

# Mi köze a Bayes-szabálynak az elmegyógyászathoz?

## Rövid bevezetés a komputációs pszichiátriába a percepció bayesiánus prediktív kódolási elméletén keresztül

PÁLFFY ZSÓFIA<sup>1</sup>, FARKAS KINGA<sup>2,3</sup> ÉS POLNER BERTALAN<sup>1</sup>

<sup>1</sup> Kognitív Tudományi Tanszék, Természettudományi Kar, Budapesti Műszaki és Gazdaságtudományi Egyetem, H-1111 Budapest, Műegyetem rkp. 3.

<sup>2</sup> Semmelweis Egyetem, Pszichiátriai és Pszichoterápiás Klinika, Budapest

<sup>3</sup> Természettudományi Kutatóközpont, Kognitív Idegtudományi és Pszichológiai Intézet, Emlékezet, Nyelv és Idegtudomány Kutatócsoport, Budapest

A komputációs pszichiátria egy ágazata a mentális zavarok heterogenitásából és komorbiditásaiból adódó problémák kezelésére tesz kísérletet, a tünetek hátterében álló fenotipikus eltérések megragadása által. Az elme működésének és zavarainak pontosabb leírásához a komputációs kognitív tudomány módszertani eszköztárát és fogalomkészletét használja. A tanulmányban bemutatjuk azt a nyelvezetet – matematikai keretrendszert –, amely a pszichózisban, illetve autizmus spektrum zavarban érintettek észlelési eltéréseinek egy lehetséges magyarázatát: a bayesiánus prediktív kódolási elméletét írja le. Eloszlásként reprezentálva az előzetes elvárásokat (prior) és az új, mintavételezett információt (likelihood), a Bayes-szabály segítségével kiszámítható a poszterior eloszlás, amellyel az előzőek alapján levont következtetés jellemezhető. A fenti fogalmakat olyan egyszerű példákon keresztül mutatjuk be, mint az érmefeldobás (cinkeltség megállapítása), vagy a betegségek diagnosztizálása eltérő gyakoriságú tünetek alapján, ugyanakkor ez a keretrendszer az elmeműködés összetettebb folyamatainak megértéséhez is hozzájárulhat. Az eddigi kutatások abba az irányba mutatnak, hogy a vizuális percepció jól vizsgálható a bayesiánus logikát követve. A likelihood információ kísérletesen manipulálható, ha az alkalmazott ingeranyag kétértelmű vagy zajos. Ezen felül instrukciók és/vagy kulcsingerek révén asszociációk, ezáltal priorok kialakítása is lehetséges (ezek az egyénre jellemző, rejtett paraméterek). A perceptuális és válaszadási modellek segítségével kognitív feladatokon nyújtott teljesítmény alapján azonosíthatók olyan paraméterek, amelyek rávilágítanak arra, hogy milyen információ-feldolgozási sajátosság áll az egyes klinikai populációk észlelési eltéréseinek hátterében. A tanulmány végén felhívjuk a figyelmet a komputációs modellezés limitációira is.

*(Neuropsychopharmacol Hung 2022; 24(1): 4–16)*

**Kulcsszavak:** kognitív tudományok, percepció, szkizofrénia, autizmus spektrum zavar

A komputációs pszichiátria széles körben elterjedt megközelítés, mégis, ha egy érdeklődő szeretne belekóstolni, magyar nyelven alig talál róla irodalmat. Az angol szakirodalomnak ráadásul – a megértést tovább nehezítendő – saját fogalomrendszere van, jelentős része a matematika nyelvét használja, így könnyen elveszettnek érezhetjük magunkat benne.

Jelen tanulmány célja, hogy bevezessen ebbe a gazdag és szerteágazó tudományterületbe. Először felvázoljuk, hogy milyen pszichiátriát érintő problémákra nyújthat megoldást a komputációs megközelítés, majd részletesebben megismerkedünk egy speciális irányzattal: a generatív modellek fogalmával és a kapcsolódó matematikai problémákkal és megoldásokkal.

A „technikaibb” részek megértéséhez elegendő a gimnáziumi matematikai tudás. Ezt követően beszélünk arról, hogy az emberi elmére miért tekinthetünk következtető gépként, s végül az észlelésre fókuszálva új megvilágításba kerül az, ahogyan a neurotipikus (azaz szokványos módon fejlődő) idegrendszer működése és a pszichopatológia elválasztható: az információfeldolgozás eltéréseinek modellezésén keresztül.

## PROBLÉMAFELVETÉS

Egy pszichiátriai diagnózis felállítása számos esetben nem egyértelmű vagy egyszerű feladat. Ennek háttérben többek között az áll, hogy a - jelenleg használatban lévő - DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Mentális zavarok diagnosztikai és statisztikai kézikönyve, 5. kiadás) (APA, 2013) által meghatározott betegségkategóriák tünetegyüttesei gyakran átfedőek (Borsboom és mtsai., 2011), ráadásul az azonos kategóriába kerülő betegek csoportja is igen heterogén lehet, ami nemcsak a diagnózisalkotást nehezíti meg, hanem a terápiás lehetőségek közötti választást is. Különös problémát okoz a heterogenitás a major pszichiátriai kórképek, például szkizofrénia esetében, ahol a tüneti kép, a betegség lefolyása vagy a gyógyszeres kezelésre való érzékenység is jelentős egyéni eltéréseket mutat (Silveira és mtsai., 2012).

Egyesek azt remélték, hogy a deskriptív szemlélet elhagyása, az agyi képpalkotás és a biológiai pszichiátria fejlődése jelentheti a megoldást a problémára. Számtalan tanulmány jelent meg, amelyben a mentális betegségek háttérében álló biológiai folyamatokat igyekeznek feltérképezni, genetikai variációk megjelenésével, neurotranszmitterek arányának felborulásával, avagy egyes agyterületek struktúrájában és/vagy funkciójában bekövetkező változásokkal magyarázva egyes mentális zavarok vagy magtünetek kialakulását (Trimble & George, 2010). Ezek a kutatások sikeresen rávilágítottak például a striatális dopaminerg rendszerben való eltérésekre szkizofréniaiban. Azonban jelentős heterogenitás figyelhető meg a zavarokon belül is a biológiai folyamatok szempontjából, példánknál maradva a striatális dopaminműködés eltérése nem jellemző mindenkire, akit szkizofréniaival diagnosztizálnak (Brugger és mtsai., 2020). Adódik a kérdés: hogyan lenne érdemes kezelni a diagnosztikus kategóriákon belüli változatosságot?

Kapur és mtsai (2012), valamint Stephan és Mathys (2014) is a valid, megbízható konfirmációs tesztek hiányában látják a problémakör alapját, vagyis azt állítják, hogy a gondok abból származnak, hogy nincs olyan egyértelmű teszt, vizsgálati eljárás, amely ré-

vén nagy pontossággal megállapítható, hogy az adott páciens jelenleg egy adott pszichiátriai kórképben szenved, és ez alapján adott kezelés javasolt számára. Természetesen az orvostudomány minden ágazata megfigyelésekből indult, azonban a pszichiátria területén a potenciális biomarkerek túlnyomó többsége nem bizonyult klinikailag elegendőnek pontosság és előrejelzés tekintetében (Moriarity & Alloy, 2021) – itt a biológiai alapú tesztek (képpalkotó-, elektrofiziológiai vagy laboratóriumi módszerek) elsősorban más betegségek kizárását teszik lehetővé (pl. pszichózis esetén központi idegrendszeri térfoglalás, temporális lebeny epilepszia, velokardiofaciális-szindróma, NMDA-receptor enkefalitisz vagy egyéb organikus elváltozások) (Keshavan & Kaneko, 2013; Lisanby és mtsai., 1998; Sommer és mtsai., 2013).

Arra van tehát szükség, hogy egy azonos diagnózissal címkézett, azonban heterogén betegcsoportot megfelelő dimenziók mentén olyan kategóriákba tudjunk osztani, melyek megbízhatóan előre tudják jelezni a betegség lefolyását, és/vagy a különböző kezelések hatékonyságát. Ezt a biológiai pszichiátria (eddig) nem tudta megadni, lehetőséget adva egy újabb ágazat, a komputációs pszichiátria előretörésének, amely alapjául a komputációs idegtudomány szolgál. A komputációs irodalomban az egyes elméletek különböző szinteken (Marr és mtsai., 1979) adhatnak magyarázatot az agyműködésre. Marr klasszikus elmélete szerint elkülöníthető komputációs, algoritmikus és implementációs szint. Számos kutatás célja megérteni a legelemibb, úgynevezett implementációs szintet, s innen kiindulva a neuronok sejtszintű folyamatai, illetve kapcsolatrendszerük alapján következtetni arra, hogy az idegrendszer, mint biológiai, komputációkra alkalmas rendszer milyen számításokat képes elvégezni. Ahogy azonban – klasszikus példát idézve – egy-egy toll felépítéséből nehezen lehet következtetni arra, hogy hogyan repülnek a madarak, úgy az agy által végzett, bonyolultabb feladatok, mint a mozgástervezés vagy az észlelés nehezen térképezhető fel pusztán alulról szerveződő megközelítés segítségével (Krakauer és mtsai., 2017; de lásd Buzsáki, 2020).

Jelen tanulmány fókuszában így elsősorban azok a kutatások állnak, melyekben az ennél absztraktabb szinteken gondolkodnak. Az algoritmikus elemzési szinten a hangsúlyt a komputációk végrehajtásához szükséges reprezentációkra és az ezekkel végzett műveletek leírására helyezik, míg a komputációs szinten a célok a problémákra adott optimális megoldás formájában jelennek meg. Ennek jelentőségét Montague és mtsai. (2012, o. 72) az emberi kogníció

komputációs fenotipizálásában látják. Megfogalmazásuk szerint a „fenotípus egy organizmus mérhető tulajdonsága, mely magába foglal különböző morfológiákat, biokémiai kaszkádokat, neurális konnektivitási mintázatokat, viselkedési mintákat stb. A kognitív fenotípus a kognitív funkciók egy olyan mintázata egy adott területen, amely különböző gondolkodási stílusok kategorizációjára alkalmas. A komputációs fenotípus egy mérhető viselkedéses vagy neurális típus, amely valamely komputációs modell tekintetében értelmezhető.“ Elgondolásuk alapján ezen komputációs fenotípusok meghatározása, majd viselkedéses jellemzők és/vagy neurális, genetikus korrelációk azonosítása révén újfajta betegségkategoriók alakíthatók.

A komputációs pszichiátria eleve igen szerteágazó tudományág, de ezen belül a modellezési lehetőségek tere is széles, mivel ahogy Huys és mtsai. (2015) rámutattak, a normatív viselkedés komputációs szempontból több ponton is tévútra léphet a problémamegoldás során (a probléma összetevőit értelmezzük rosszul, a megoldási módszer nem optimális, vagy nem sikerült adaptálódni egy új környezethez). Tehát a komputációs pszichiátria irodalmában sokféle modellezési megközelítéssel találkozhatunk, például neurális hálókkal és áramkörökkel (neural networks and circuits), drift-diffúziós modellekkel (drift-diffusion models), megerősítéses tanulási modellekkel (reinforcement learning models), illetve bayesi modellekkel és prediktív kódolással (Series, 2020). Jelen tanulmányban a téma egy szűkebb problémakörére helyezük a hangsúlyt: az észleléssel, mint modellen alapuló (model based), optimális, bayesiánus elvekre épülő információfeldolgozással és kiemelten a bayesiánus prediktív kódolási elmélettel fogunk részletesebben foglalkozni. Ahogy később bemutatjuk, a valószínűségszámítás szabályait követő bayesiánus következtetés egy optimális megoldás az érzékelt információkat okozó folyamatok megállapítására, s ennek egy algoritmikus realizációja a prediktív kódolás lehet (Aitchison & Lengyel, 2017). Sterzer és kollégái (2018) amellett érvelnek, hogy ez a keretrendszer elég komplex ahhoz, hogy sikeresen egyesítse a pszichózis neurobiológiai magyarázatait a klinikai leírásokkal és a szubjektív élményvilággal. Ehhez szorosan kapcsolódva Lawson és kollégái (2014) pedig az autizmusban érintettek percepciói eltéréseinek hátterét vázolták fel, szintén a bayesiánus formalizáció és a neurális elváltozások összekötésével.

A komputációs pszichiátria tehát a mentális zavarok heterogenitásából és komorbiditásaiból adódó problémák kezelésére tesz kísérletet, a tünetek hátterében

álló fenotipikus eltérések megragadása által. Az elme működésének és zavarainak pontosabb leírásához a komputációs kognitív tudomány módszertani eszköztárát és fogalomkészletét használja. A következő fejezetben bemutatjuk azt a nyelvezetet – matematikai keretrendszer –, amely a pszichózisban, illetve autizmus spektrum zavarban érintettek észlelési eltéréseinek egy lehetséges magyarázatát: a bayesiánus prediktív kódolás elméletét írja le.

## BEVEZETÉS A BAYESI MODELLEZÉSBE

### *Generatív modellek: milyen eséllyel dobunk fejet egy érmével?*

A kapcsolódó irodalmat olvasva nem megerőltető a generatív modellek fogalma. Generatívnak egy olyan folyamatot hívhatunk, ahol bizonyos feltételek teljesülése esetén, adott paraméterek mellett meg tudjuk állapítani, hogy (az előzőleg azonosított) kimenetek közül melyik milyen valószínűséggel fog bekövetkezni. Amint azt később látni fogjuk, a generatív modellek segítségével vizsgálható az észlelés, ami felfogható annak a problémának a megoldásaként, hogy a külvilágban milyen tárgy (paraméterek) okozhatta az érzéketeket (kimenetel), és generatív modelleket használhatunk olyankor is, amikor egy vizsgálati személy gombnyomásai (kimenetel) alapján szeretnénk következtetni mentális reprezentációira (paraméterek).

Bevezetesként vegyünk egy egyszerűbb jelenséget, és tekintsünk egy szabályos érme feldobására generatív folyamatként: ennek lehetséges kimenetelei a 'fej' és az 'írás', amelyek azonos valószínűséggel következnek be (eltekintve attól, hogy elhanyagolható valószínűséggel az élére is érkezhetsz). Paraméternek tekinthető a cinkeltség mértéke, mivel egy szabálytalan, cinkelt érme esetén nem azonos eséllyel kapunk fejet vagy írást. Matematikailag az érmedobás, mint minden kétféle kimenettel bíró folyamat, Bernoulli-eloszlásból származó mintavételezéssel modellezhető. Egy Bernoulli-folyamat kimenetelét egyetlen paraméter határozza meg, melynek értéke azonos esélyek (szabályos érme) esetén 0.5.

A generatív modell tehát arra a kérdésre ad választ, hogy milyen eséllyel lesz az eredmény fej vagy írás, ha feldobok egy szabályos vagy egy ismert mértékben cinkelt érmét. Matematikailag ezt feltételes valószínűség segítségével önthetjük formába:  $P(\text{kimenetel} \mid \text{paraméterek})$ . Ez a kifejezés azt jelzi, hogy az egyes kimenetek valószínűsége függ a mögött szereplő változók (feltételek) értékétől. Példánk esetén pl. kiszámolható, hogy  $P(\text{fej} \mid \text{szabályos érme})=0.5$ , de

$P(\text{írás} \mid 0.6\text{-ra cinkelt érme})=0.4$ . Fontos megjegyezni, hogy egy ilyen kifejezés függvényként többféleképpen értelmezhető. Ha az adott modellben a paraméterek értékét fixen tartjuk, akkor a kimenetelt véve változónak, valóban valószínűségeket (az angolban probability-t) számolunk. Ha az összes behelyettesítés esetén kapott értékeket összeadjuk, 1-et kapunk, mivel ezek az események egymást kizárják, s lefedik a lehetséges események terét – valamelyik mindenképpen bekövetkezik:  $P(\text{fej} \mid \text{paraméter}) + P(\text{írás} \mid \text{paraméter}) = 1$ . Vagyis  $P(x \mid \text{paraméter})$  egy valószínűségi eloszlás. Ehhez hasonló, ettől mégis élesen elkülöníthető az, ha valamelyik kimenetelt rögzítjük előre, és azt szeretnénk megtudni, hogy ez melyik paraméter esetén milyen valószínűséggel következik be. Angol mintára ezt a függvényt likelihood-nak nevezzük, és a következőképpen írható le:  $P(\text{kimenetel} \mid x)$ . Ez nem valószínűségi eloszlás, mivel, ha összegezzük a lehetséges  $x$ -ekre kapott értékeket, nem feltétlen lesz 1 az eredmény.

#### **Az inverz probléma: vajon mennyire cinkelt az érménk?**

Általában nem tudjuk egy érméről ránézésre, hogy cinkelt-e, mégis megállapíthatjuk azt mintavételezés segítségével, a generatív modell ismeretében, hogy mennyire kéne biztosnak lenni abban, hogy szabályos. Ezt hívjuk inverz problémának, amikor a minta (vagyis pl. érmedobások) és a generatív modell (pl. az egyes érmefeldobások  $p$  paraméterű Bernoulli-eloszlásból, így a kapott fejek száma 100 érmefeldobás esetén  $N=100$  és  $p$  paraméterű binomiális eloszlásból származik) alapján következtetünk a paraméterek értékeire (vagyis pl. a cinkeltség mértékére,  $p$ -re). Ha például 100-szor feldobjuk az érmét, és ebből csak 30 alkalommal lett fej, akkor legtöbbször cinkeltségre gyanakodna (mivel bár ez előfordulhat szabályos érme esetén is, azt sejtjük, hogy csupán kis valószínűséggel). Ahhoz, hogy megtudjuk, matematikailag mekkora esély is van erre (hogy egy szabályos érmével 100-ból 30 esetben dobjunk fejet), kiszámolhatjuk a Binomiális( $x, 100, p$ ) eloszlás szerinti valószínűséget a mintában található  $x=30$  fej és  $p=0.5$  helyettesítéssel. Ennek értéke megközelítőleg 0.00002, vagyis tényleg kicsi az esély rá. A likelihood függvény maximumhelyét megkeresve megállapíthatjuk, hogy a legvalószínűbb paraméter, ami mellett ilyen eredménye lett volna a dobásainknak, a 0.3. Ha azonban erős az az előzetes elvárásunk, hogy az érme szabályos (0.5 „cinkeltségű”), akkor egy 100 elemű minta önmagában nem eléggé meggyőző, hiszen tudjuk, hogy a relatív

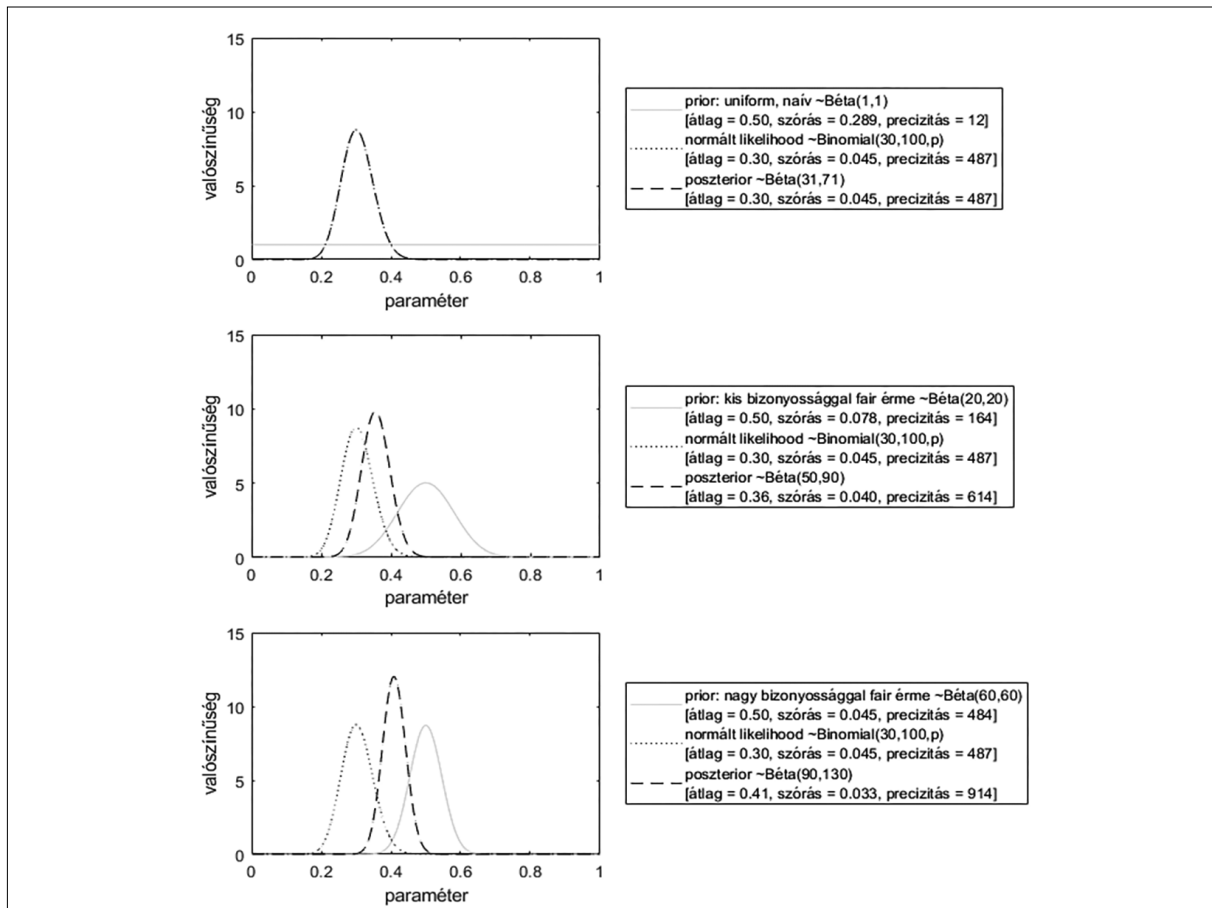
gyakoriság a pontos valószínűségekhez csak végtelen sok mintavétel után tart. Ezért a kísérlet elvégzése után a cinkeltséget vélhetően 0.3 és 0.5 közé becsülénk. Ez a gondolatmenet is precízen formalizálható, ha segítségül hívjuk a Bayes-szabályt.

#### **A Bayes-szabály**

A Bayes-szabály (Efron, 2013) azt mondja ki, hogy egy változó (nevezzük A-nak) adott értékének bekövetkezési valószínűsége egy másik változó (legyen B) értékének ismeretében, arányos a másik változó (B) ismert értékének bekövetkezési valószínűségével az első változó (A) ismeretében; a rátát pedig az első (A) és a második (B) változó (feltétel nélküli) valószínűségeinek aránya adja. Vagyis:  $P(A|B) = P(B|A) * P(A) / P(B)$ . Az inverz probléma szerinti behelyettesítés elvégzésével a következő formulát kapjuk:  $P(\text{paraméter} \mid \text{minta}) = P(\text{minta} \mid \text{paraméter}) * P(\text{paraméter}) / P(\text{minta})$ . A mintát ismerjük, ez az érmefeldobásaink eredménye. Ha az a cél, hogy megkeressük azt a paramétert, amely a minta alapján a legvalószínűbb, akkor a baloldalon szereplő  $P(x \mid \text{minta})$  függvény maximumhelyére (MAP, azaz maximum a posteriori becslés), de gyakran az eloszlás várható értékére és szórására vagyunk kíváncsiak (ahogy később látni fogjuk, a pontbecsléssel szemben ez a becslés bizonytalanságát is megragadja). Ez ekvivalens azzal, ha a jobboldali  $P(\text{minta} \mid x) * P(x) / P(\text{minta})$  függvénnyel tennénk ugyanezt. Érdemes észrevenni, hogy elhagyhatjuk a  $P(\text{minta})$  nevezőt, mivel ez nem függ a paramétertől, így csak magát a maximumértéket befolyásolja, azt nem, hogy azt a függvény hol veszi fel, s jelen esetben csak ez számít. A megmaradt tényezők közül az egyik a likelihood:  $P(\text{minta} \mid x)$ ; a másikat pedig prior valószínűségnek nevezzük, mivel ez utóbbi az adatgyűjtést megelőző tudásunkat tartalmazza a paraméter lehetséges értékeinek valószínűségéről:  $P(x)$ . A bal oldalon lévő kifejezésre poszterior valószínűségként hivatkozunk, mivel ez a mintavételezést követő vélekedésünket foglalja magában:  $P(\text{paraméter} \mid \text{minta})$ . A poszterior eloszlás tehát arányos a likelihood és a prior eloszlások pontonkénti szorzatával:  $P(\text{paraméter} \mid \text{minta}) \propto P(\text{minta} \mid \text{paraméter}) * P(\text{paraméter})$ .

A példafeladatban arra kerestük a választ, hogy melyik paraméter esetén maximális a  $P(\text{paraméter} \mid 100 \text{ dobásból } 30 \text{ fej})$  poszterior, és most már tudjuk, hogy ehhez a  $P(100 \text{ dobásból } 30 \text{ fej} \mid \text{paraméter})$  likelihoodot és a  $P(\text{paraméter})$  priort elég tudni. A likelihood egyértelműen adódik a generatív folyamatból, korábban kiszámoltuk a Binomiális(30, 100,  $p$ ) eloszlás alapján. A prior azonban nem ilyen egy-

**1. ábra** A prior eloszlás precizitása meghatározza a poszterior alakját. Naiv prior esetén a poszterior egybeesik a likelihooddal, a precízebb priorok viszont fokozatosan maguk felé vonzzák az eloszlás csúcsát. Az illusztráció a MATLAB programmal készült.



értelmű. Van, hogy nincs prior tudásunk, akkor ezt a tényezőt el is hagyhatjuk, és kizárólag a likelihoodot maximalizáljuk. Ez ugyanazt az eredményt hozza, mintha lenne 'naiv' tudásunk, miszerint a keresett paraméter 0 és 1 közötti, de ezen belül mind ugyanakkora valószínűséggel lehetséges (ld. 1. ábra felső diagram). De formalizálható az is, hogy fair értémet feltételezünk, ekkor az eloszlás csúcsa 0.5-ben kell legyen. Mivel továbbra is csak 0 és 1 közötti értéket keresünk, a béta-eloszlás lehet jó megoldás. Az eloszlások maximumhelye vagy várható értéke önmagában csak pontbecslés, amihez a szórás fontos többletinformációt ad: ez mutatja a becsléssel kapcsolatos bizonytalanság mértékét. Ha eleve nyitottak voltunk a mintavétel előtt arra, hogy nem pontosan fair az érme, akkor a prior eloszlás „lapos”, ilyenkor a 0.5 nem sokkal valószínűbb, mint a környezete (ld. 1. ábra középső diagram). Minél biztosabbak voltunk abban, hogy fair az érme, annál „csúcsosabb” a prior eloszlás (ld. 1. ábra alsó diagram). A likelihood szórást az elemszám határozza meg: minél több mintát

vettünk, annál szűkebb az eloszlás, és intuitívan annál biztosabbak lehetünk abban, hogy a minta alapján jó becsléseket tudunk végezni.

Figyeljük meg az 1. ábrán, hogy a poszterior eloszlás paramétereit hogyan befolyásolja a likelihood és a prior szórása! Mivel a szórás valamelyest konstraintív (minél nagyobb, annál kevésbé vagyunk biztosak), szokás a precizitást (precision), vagyis a szórásnégyzet inverzét is megadni (ez minél nagyobb, annál biztosabbak vagyunk a becslt értékben). A poszterior eloszlás maximumhelye, illetve várható értéke minden esetben a likelihood és a prior eloszlások értékei között helyezkedik el, még hozzá a precízebb eloszláshoz közelebb. Ha nem érmedobásokról van szó, hanem valamilyen folytonos eloszlású változó értékeiről, akkor igen gyakran normális eloszlásokat feltételezhetünk, és ezek esetén a poszterior maximumhelye – amely az eloszlás szimmetriája miatt egyben a várható érték is - pontosan a prior és likelihood átlagok precizitásával (inverz varianciával) súlyozott átlaga (Dienes, 2013).

A Bayes-szabály matematikai megfogalmazásban absztraktnak tűnhet, de valójában az orvosok – akár tudtukon kívül, de mégis – rendszeresen alkalmazzák a gyakorlatban, ahogy erre Hall már a XX. század közepén rávilágított (Hall, 1967). Amikor ugyanis egy páciens adott panaszokkal fordul hozzájuk, és az azt okozó betegséget igyekeznek megállapítani, akkor az orvos feladata a [betegség  $\rightarrow$  tünetek] generatív modell inverz problémájának megoldása, azaz a  $P(\text{betegség}|\text{tünetek})$  poszterior eloszlás maximumhelyét keresi. A szakkönyvek a likelihood függvény kiértékelését teszik lehetővé, hiszen azt mondják meg, hogy a betegség ismeretében mennyire valószínű egyes tünetek jelenléte:  $P(\text{tünetek}|\text{betegség})$  (pl. a köhögés valószínű megfázás és tuberkulózis esetén is), valamint az egyes betegségek tünettől független (a priori) előfordulási valószínűségeit adják meg:  $P(\text{betegség})$  (pl. a megfázás gyakori, a tuberkulózis szerencsére ma már nagyon ritka). Ezeket a gyakoriságokat együtt kezelve (mintegy a valószínűségeket összeszorozva) állapítják meg a legvalószínűbb betegséget (pl. a köhögés alapján (poszterior) valószínűbbnek fogják ítélni a megfázást, mint a tuberkulózist). [Természetesen a valóságban több tünet alapján, és több betegség közül kell döntést hozni, ez egy egyszerű példa.]

Láthattuk, hogy eloszlásként reprezentálva az előzetes elvárásokat (prior) és az új, mintavételezett információt (likelihood), a Bayes-szabály segítségével kiszámítható a poszterior eloszlás, amellyel az előzőek alapján levont következtetés jellemezhető. Igaz ez olyan egyszerű példák esetén, mint az érmefeldobás (cinkeltég megállapítása), vagy a betegségek diagnosztizálása eltérő gyakoriságú tünetek alapján, ugyanakkor ez a keretrendszer az elméműködés összetettebb folyamatainak megértéséhez is hozzájárulhat.

## A BAYESI MODELLEK ALKALMAZÁSA ÉS A PREDIKTÍV KÓDOLÁS

### *Bayesiánus folyamatok a kognitív idegtudomány fókuszában*

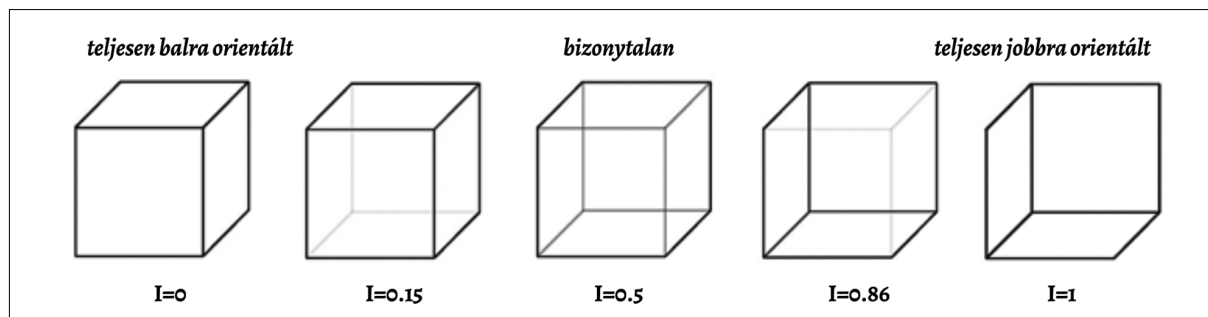
Egyre szélesebb körben elfogadott az a szemlélet, miszerint az agy a tapasztalatok alapján generatív modelleket alkot a környezet szabályosságairól, s az észlelés során valójában ennek segítségével következtet az érzékek háttérben álló jelenségekre (Friston, 2010). A korábban bevezetett fogalmakat alkalmazva: az észlelet nem más, mint a poszterior becslés, melyet a korábbi tapasztalatokon alapuló prior elvárások

és az aktuálisan beérkező információkból kialakult érzéklet, a likelihood becslés alapján számol elmenk. Általános feltételezés továbbá, hogy a környezet (vagy egy adott kognitív feladat) efféle absztrakt reprezentációja folyamatosan frissítésre kerül, mivel az ágensek célja minél pontosabb predikciók megalkotása - vagyis folyamatosan tanulunk.

A különböző percepcióhoz kötődő tanulási folyamatokat klasszikusan kétféle kategóriába sorolták: egyrészt megvalósulhat perceptuális tanulás, amikor sok próbán át, visszajelzések révén elsajátítjuk a hasonló szenzoros információk diszkriminációjához szükséges információkat, míg a statisztikai tanulás az addig ismeretlen szenzoros élmények, gyakori téri vagy időbeli mintázatok akár egészen komplex reprezentációjának kialakulásához vezető automatikus, implicit mechanizmus. Ahogy azonban Fiser és Lengyel rámutatnak (2019), a két folyamat valójában leírható egy egyesített, bayesiánus keretrendszerben. Az efféle probabilisztikus tanulás ismérve, hogy az információfeldolgozást a szenzoros „bottom-up” (alacsonyabb szintről, részletekből kiinduló) és az eseményalapú „top-down” (magasabb szintről, komplexebb elképzeléseken alapuló) hatások relatív bizonytalanságukon alapuló kombinálásaként írja le. Ennek hierarchikus kiterjesztései a Hierarchikus Bayesiánus Modellek, amelyben a szenzoros evidenciákat egyrészt a megjelenített stimulus és az annak érzékelését befolyásoló, egyénre jellemző szenzoros paraméterek befolyásolják (a perceptuális tanulás ezeken tud érvényesülni), de eközben a megjelenített stimulus maga is egy generatív folyamat eredménye, amit egy készlet, s az arra jellemző probabilisztikus szabályosságok határoznak meg (itt érvényesül a statisztikai tanulás).

A prediktív kódolási elmélet ezt úgy egészíti ki, hogy a frissítés hajtóerejének a predikciós hibát, vagyis a predikciók és az input eltérését tekinti, mivel optimális működés során ennek hosszú távú minimalizálására törekszünk (Aitchison & Lengyel, 2017). A prediktív kódolás alapvetően egy algoritmus, segítségével a (neurális) reprezentációk szintjén írható le (Marr algoritmikus szintjén) az, hogy hogyan keletkezik, majd kerül felhasználásra a predikciós hiba. Amikor bayesiánus következtetés a cél, a predikció kiszámítása a Bayes-szabály segítségével történik. Fontos megjegyezni, hogy a kettő létezhet a másik nélkül, vagyis prediktív kódolással megvalósíthatóak más komputációs célok, illetve a bayesiánus következtetés is létrejöhet más reprezentációkon keresztül, de a tanulmányban később kiemelt vizsgálatok ebben a közös keretrendszerben értelmezhetőek.

**2. ábra** Bistabil észlelés. A Necker-kocka egyes változatai a szenzoros információk (likelihood) alapján értelmezhetők egy kocka felül- vagy alulnézeti képeként (poszterior), de a tökéletesen bizonytalan középső ábrát általában inkább felülnézetként értelmezzük, mivel többször találkozunk ilyen alakokkal (prior). (Maksimenko és mtsai., 2017)



Számos tanulmány rámutatott arra, hogy nemcsak az egyes modalitásokból érkező információ észlelése/értelmezése (Fiser és mtsai., 2010; Petzschner és mtsai., 2015; Yuille & Kersten, 2006), hanem a multimodális integráció (Alais & Burr, 2004; Ernst & Banks, 2002), a nyelvhasználat (Brent, 1996), egyes motoros funkciók (Heald és mtsai., 2021; Nagengast és mtsai., 2009), sőt, memória- figyelmi- és tanulási folyamatok (Houlsby és mtsai., 2013; Norman és mtsai., 2008; Orbán és mtsai., 2008; Shettel és mtsai., 2006) is jól modellezhetőek bayesiánus formulák segítségével. De nézzük meg kicsit közelebbről a vizuális percepció vizsgálatát!

### A vizuális percepció vizsgálata

Könnyen látható, hogy a látórendszernek folyamatos feladata a vizuális modalitást jellemző kétértelműség/bizonytalanság feloldása, mivel a háromdimenziós környezet retinális képe már csak kétdimenziós vetület, amely nem injektív leképezés (a térbeli elrendezésből egyértelműen következik a vetület, de egy vetület több különböző térbeli elrendezésből származhat, ld. 2. ábra). Mivel tehát a szenzoros információ (likelihood) önmagában nem magyarázza az egyértelmű (poszterior) észleletet, amelyet általában megélünk, logikus levonni a következtetést, miszerint ezt a folyamatot előzetes, prior tapasztalataink segítik.

A priorok észlelésre gyakorolt hatására a leglátványosabban olyan helyzetekben derül fény, amikor a szenzoros információ különösen zajos. Ez a hétköznapokban jelentheti azt, hogy például félhomályban tévesen ítélünk meg alakokat, de különböző vizuális illúziók, szándékosan kétértelmű ingerek esetén is „becsaphat minket a szemünk”. Vegyük például Schmack és mtsai. (2013) tanulmányát, akik azt

vizsgálták, hogy a téveszmékre való hajlam hogyan függ össze az észlelés eltéréseivel. Kísérletük vizuális ingeranyaga olyan mozgó pontok felhőjéből állt, amelyek egy gömbfelszín forgásának illúzióját keltették úgy, hogy a pontok egy része az egyik, míg a többi a másik irányú forgást szimulálta – a résztvevők időnként változó, de adott ideig mindig egyirányú forgás észleléséről számoltak be. Amikor azt az instrukciót kapták, hogy olyan szemüveget viselnek, amely befolyásolja az irányok észlelését, akkor nemcsak a tanulási fázis alatt számoltak be többször arról, hogy a pontok nagyobb százaléka által követett irányt észlelték, hanem a teszt fázis alatt is, holott utóbbi esetben kiegyensúlyozott volt az inger (a téveszmékkel kapcsolatos eredményekről később lesz majd szó). A kétértelműség előidézése mellett egy másik lehetőség a stimulusok zajosságának növelése, ezzel csökkentve a likelihood precizitást. Tekintsük például Lawson és mtsai. (2017) kísérletét, ahol a résztvevőknek nagyon, közepesen vagy egyáltalán nem zajos képekről kellett döntést hozniuk: házat vagy arcot látnak. Itt a priorok kialakítását kulcsingerek segítették: egy alacsonyabb vagy magasabb hangot játszottak le a képek előtt, melyek probabilisztikus viszonyban álltak az utánuk megjelenő képekkel (pl. magas hangot általában ház követett). A reakcióidő adatokból kiolvasható volt, hogy a kísérleti személyek lassabban válaszoltak a váratlan (kisebb valószínűségű), mint a várható (nagyobb valószínűségű) kimenetek esetén, továbbá a zajosság mértékének növelésével is megnövekedett reakcióidőket tapasztaltak – látható tehát, hogy mind a likelihood, mind a priorok befolyásolása mérhető viselkedésbeli eltéréseket eredményez. A kihívást az jelenti, hogy ezeket megfelelően integráljuk a modellbe.

**Percepciókutatás, avagy a megfigyelő megfigyelése**

Amikor kutatóként a (vizuális) percepció során zajló agyműködést szeretnénk megfejteni, kettős bayesiánus formulákkal érdemes dolgozni, mivel ilyenkor lényegében megfigyeljük a megfigyelőt (Daunizeau és mtsai., 2010). Ugyanis, ahogy a kiváló idegtudományi kutatók rámutatnak tanulmányukban, ahhoz, hogy egy kognitív feladat során szerzett viselkedéses vagy neurofiziológiás adatból következtetéseket vonhassunk le az alany agyában zajló folyamatokról, két, egymással szoros kapcsolatban álló modellt is meg kell alkotni. Egyrészt ki kell dolgozni az adott feladatra jellemző úgynevezett perceptuális (észlelési) modellt, amely azt írja le, hogy a kísérleti személy, optimális bayesi megfigyelőként, hogyan dolgozza fel a bemutatott ingereket és együttjárásaikat. Ezen felül egy válaszadási modellre is szükség van, amely meghatározza a megfigyelésekre adott válaszokat, hiszen mi külső megfigyelőként csak az utóbbihoz férünk hozzá - általában gombnyomásokhoz és a reakció-időkhöz, de az agyi aktivitás mérései is ide tartoznak.

A válaszadási modell szoros kapcsolatban áll az észlelési modellel, ez jól látható abból is, hogy ha olyan ingerrel találkozunk, amelyre előzőleg számítottunk (nagy volt a prior valószínűsége), akkor gyorsabb és pontosabb reakciók várhatóak (Lawson és mtsai., 2017; Pálffy és mtsai., 2021). Minden aktivitás, minden cselekvés energiabefektetéssel jár, de az egyes célok (pl. gyors reakció az életben maradásért, helyes válasz a feladat során) elérése jutalmazó. Valamilyen veszteség-, vagy hasznosságfüggvénnyel (loss / utility function) formalizálható az, hogy egy adott állapotban (külső-belső környezet tulajdonságai alapján) mely cselekvés milyen kimenettel, s ezáltal értékkel bír. Az embereket optimális döntéshozónak tekintve feltételezhetjük, hogy a lehető legnagyobb hasznossággal (legkisebb veszteséggel) kecsgető akciót választják (Daunizeau és mtsai., 2010; Steyvers és mtsai., 2009). Ez az optimális cselekvés természetesen függ attól, hogy milyen állapotot tekintenek a legnagyobb valószínűségűnek (~valósnak) az adott pillanatban, amit elsősorban a perceptuális modell alapján számolt poszterior eloszlás határoz meg.

Már az egyik legegyszerűbb, de reakcióidő adatokra kiválóan illeszkedő válaszadási modell, a LATER (Carpenter, 1981; útmutatóért lásd Noorani & Carpenter, 2016) esetén is azt látjuk, hogy az adatokban jelenlévő variancia magyarázatára a szenzoros információ (prior szerinti) valószínűsége mellett további paraméterekre is szükség van (amik személyenként változhatnak). Komplex feladatok és inger-

anyagok esetén jellemzően a perceptuális modell is csak több paraméter felhasználásával konstruálható, így a kísérleti adatok elemzésekor a poszterior valószínűség  $P(\text{perceptuális modell paraméterek, válaszadási modell paraméterek} \mid \text{válasz adatok, stimulusok})$  alakú. Ennek kiszámítása technikailag lényegesen nehezebb, mégis elvi háttérét tekintve a korábban említett cinkeltségi paraméter kereséséhez hasonló feladat: likelihood és prior eloszlások alapján tesszük. Jelentős eltérést az jelent, hogy míg a példában a prior és likelihood eloszlások alakja miatt a poszterior eloszlás is nevezetes volt, ezért könnyedén paraméterezhető és kiértékelhető, addig az összetettebb modelleknél ez ritkán valósul meg, s az inverz probléma nem oldható meg approximációs, azaz közelítő algoritmusok használata nélkül. Továbbá felhívnánk a figyelmet arra, hogy az érmedobási példához képest további komplexitást jelent az, hogyha fokozatos tanulást feltételezünk. Ilyenkor a háttérben álló paramétereket minden egyes beérkező információ (pl. érmedobás) után inferálni kell, ezt gyakran hierarchikus struktúra bevezetésével oldják meg. Ezen részletek nagyobb mélységben való megvitatása nem célja ezen tanulmánynak, de az érdeklődő olvasó számára javasoljuk Daunizeau, illetve Mathys munkáit (Daunizeau és mtsai., 2010; Mathys és mtsai., 2011).

**Egyének közötti eltérések – klinikai relevancia**

A komputációs modellezés egyik célja olyan normatív modellek feltárása, melyek alapján általános következtetések vonhatóak le az elme működéséről, de eszköztára ezen túlmutat: csoportok vagy egyének közötti különbségek feltárására is alkalmas. Különös jelentőséggel bír ez az észlelés területén, ahol a szenzoros információk és a prior hiedelmek súlyozásában igen nagy egyéni különbségek vannak, amik nem csak megbízhatóan mérhetőek (Pálffy és mtsai., 2021), de egyesek szerint bizonyos pszichiátriai kórképek magtüneteinek háttérében is ennek a rendszernek az átlagostól eltérő működése állhat. A következőkben a pszichotikus tünetek egy specifikus magyarázatával fogunk foglalkozni, és csak megjegyezzük, hogy ez a szkizofrénia és a pszichózisok komputációs pszichiátriai irodalmának csak egy szelete, ráadásul ezen a területen belül is sok kérdés nyitva van még. A negatív tüneteket például a megerősítéses tanulás különféle eltéréseihez kötötték, valamint a pozitív tünetek magyarázatára is egy sor további elmélet született (pl. dopamin, aberráns jelentőségadás és jutalompredikációs hiba; gátló interneuronok károsodása és körkörös következtetés), a kognitív tünetek magyarázataként



pedig felmerült, hogy a prefrontális kéregben a serkentés/gátlás egyensúlya eltolódik, ami miatt a neurális hálók kevésbé stabilak (Adams, 2020).

A pszichózisban érintettek érzékcsalódásai, illetve téveszméi a bayesiánus prediktív kódolási elmélet Sterzer és mtsai. (2018) szerint egy többszintű modellben, a likelihood információ (alacsonyabb szintű prior elvárásokhoz viszonyított) nagyobb precizitásával magyarázhatók. Ez ugyanis azt okozza, hogy az észlelet, a poszterior becslés közelebb lesz a szenzoros bemenethez, mint az eredeti elváráshoz (pl. a vizuális illúzióknak kevésbé dőlnek be a szkizofréniával diagnosztizált személyek (Dima és mtsai., 2009)), ami által a későbbi (mindig az előző poszteriorból formálódó) priorok megbízhatatlanok, pontatlanok lesznek. Az elmélet szerint a téveszmék kialakulásáért pontosan ezek a szenzoros információ értelmezéséhez szükséges alacsonyabb szintű (perceptuális, az érzékletek alapvető tulajdonságaihoz kapcsolódó) prior elvárások pontatlansága felelős, mivel így a perceptuális zaj feloldásához magasabb szintű (konceptuális, az érzékeltekhez kevésbé kapcsolódó) prior hiedelmek szükségesek. A pszichózisban jellemző másik percepcióhoz kapcsolódó tünet, a hallucinációk megjelenése e modell szerint azzal magyarázható, hogy az erős magasabb szintű elvárások felülírják a szenzoros információt. Gyakori tünet továbbá pszichózisban a befolyásoltatás érzete, amikor a betegek a saját maguk által kivitelezett akcióknak, gondolatoknak külső forrást tulajdonítanak. Ez szintén magyarázható az alacsonyabb szintű priorok (efferens kópiák) csökkent súlyozásával.

Már említettük azt a kutatást, melyben egy pontfelhő forgásirányának észlelését befolyásolta az, hogy az instrukciók (és egy tanulási fázis) szerint az egyik irányt erősítő szemüveget viseltek a kísérleti személyek, holott az inger kiegyensúlyozottan kétértelmű volt. A kutatók itt azt is megfigyelték, hogy szkizofréniával élő, téveszméket megelő embereknél ez a hatás a kontrollcsoportnál nagyobb mértékben jelent meg (Schmack és mtsai., 2015), továbbá ez jellemezte azokat is, akik ugyan nem rendelkeztek diagnózissal, de kérdőíves válaszaik alapján téveszmékre hajlamosabbak voltak (Schmack és mtsai., 2013). Ez beleillik az elméletbe, miszerint a téveszmék együttjárnak a magasabb szintű prior elvárások erősebb súlyozásával. Az, hogy ugyanezen személyek a tanulás előtt gyakrabban jelezték az észlelt irány megváltozását is, azt jelzi, hogy az - aktuális percepcióval ellentétes irányba mozgó pontok révén kialakult - predikciós hiba hatása valóban erősebben érvényesült náluk, s a magasabb szintű elvárások (szemüveghez társított

irány) nélkül az észlelésüket elsősorban a likelihood információ határozta meg.

A pszichózis mellett az autisztikus észlelést is lehetséges a prediktív kódolás keretrendszerében értelmezni. Pellicano és Burr (2012) amellett érvelnek, hogy a szenzoros információ precizitásának felülértékelése, ezáltal a prior elvárások kisebb befolyásoló ereje révén az autizmussal élő személyek valójában a világot a neurotipikus személyekhez képest jobban úgy látják, amilyen az valójában. Ez előnyt is nyújthat, mivel jobban teljesítenek olyan feladatokban, ahol a kontextus zavart okoz (pl. egy képben elrejtett alakzatok keresése), ugyanakkor hátrányt jelent akkor, ha a kontextus segíthetne feldolgozni a beérkező információkat (pl. a tárgyak árnyékai számukra csak a zaj mértékét növelik, nem segítik a felismerést). Mivel a poszterior meghatározza a jövőbeli prior elvárásokat, ezért, ha a poszterior jobban befolyásolja az aktuális szenzoros információt, mint a korábbi tudás, akkor nehéz lesz megtalálni az általános összefüggéseket. Ehelyett a poszterior túlságosan tükrözni fogja a szenzoros adatok esetlegességeit, hiszen az autizmussal élő személyek felülértékelik a szenzoros adatok pontosságát, azaz alábecsülik a szenzoros zajt. Tanulás szempontjából ez azt jelenti, hogy a tapasztalat-függő plaszticitás nehezebben valósul meg számukra, így rövidtávon kevésbé torzítja az észlelésüket egy korábbi stimulus (pl. egy arc után egy másik), de a priorok hosszabb távú segítő hatása is elmarad, mivel a predikciós hibák túlzott súlyozása révén mindig egy adott környezethez kalibrálnak (mintha minden nap egy új világban ébrednének fel).

Elméletük vizsgálatára Lawson és mtsai. (2017) a már korábban említett perceptuális kategorizációs feladatokat úgy alakították ki, hogy a kulcsingerek prediktív erejét szisztematikusan változtatva a szenzoros zajon túl úgynevezett probabilisztikus bizonytalanságot, valamint környezeti bizonytalanságot is bevezettek. Egy ilyen helyzetben egy váratlan kimenetel értelmezhető „szabályba illő”, de alacsony valószínűségű eseményként, vagy a környezet megváltozásaként is. Az adatokra legjobban illeszkedő perceptuális (és ezen keresztül a válaszadási) modelljük hierarchikus szerkezetű volt, ahol - a látórendszerhez hasonlatos módon - a magasabb szinteken kerültek kiszámításra az alacsonyabb szintek predikciói, amelyek az alsó szintre érkező új információk, pontosabban az esetlegesen fellépő predikciós hibák révén dinamikusan frissültek. A modellben az alsó szinten a döntéssel kapcsolatos bizonytalanság jelent meg (pl. milyen valószínű, hogy házat mutatnak?). Ennek precizitását (ezáltal a predikciós hiba súlyát) a következő

szint határozta meg, ahol az asszociációk reprezentálódtak (pl. milyen valószínű, hogy alacsony hang után házat mutatnak?). Ennek precizitását pedig a harmadik, legfelső szint határozta meg, ahol a környezeti bizonytalanság reprezentációja volt (milyen valószínű, hogy megváltozott a szabály?). A kutatás eredményeképp azt találták, hogy az autizmussal élő személyek számára kevésbé volt meglepő a váratlan kimenetel, mint a normál populációból vett minta számára (nem válaszoltak lassabban ezeken a próbákon), amelynek háttérében az áll, hogy számukra a priorok precizitása alacsonyabb volt (nem volt erős elvárásuk a másik alakzatra). Az eltérés az illesztett modell felső szintjén jelent meg: az autizmussal élők nagyobbra becsülték a környezeti bizonytalanságot, a kisebb valószínűségű kimenetek számukra nem a probabilisztikus szabálynak mondtak ellen, hanem azt a következtetést vonták le, hogy megváltozott a környezet, új szabály jelent meg.

Az eddigi kutatások tehát abba az irányba mutatnak, hogy a vizuális percepció jól vizsgálható a bayesiánus logikát követve. A likelihood információ kísérletesen manipulálható, ha az alkalmazott ingeranyag kétértelmű vagy zajos. Ezen felül instrukciók és/vagy kulcsingerek révén asszociációk, ezáltal priorok kialakítása is lehetséges (ezek az egyénre jellemző, rejtett paraméterek). A perceptuális és válaszadási modellek segítségével kognitív feladatokon nyújtott teljesítmény alapján azonosíthatók olyan paraméterek, amelyek rávilágítanak arra, hogy milyen információfeldolgozási deficit áll az egyes klinikai populációk észlelési eltéréseinek háttérében.

## ZÁRÓ GONDOLATOK ÉS LIMITÁCIÓK

A hatékony orvoslás egy fontos feladata a fenotipikus tünetek és a háttérben álló biológiai mechanizmusok közötti kapcsolat feltárása, amely számos pszichiátriai kórkép esetén napjainkban is megoldásra váró probléma. A heterogeneitás, amely egyes betegségkategoriókat jellemez mind genetikai, idegrendszeri, kognitív és tüneti szinten (Joyce & Roiser, 2007), motivációt ad arra, hogy megtaláljuk a közös varianciában megnyilvánuló összefüggéseket (Insel, 2010). A tünetek háttérében álló folyamatok, befolyásoló faktorok ismeretében nyílik lehetőség a személyre szabott orvoslásra, a hatékony intervencióra és kezelésre (Kapur és mtsai., 2012).

A komputációs pszichiátria egy ágazata kognitív fenotípusok azonosításával olyan átfogó mechanisztikus magyarázatot ajánl, amelyek ellenőrzése a kutató-

sok széles spektrumán át lehetséges (és szükséges). A bayesiánus prediktív kódolási elmélet például biológiai rendszerek optimális működését alapul véve írja le az információfeldolgozás normatív modelljét. Továbbá, a rendszer alacsonyabb, illetve magasabb szintjén történő eltérésekkel magyarázható az, hogy miért észleli valaki máshogy a környezetét, mint ami az átlag populációban megszokott. Amikor például Lawson és mtsai. (2014) arra hívják fel a figyelmet, hogy optimális inferencia nem valósulhat meg, ha a szenzoros likelihood és/vagy a prior vélekedések precizitását rosszul becsüljük meg, akkor ez a komputációs szinten megfogalmazott állítás, amely szisztematikusan vizsgálható kognitív tesztekre adott válaszok elemzésével. Ezen elvek mentén magyarázhatók a téveszmék és hallucinációk, mint felerősödött magasabb szintű priorok, amelyek elnyomják a szenzoros információkat; vagy az autisztikus észlelés, szenzoros túltelítődés, amikor az ismeretlen, új információk által kiváltott predikciós hibák kórosan nagy súllyal jeleznek a felsőbb szintek számára. Az elméleti keretbe illeszthetők a folyamatokat implementáló idegrendszeri mechanizmusok is: a szenzoros információ túlzott súlyozását okozhatja a top-down zajcsökkentő jelek csillapodása, neurotranszmitterek egyensúlyának felborulása miatt (Wood és mtsai., 2021).

Limitációként azonban meg kell említeni, hogy bár a felmutatott problémára a keretrendszer magyarázó ereje igen magas, nem feltétlenül zár ki más lehetőségeket. Mint már korábban írtuk, maga a bayesiánus következtetés, mint optimális komputációs cél, megvalósulhat a prediktív kódolástól eltérő algoritmusokon keresztül is (ahogy utóbbi célja sem minden esetben az előbbi megvalósítása) (Aitchison & Lengyel, 2017).

Általánosságban elmondható, hogy a komputációs megközelítés alapvetően az elme modellezését jelenti, célja normatív számítási lehetőségek megtalálása olyan problémákra, amelyekkel az idegrendszer maga is rendszeresen megbirkózik. A statisztikailag optimálistól azonban valamelyest eltérhet a valós viselkedés. Az egyének vagy csoportok közötti eltérésekre kíváncsi kutató feladata elméletileg megalapozott modellt alkotni, majd ehhez megtalálni azt a feladatot, módszert, amely révén számolhatóak a hipotézisekhez köthető paraméterek. Nem szabad azonban elfelejteni azt, hogy noha ezeket a paramétereket képletek írják le, általában közelítő eljárások révén kerülnek kiszámításra, és ugyanúgy fontos ellenőrizni a reliabilitásukat és validitásukat, mint bármely más mérőeszköz esetén (Brown és mtsai., 2020; Moriarity & Alloy, 2021).

**KÖSZÖNETNYILVÁNÍTÁS**

Az Innovációs és Technológiai Minisztérium ÚNKP-21-3 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült (P.Zs.). A munka a BME Biotechnológia projekt (FIKP-BIO) és az NKFI/OTKA K 128599 támogatásával készült el (P.B.). Kiemelten köszönjük Orbán Gergő kéziratával kapcsolatos meglátásait, szakmai támogatását, valamint az anonim bíráló javaslatait.

**LEVELEZŐ SZERZŐ:** Pálffy Zsófia

Kognitív Tudományi Tanszék, Természettudományi Kar,  
Budapesti Műszaki és Gazdaságtudományi Egyetem,  
H-1111 Budapest, Műegyetem rkp. 3.  
E-mail: zsofiapalfy@edu.bme.hu

**AJÁNLOTT BEVEZETŐ OLVASMÁNYOK**

1. Komputációs pszichiátriai bevezető könyv: Series, P. (Szerk.). (2020). *Computational Psychiatry: A Primer*. The MIT Press.
2. Az idegtudományi modellalkotás javasolt lépései: Blohm, G., Kording, K. P., & Schrater, P. R. (2020). A How-to-Model Guide for Neuroscience. *eNeuro*, 7(1).  
<https://doi.org/10.1523/ENEURO.0352-19.2019>
3. Matematikai jellegű bevezető a szabad energián alapuló modellezésbe: Bogacz, R. (2017). A tutorial on the free-energy framework for modelling perception and learning. *Journal of Mathematical Psychology*, 76, 198–211.  
<https://doi.org/10.1016/j.jmp.2015.11.003>
4. Mire érdemes figyelni viselkedéses adatokra illesztett komputációs modell alkotásakor: Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife*, 8, e49547. <https://doi.org/10.7554/eLife.49547>

**IRODALOM**

1. Adams, R. A. (2020). 6 Psychosis and Schizophrenia from a Computational Perspective. In P. Series (Szerk.), *Computational Psychiatry: A Primer*. The MIT Press.
2. Aitchison, L., & Lengyel, M. (2017). With or without you: Predictive coding and Bayesian inference in the brain. *Current Opinion in Neurobiology*, 46, 219–227.  
<https://doi.org/10.1016/j.comb.2017.08.010>
3. Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, 14(3), 257–262. <https://doi.org/10.1016/j.cub.2004.01.029>
4. World Psychiatric Association, A. P. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5\*)*. American Psychiatric Pub.
5. Borsboom, D., Cramer, A. O. J., Schmittmann, V. D., Epskamp, S., & Waldorp, L. J. (2011). The Small World of Psychopathology. *PLOS ONE*, 6(11), e27407.  
<https://doi.org/10.1371/journal.pone.0027407>
6. Brent, M. R. (1996). Advances in the computational study of language acquisition. *Cognition*, 61(1), 1–38.  
[https://doi.org/10.1016/S0010-0277\(96\)00779-2](https://doi.org/10.1016/S0010-0277(96)00779-2)
7. Brown, V. M., Chen, J., Gillan, C. M., & Price, R. B. (2020). Improving the Reliability of Computational Analyses: Model-Based Planning and Its Relationship With Compulsivity. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 5(6), 601–609. <https://doi.org/10.1016/j.bpsc.2019.12.019>
8. Brugger, S. P., Angelescu, I., Abi-Dargham, A., Mizrahi, R., Shahrezaei, V., & Howes, O. D. (2020). Heterogeneity of Striatal Dopamine Function in Schizophrenia: Meta-analysis of Variance. *Biological Psychiatry*, 87(3), 215–224.  
<https://doi.org/10.1016/j.biopsych.2019.07.008>
9. Buzsáki, G. (2020). The Brain–Cognitive Behavior Problem: A Retrospective. *eNeuro*, 7(4), ENEURO.0069-20.2020.  
<https://doi.org/10.1523/ENEURO.0069-20.2020>
10. Carpenter, R. H. S. (1981). *Oculomotor Procrastination*. In *Eye Movements*. Routledge.
11. Daunizeau, J., Ouden, H. E. M. den, Pessiglione, M., Kiebel, S. J., Stephan, K. E., & Friston, K. J. (2010). Observing the Observer (I): Meta-Bayesian Models of Learning and Decision-Making. *PLOS ONE*, 5(12), e15554.  
<https://doi.org/10.1371/journal.pone.0015554>
12. Dienes Z. (2013). Mitől tudomány a pszichológia? (1.). Akadémiai Kiadó Zrt.
13. Dima, D., Roiser, J. P., Dietrich, D. E., Bonnemann, C., Lanfermann, H., Emrich, H. M., & Dillo, W. (2009). Understanding why patients with schizophrenia do not perceive the hollow-mask illusion using dynamic causal modelling. *NeuroImage*, 46(4), 1180–1186. <https://doi.org/10.1016/j.neuroimage.2009.03.033>
14. Efron, B. (2013). Bayes' Theorem in the 21st Century. *Science*, 340(6137), 1177–1178. <https://doi.org/10.1126/science.1236536>
15. Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429–433. <https://doi.org/10.1038/415429a>
16. Fiser, J., Berkes, P., Orbán, G., & Lengyel, M. (2010). Statistically optimal perception and learning: From behavior to neural representations. *Trends in Cognitive Sciences*, 14(3), 119–130.  
<https://doi.org/10.1016/j.tics.2010.01.003>
17. Fiser, J., & Lengyel, G. (2019). A common probabilistic framework for perceptual and statistical learning. *Current Opinion in Neurobiology*, 58, 218–228.  
<https://doi.org/10.1016/j.comb.2019.09.007>
18. Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.  
<https://doi.org/10.1038/nrn2787>
19. Hall, G. H. (1967). THE CLINICAL APPLICATION OF BAYES' THEOREM. *The Lancet*, 290(7515), 555–557.  
[https://doi.org/10.1016/S0140-6736\(67\)90514-4](https://doi.org/10.1016/S0140-6736(67)90514-4)
20. Heald, J. B., Lengyel, M., & Wolpert, D. M. (2021). Contextual inference underlies the learning of sensorimotor repertoires. *Nature*, 600(7889), 489–493.  
<https://doi.org/10.1038/s41586-021-04129-3>
21. Housby, N. M. T., Huszár, F., Ghassemi, M. M., Orbán, G., Wolpert, D. M., & Lengyel, M. (2013). Cognitive Tomography Reveals Complex, Task-Independent Mental Representations. *Current Biology*, 23(21), 2169–2175.  
<https://doi.org/10.1016/j.cub.2013.09.012>
22. Huys, Q. J. M., Guitart-Masip, M., Dolan, R. J., & Dayan, P. (2015). Decision-Theoretic Psychiatry. *Clinical Psychological Science*, 3(3), 400–421. <https://doi.org/10.1177/2167702614562040>
23. Insel, T. R. (2010). Rethinking schizophrenia. *Nature*, 468(7321), 187–193. <https://doi.org/10.1038/nature09552>

24. Joyce, E. M., & Roiser, J. P. (2007). Cognitive heterogeneity in schizophrenia. *Current opinion in psychiatry*, 20(3), 268–272. <https://doi.org/10.1097/YCO.0b013e3280ba4975>
25. Kapur, S., Phillips, A. G., & Insel, T. R. (2012). Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Molecular Psychiatry*, 17(12), 1174–1179. <https://doi.org/10.1038/mp.2012.105>
26. Keshavan, M. S., & Kaneko, Y. (2013). Secondary psychoses: An update. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, 12(1), 4–15. <https://doi.org/10.1002/wps.20001>
27. Krakauer, J. W., Ghazanfar, A. A., Gomez-Marín, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*, 93(3), 480–490. <https://doi.org/10.1016/j.neuron.2016.12.041>
28. Lawson, R. P., Mathys, C., & Rees, G. (2017). Adults with autism over-estimate the volatility of the sensory environment. *Nature neuroscience*, 20(9), 1293–1299. <https://doi.org/10.1038/nn.4615>
29. Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00302>
30. Lisanby, SH., Kohler, C., Swanson, CL., Gur, RE. (1998). Psychosis Secondary to Brain Tumor. *Seminars in Clinical Neuropsychiatry*, 3(1), 12–22.
31. Maksimenko, V. A., Runnova, A. E., Zhuravlev, M. O., Makarov, V. V., Nedayvozov, V., Grubov, V. V., Pchelintceva, S. V., Hramov, A. E., & Pisarchik, A. N. (2017). Examples of Necker cube images. <https://doi.org/10.1371/journal.pone.0188700.g001>
32. Marr, D., Poggio, T., & Brenner, S. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 204(1156), 301–328. <https://doi.org/10.1098/rspb.1979.0029>
33. Mathys, C., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A Bayesian Foundation for Individual Learning Under Uncertainty. *Frontiers in Human Neuroscience*, 5. <https://doi.org/10.3389/fnhum.2011.00039>
34. Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, 16(1), 72–80. <https://doi.org/10.1016/j.tics.2011.11.018>
35. Moriarity, D. P., & Alloy, L. B. (2021). Back to Basics: The Importance of Measurement Properties in Biological Psychiatry. *Neuroscience & Biobehavioral Reviews*, 123, 72–82. <https://doi.org/10.1016/j.neubiorev.2021.01.008>
36. Nagengast, A. J., Braun, D. A., & Wolpert, D. M. (2009). Optimal Control Predicts Human Performance on Objects with Internal Degrees of Freedom. *PLOS Computational Biology*, 5(6), e1000419. <https://doi.org/10.1371/journal.pcbi.1000419>
37. Noorani, I., & Carpenter, R. H. S. (2016). The LATER model of reaction time and decision. *Neuroscience & Biobehavioral Reviews*, 64, 229–251. <https://doi.org/10.1016/j.neubiorev.2016.02.018>
38. Norman, K. A., Detre, G., & Polyn, S. M. (2008). Computational Models of Episodic Memory. In R. Sun (Szerk.), *The Cambridge Handbook of Computational Psychology* (o. 189–225). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816772.011>
39. Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences*, 105(7), 2745–2750. <https://doi.org/10.1073/pnas.0708424105>
40. Pálffy, Z., Farkas, K., Csukly, G., Kéri, S., & Polner, B. (2021). Cross-modal auditory priors drive the perception of bistable visual stimuli with reliable differences between individuals. *Scientific Reports*, 11(1), 16943. <https://doi.org/10.1038/s41598-021-96198-7>
41. Pellicano, E., & Burr, D. (2012). When the world becomes ‘too real’: A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, 16(10), 504–510. <https://doi.org/10.1016/j.tics.2012.08.009>
42. Petzschner, F. H., Glasauer, S., & Stephan, K. E. (2015). A Bayesian perspective on magnitude estimation. *Trends in Cognitive Sciences*, 19(5), 285–293. <https://doi.org/10.1016/j.tics.2015.03.002>
43. Schmack, K., Castro, A. G.-C. de, Rothkirch, M., Sekutowicz, M., Rössler, H., Haynes, J.-D., Heinz, A., Petrovic, P., & Sterzer, P. (2013). Delusions and the Role of Beliefs in Perceptual Inference. *Journal of Neuroscience*, 33(34), 13701–13712. <https://doi.org/10.1523/JNEUROSCI.1778-13.2013>
44. Schmack, K., Schnack, A., Priller, J., & Sterzer, P. (2015). Perceptual instability in schizophrenia: Probing predictive coding accounts of delusions with ambiguous stimuli. *Schizophrenia Research: Cognition*, 2(2), 72–77. <https://doi.org/10.1016/j.scog.2015.03.005>
45. Series, P. (2020). 2 Methods of Computational Psychiatry: A Brief Survey. In P. Series (Szerk.), *Computational Psychiatry: A Primer*. The MIT Press.
46. Shettel, M., Vecera, S., & Mozer, M. C. (2006). Top-Down Control of Visual Attention: A Rational Account. *Advances in Neural Information Processing Systems*, 18. <https://proceedings.neurips.cc/paper/2005/hash/4f1f29888cabf5d454f866fe457737a23-Abstract.html>
47. Silveira, C., Marques-Teixeira, J., & de Bastos-Leite, A. J. (2012). More Than One Century of Schizophrenia: An Evolving Perspective. *The Journal of Nervous and Mental Disease*, 200(12), 1054–1057. <https://doi.org/10.1097/NMD.0b013e318275d249>
48. Sommer, I. E., de Kort, G. A. P., Meijering, A. L., Dazzan, P., Hulshoff Pol, H. E., Kahn, R. S., & van Haren, N. E. M. (2013). How frequent are radiological abnormalities in patients with psychosis? A review of 1379 MRI scans. *Schizophrenia Bulletin*, 39(4), 815–819. <https://doi.org/10.1093/schbul/sbs037>
49. Stephan, K. E., & Mathys, C. (2014). Computational approaches to psychiatry. *Current Opinion in Neurobiology*, 25, 85–92. <https://doi.org/10.1016/j.conb.2013.12.007>
50. Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., & Corlett, P. R. (2018). The Predictive Coding Account of Psychosis. *Biological Psychiatry*, 84(9), 634–643. <https://doi.org/10.1016/j.biopsych.2018.05.015>
51. Steyvers, M., Lee, M. D., & Wagenmakers, E.-J. (2009). A Bayesian analysis of human decision-making on bandit problems. *Journal of Mathematical Psychology*, 53(3), 168–179. <https://doi.org/10.1016/j.jmp.2008.11.002>
52. Trimble, M. R., & George, M. (2010). *Biological Psychiatry*. John Wiley & Sons.
53. Wood, E. T., Cummings, K. K., Jung, J., Patterson, G., Okada, N., Guo, J., O'Neill, J., Dapretto, M., Bookheimer, S. Y., & Green, S. A. (2021). Sensory over-responsivity is related to GABAergic inhibition in thalamocortical circuits. *Translational Psychiatry*, 11, 39. <https://doi.org/10.1038/s41398-020-01154-0>
54. Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences*, 10(7), 301–308. <https://doi.org/10.1016/j.tics.2006.05.002>

# What Connects Bayes's Rule to Mental Health?

## Introduction to Computational Psychiatry Through the Predictive Coding Account of Perception

A branch of computational psychiatry aims to handle the problems which arose due to the heterogeneity and comorbidity of mental illnesses by capturing phenotypic differences behind symptoms. In order to more accurately describe the mechanisms of the mind and its disturbances, it uses the tools and methods of computational cognitive science. In this paper, we introduce the reader to the (mathematical) language in which (Bayesian) predictive coding was written, which holds one of the possible explanations of the perceptual differences found on the autism or the psychosis spectrum. By representing prior knowledge and sampled sensory information as (prior and likelihood, respectively) distributions, the Bayes-rule allows us to calculate the inferred posterior distribution. The aforementioned terms are introduced through simple examples, such as coin tosses (inferring how biased it is) or making a diagnosis, however, this framework might also help us understand the more complex mechanisms of the mind. Studies show that visual perception can be examined following Bayesian formalisation. The likelihood information can be manipulated with bistable or noisy stimuli, whereas through instructions and/or cues, prior expectations are formed. With the use of perceptual and response models fitted to cognitive task performance, it becomes possible to identify parameters which shed light on the features of information processing that characterise the perceptual alterations in certain clinical populations. At the end of the paper we draw attention to the limitations of computational modelling.

**Keywords:** cognitive sciences, perception, schizophrenia, autism spectrum disorder