Bar-Hillel. (1973). "On the subjective probability of compound events," *Organizational Behavior and Human Performance*, **9**, 396-406.

Bartha, P., Johns, R. (2001). "Probability and symmetry," *Philosophy of Science*, **68**, 109-122.

Bayes, T. (1763). "An essay towards solving the problem in the doctrine of chances," *Philosophical Transactions of the Royal Society of London*, **53**, 370-418.

Bernoulli, D. (1738). "Specimen theoriae novae de mensura sortis,", in: *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, **53**; Sommer, L. (trans.) "Exposition of a new theory on the measurement risk," in: *Econometrica*, **22**, 1954.

Bernoulli, J. (1713). *Ars conjectandi*, in: *Die Werke von Jacob Bernoulli* Basle Naturforschende Gesellshaft, 3. Vol., Basel, 1969-75.

Bertrand, J. L. F. (1889). *Calcul des probabilités*, Paris: Gauthier-Villars.

Billingsley, P. (1995). *Probability and Measure*, New York: John Wiley.

Black, M. (1967). "Probability," in: Paul Edwards (ed.) *The Encyclopedia of Philosophy*, New York: Macmillan.

Boltzmann, L. (1872). "Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen," Hasenöhrl F. (ed.), *Wissenchaftlichen Abhandlungen*, 1909, 1. Vol., 23.; reprint New York, 1969.

Boltzmann, L. (1877). "Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung resp. dem Sätzen über das Wärmegleichgewicht," Hasenöhrl F. (ed.), *Wissenchaftlichen Abhandlungen*, 1909, 2. Vol., 42.; reprint New York, 1969.

Boole, G. (1852). *Studies in Logic and Probability*, London: Watts and Co.

Borel, E. (1898). *Leçons sur la théorie des fonctions*, Paris: Gauthier-Villars.

Borel, E. (1909). "Les probabilités dénombrales et leur applications arithmétiques," *Rendiconti del Circolo Matematico di Palermo*, **27**, 247-270.

Cardano, G. (1663). *De ludo aleae*, in: *Opera Omnia*, 1. Vol., Amsterdam, 1966.

Carnap, R. (1950). *Logical Foundations of Probability*, Chicago: University of Chicago Press.

Carnap, R. (1952). *The Continuum of Inductive Methods*, Chicago: University of Chicago Press.

Carnap, R., W. Stegmüller (1959). *Induktive Logik und Wahrscheinlichkeit*, Wien: Springer.

Carnap, R. (1963). *Replies and Systematic Exposition*, in: P. A. Schlipp (ed.) *The Philosophy of Rudolf Carnap, The Library of Living Philosophers*, Cambridge: Cambridge University Press.

Carnap, R. (1971). "A Basic System of Inductive Logic, Part 1," in: Carnap R., R. C. Jeffrey (ed.) *Studies in Inductive Logic and Probability*, 1. Vol., Berkeley: University of California Press, 33–165.

Carnap, R. (1980). "A Basic System of Inductive Logic, Part 2," in: Jeffrey R. C. (ed.) *Studies in Inductive Logic and Probability*, 2. Vol., Berkeley: University of California Press, 7-155.

Champernowne, D. G. (1933). "The construction of decimal normal numbers in the scale of ten," *Jour. Lond. Math. Soc.*, **8**, 254-260.

Chapman, L. J., L. P. Chapman (1969). "Illusory correlation as an obstacle to the use of valid psychodiagnostic signs," *Journal of Abnormal Psychology*, **74**, 271-280.

Church, A. (1940). "The concept of a random sequence," *Bull. Am. Math. Soc.*, **46**, 130-135.

Chow, Y. S., H. Teicher (1988). *Probability Theory*, Berlin: Springer.

Condorcet, M. J. A. N. (1785/2009). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Paris: Kessinger Publishing Co.

Copeland, A. H. (1932). "The theory of probability from the point of view of admissible numbers," *Ann. Math. Stat.*, **3**, 143-156.

Cramér, H. (1946). *Mathematical Methods of Statistics*, Princeton: Princeton University Press.

Daston, L. (1988). *Classical Probability in the Enlightenment*, Princeton: Princeton University Press.

Daston, L. (2006). "Probability and Evidence" in: Park, K., L. Daston (eds.), *The Cambridge History of Science, Early Modern Science*, Cambridge: Cambridge University Press.

Davidson, D. (1980). *Essays on Actions and Events*, New York: Oxford University Press.

De Moivre, A. (1718). *The Doctrine of Chances, or a Method of Calculating the Probability of Events in Play*, London; reprint New York, 1967.

De Morgan, A. (1847). *Formal Logic: or, The Calculus of Inference, Necessary and Probable*, London: Taylor and Walton.

Diaconis, P. (1977). "Finite forms of de Finetti's theorem on exchangeability," *Synthese*, **36**, 271-281.

Doob, J. (1996). "The development of rigor in mathematical probability theory (1900-1950)," *American Mathematical Monthly*, 586-595.

Eagle, A. (2004). "Twenty-one arguments against propensity analyses of probability," *Erkenntnis*, **60**, 371–416.

Earman, J. (1986). *A Primer on Determinism*, Dordrecht: D. Reidel.

Earman, J. (1992). *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, Cambridge, MA: MIT Press.

Ellis, J. (1849). "On the foundations of the theory of probability," *Transactions of the Cambridge Philosophical Society*, **9**, 605-607.

Fermat, P. (1654). Pascal–Fermat-levelezés, in: Tannery P., C. Henry (eds.), *Œuvres complètes*, 2. Vol., Paris, 1891-1922.

Fetzer, J. H. (1971). "Dispositional probabilities," *Boston Studies in the Philosophy of Science*, **8**, 473-82.

Fetzer, J. H. (1973). "Physical probabilities: a conceptual dilemma," (kézirat).

Fetzer, J. H. (1981). *Scientific Knowledge: Causation, Explanation, and Corroboration*, Boston Studies in the Philosophy of Science, 69. Vol., Dortrecht: D. Reidel.

Finetti, B. de (1931). "Probabilismo," *Logos*, **14**, 163-219.; angolul: Di Maio M. C., M. C. Galavotti, R. C. Jeffrey (trans.): "Probabilism: A Critical Essay on the Theory of Probability and on the Value of Science," *Erkenntnis*, **31**, 169-223. 1989.

Finetti, B. de (1937). "La prévision: ses lois logiques, ses sources subjectives," *Annales de l'Institute Henry Poincaré*, **7**, 1-68.; angolul: Kyburg, H. R. Jr., H. E. Smokler (eds.) *Studies in Subjective Probability*, New York: Krieger, 1980.

Finetti, B. de (1972). *Probability, Induction and Statistics*, London: John Wiley.

Finetti, B. de (1974). *Theory of Probability: A Critical Introductory Treatment*, 2. Vol. Machi, A., A. Smith (trans.), New York: John Wiley.

Finetti, B. de (1995). *Filosofia della Probabilitá*, Il Saggiatore, Milano.

Frechét, M. (1915). "Sur l'intégrale d'une fonctionelle étendue à un ensemble abstrait," *Bulletin de le Société mahtématique de France*, **43**, 248-265.

Galilei, G. (1623?). *Sopra le scoperte dei dadi* in: Favaro A. A. (ed.), *Opere*, 8. Vol., Firenze, 1890-1909.

Giere, R. N. (1973). "Objective single-case probabilities and the foundation of statistics," in: Suppes, P. et al. (eds.) *Logic, Methodology and Philosophy of Science IV.*, Amsterdam: North Holland, 467-483.

Giere, R. N. (1976a). "A Laplacean formal semantics for single case probabilities," *Journal of Philosophical Logic*, **5**, 321-353.

Giere, R. N. (1976b). "Empirical probability, objective statistical methods and scientific inquiry," in: Harper, W., C. A. Hooker (eds.) *Foundations of Probability Theory, Statistical Inference and Statistical Theories in Science*, Dordrecht: D. Reidel, 2. Vol., 63-101.

Gillies, D. (2000a). *Philosophical Theories of Probability*, London: Routledge.

Gillies, D. (2000b). "Varieties of propensity," *British Journal for the Philosophy of Science*, **51**, 807-835.

Graunt, J. (1662). *Natural and Political Observations mentioned in a following Index, and made upon the Bills of Mortality*, London, reprint: Willcox W. F. (ed.) *Natural and Political Observations mentioned in a following Index, and made upon the Bills of Mortality by John Graunt*, Baltimore: John Hopkins, 1939.

Hacking, D. (1965). *Logic of Statistical Inference*, Cambridge: Cambridge University Press.

Hacking, D. (1971). "The Leibniz-Carnap program for inductive logic," *The Journal of Philosophy*, *68*, 597-610.

Hacking, D. (1975). *The Emergence of Probability*, Cambridge: Cambridge University Press.

Hacking, D. (2001). *An Introduction to Probability and Inductive Logic*, Cambridge: Cambridge University Press.

Hájek, A. (2003). "What conditional probability could not be," *Synthese*, **137**, 273–323.

Hájek, A. (2007). "The reference class problem is your problem too," *Synthese*, **156**, 563–585.

Halley, E. (1693). "An estimate of the degrees of mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslau; with an attempt to ascertain the price of annuities upon lives," *Philosophical Transactions of the Royal Society of London*, **17**, 596-610; 654-6.

Hilbert, D. (1976). "Mathematical problems. Lecture delivered before the International Congress of Mathematicians at Paris in 1900," *Proceedings of Symposia in Pure Mathematics*, **28**, 1-34.

Hitchcock, C. (2002). "Probability and Chance," in: Smelser N. J., P. B. Baltes (eds.): *International Encyclopedia of the Social and Behavioral Sciences*, 18. Vol., London: Elsevier, 12089 - 12095.

Humphreys, P. W. (1985). "Why propensities cannot be probabilities," *Philosophical Review*, **94**, 557-570.

Huoranszki, F. (2001). *Modern metafizika*, Budapest: Osiris.

Huygens, C. (1657). *De ratiociniis in aleae ludo*, Amsterdam, in: *Œuvres complètes*, Société Hollandaise des Sciences, 22. Vol., The Hague, 1888-1967.

Janes, E. T. (2003). *Probability Theory: The Logic of Science*, Cambridge: Cambridge University Press.

Jeffrey, C. (1965). *The Logic of Decision*, New York: McGraw-Hill; második kiadás: Chicago: University of Chicago Press, 1983.

Jeffrey, C. (2004). *Subjective Probability: The Real Thing*, Cambridge: Cambridge University Press.

Jeffreys, H. (1939). *Theory of Probability*, Oxford: Oxford University Press.

Jevons, W. S. (1873). *The Principles of Science*, London: Macmillen.

Kahneman, D., A. Tversky (1972). "Subjective probability: a judgment of representativeness," *Cognitive Psychology*, **3**, 430-454.

Kahneman, D., A. Tversky (1973). "On the psychology of prediction," *Psychological Review*, **80**, 237-251.

Kahneman, D., P. Slovic, A. Tversky (eds.) (1982). *Judgment Under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.

Kamke, E. (1932). "Über neuere Begründungen der Wahrscheinlichkeitsrechnung," in: Braithwaite R. B. (ed.), *Jahresbericht der Deutsche Mathematiker-Vereinigung*, **42**, 14-27.

Keynes, J. M. (1921). *A Treatise on Probability*, London: Macmillen.

Kim, J. (1998). *Mind in a Physical World*, Cambridge MA: MIT Press.

Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer; magyarul: Zibolen Endre (ford.) *A valószínűségszámítás alapfogalmai*, Budapest: Gondolat, 1982.

Kolmogorov, A. N. (1965). "Three approaches to the quantitative definition of information," *Problemy Perdaci Informacii*, **1**, 4-7.

Kyburg, H. E., H. Smokler (eds.) (1964). *Studies in Subjective Probability*, New York: John Wiley.

Kyburg, H. E. (1974). "Propensities and probabilities," *British Journal for the Philosophy of Science*, **25**, 352-375.

Lambalgen, M. (1987). "Von Mises' definition of random sequences reconsidered," *Journal of Symbolic Logik*, **52**, 725-55.

Lambalgen, M. (1987). *Random sequences*, Dissertation, Department of Mathematics, University of Amsterdam.

Lambalgen, M. (1996). "Randomness and foundations of probability: von Mises' axiomatization of random sequences," in: Fergusson T. et al. (eds.), *Probability, Statistic and Game Theory, papers in honor of David Blackwell*, Institute for Mathematical statistics.

Laplace, P. S. (1795). *Essaie philosophique sur les probabilités*, Amsterdam. in: *Œuvres complètes*, Académie des Sciences, 7. Vol., Paris, 1878-1912.

Laplace, P. S. (1812). *Théorie analytique des probabilités*, in: *Œuvres complètes*, Académie des Sciences, 7. Vol., Paris, 1878-1912.

Lebesgue, H. (1904). *Leçons sur l'intégration et la recherche des fonctions primitives*, Paris: Gauthier-Villars.

Lehman, R. S. (1955). "On confirmation and rational betting," *The Journal of Symbolic Logic*, **20**, 251-262.

Leibniz, G. W. (1763). *De conditionibus*, in: Gerdhardt C. I. (ed.), *Die philosophischen Schriften von G. W. Leibniz*, 7. Vol., Berlin, 1875-1890.

Levi, I. (1980). *The Enterprise of Knowledge: An Essay an Knowledge, Credal Probability and Chance*, Cambrige MA: MIT Press.

Lewis. D. (1972). "General semantics," in: D. Davidson, Harman, G. (eds.) *Semantics of Natural Language*, Dordrecht: D. Reidel.

Lewis. D. (1986). "A Subjectivist's Guide to Objective Chance," reprint with postscripts added of (1980). *Philosophical Papers*, 2. Vol. Oxford: Oxford University Press, 83–132.

Lewis. D. (1994). "Humean Supervenience Debugged," *Mind*, **103**, 473–490.

Luce, R. D., H. Raiffa (1954). *Games and Decisions: Introduction and Critical Survey.*, New York: John Wiley.

Maher. P. (1993). *Betting on Theories*, Cambridge: Cambridge University Press.

Martin-Löf. P. (1966). "The definition of random sequences," *Information and Control*, **9**, 602–619.

Maxwell, J. C. (1860). "Illustrations of the dynamical theory of gases," *Philosophical Magazine*, **19**, 19-32; **20**, 21-37.

Maxwell, J. C. (1867). "On the dynamical theory of gases," *Philosophical Transactions of the Royal Society of London*, **157**, 49-88.

Mellor, D. H. (1971). *The Matter of Chance*, Cambridge: Cambridge University Press.

Mellor, D. H. (2005). *Probability: A Philosophical Introduction*, London: Routledge.

Miller, D. W. (21994). *Critical Rationalism. A Restatement and Defence*, Chicago: Open Court.

Mises, R. von (1928/51). *Wahrscheinlichkeit, Statistik und Wahrheit*, Berlin: Springer.

Mises, R. von (1931). *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik*, Wien: Deuticke.

Mises, R. von (1964). *Mathematical Theory of Probability and Statistics*, New York: John Wiley.

Neumann, J., O. Morgenstern (1964). *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.

Pacioli, L. (1494). *Summa de arithmetica, geometria, proportioni et proportionalitá*, Venezia.

Pascal, B. (1954). *Pensées*, in: Chevalier J. (ed.) *Œuvres complètes*, Paris: Gallimard; magyarul: Pődör László (ford.) *Gondolatok*, Budapest: Gondolat, 1983.

Petty, W. (1666). "Review of Graunt," *Le Journal des Sçavans*, **2**, 359-70.

Peirce, C. S. (1910). "Notes on the Doctrine of Chances," in: Buchler J. (ed.), *Philosophical Writings of Peirce*, New York: Dover.

Plato, J. von (1994). *Creating Modern Probability*, Cambridge: Cambridge University Press.

Poisson, S. (1837). *Recherches sur la probabilité des jugements en matière criminelle et en matière civile, precédées des règles générales du calcul des probabilités*, Paris.

Popper, K. (1935). *Logik der Forschung*, Berlin: Springer; magyarul: Petri Gy., Szegedi P. (ford.), *A tudományos kutatás logikája,* Budapest: Európa, 1997.

Popper, K. (1957). "The propensity interpretation of the calculus of probability, and the quantum theory," in: Körner S. (ed.), *Observation and Interpretation*, London: Butterworth, 65–70.

Popper, K. (1959). "The propensity interpretation of probability," *British Journal for the Philosophy of Science*, **10**, 25–42.

Popper, K. (1983). *Realism and the Aim of Science*, New Jersey: Rowman and Littlefield.

Popper, K. (1990). *The World of Propensities*, Bristol: Thoemmes.

Price, R. (1767). *Four Dissertations*, London: Millar and Cadell, 1811.

Ramsay, F. (1926). "Truth and probability," in: Braithwaite R. B. (ed.), *The Foundations of Mathematics and other Logical Essays*, London: Kegan Paul, 1931.

Rédei, M., S. J. Summers (2007). "Quantum probability theory," *Studies in the History and Philosophy of Modern Physics*, **38**, 390-417.

Reichenbach, H. (1932). "Axiomatik der Wahrscheinlichkeitsrechnung," *Math. Zeitschrift*, **34**, 568-619.

Reichenbach, H. (1949). *The Theory of Probability*, Berkeley: University of California Press.

Reichenbach, H. (1956). *The Direction of Time*, Berkeley: University of California Press.

Rényi, A. (1970). *Foundations of Probability*, San Francisco: Holden-Day.

Ross, S. M. (2009). *A First Course in Probability*, Prentice Hall.

Salmon, M. H., J. Earman, C. Glymour, J. Lennox (eds.) (1992). *Introduction to the Philosophy of Science*, Prentice-Hall.

Salmon, W. C. (1966). *The Foundations of Scientific Inference*, Pittsburgh: University of Pittsburgh Press.

Salmon, W. C. (1979). "Propensities: a discussion review of D. H. Mellor *The Matter of Chance*," *Erkenntnis*, **14**, 183-216.

Savage, L. J. (1954). *The Foundations of Statistics*, New York: Dover.

Schervish, M. J., Seidenfeld, T., and Kadane, J. B.. (2000). "How sets of coherent probabilities may serve as models for degrees of incoherence," *Journal of Uncertainty, Fuzziness, and Knowledge-based Systems*, **8 (3)**, 347-356.

Schneider, S. (2005). "Events," *Internet Encyclopedia of Philosophy*, URL = http://www.iep.utm.edu/events.

Sklar, H. (1993). *Physics and Chance*, Cambridge: Cambridge University Press.

Stegmüller, W. (1973). *Personelle und Statistische Wahrscheinlichkeit*, Zweiter Halbband. *Statistisches Schliessen, Statistische Begründung, Statistische Analyse,* Berlin: Springer.

Suppes, P. (1973). "New foundations of objective probability: Axioms for propensities," in Suppes P., L. Henkin, G. C. Moisil, A. Joja (eds.), *Logic, methodology, and philosophy of science IV: Proceedings of the Fourth International Congress for Logic, Methodology and Philosophy of Science, Bukarest, 1971*, Amsterdam: North-Holland, 515-529.

Suppes, P. (1984). *Probabilistic Metaphysics*, Oxford: Blackwell.

Suppes, P., M. Zanotti (1996). *Foundations of Probability with Applications: Selected Papers 1974-1995*, Cambridge: Cambridge University Press.

Szabó, G. (2010). "Lewis, valószínűség, *Principal Principle*," *Világosság*, ???.

Szabó, L. E. (2002). *A nyitott jövő problémája*, Budapest: Typotex.

Szabó, L. E. (2003). "Formal systems as physical objects: a physicalist accou:nt of mathematical truth," *International Studies in the Philosophy of Science*, **17**, 117-125.

Szabó, L. E. (2009). "What remains of probability?," in: Dieks D., W. Gonzalez, S. Hartmann, M. Weber, F. Stadler and T. Uebel (eds.): *The Present Situation in the Philosophy of Science*, New York: Springer.

Székely, J. G. (2004). *Paradoxonok a véletlen matematikájában*, Budapest: Typotex.

Tartaglia, N. (1556). *General trattato di numeri et misure*, Venezia.

Tornier, E. (1933). "Comment," *Math. Ann.* **108**, 320.

Torretti, R. (1990). *Creative Understanding: Philosophical Reflections on Physics*, Chicago: University of Chicago Press.

Tőzsér, J. (2009). *Metafizika*, Budapest: Akadémiai.

Uffink, J. (1999). "The principle of the common cause faces the Bernstein paradox," *Philosophy of Science* **66 (3)**, 525.

Van Fraassen, B. C. (1980). *The Scientific Image*, Oxford: Clarendon.

Venn, J. (1866). *The Logic of Chance*, London: Macmillan.

Ville, J. (1939). *Étude critique de la notion de collectif*, Paris: Gauthiers-Villars.

Wald, A. (1937). "Die Widerspruchsfreiheit des Kollektivbegriffes der Wahrscheinlichkeitsrechnung," *Ergebnisse eines mathematischen Kolloquiums*, 8. Vol., 38-72.

Wittgenstein, L. (1922/56). *Tractatus Logico-Philosophicus*, London: Routledge & Kegan Paul; magyarul: Márkus Gy. (ford.) *Logikai-filozófiai értekezés*, Budapest: Akadémiai, 1989.

Név- és tárgymutató