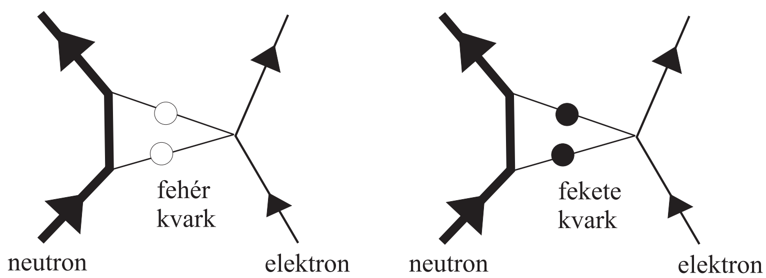
Valószínűség

Hraskó Péter, PTE Elméleti Fizika Tanszék

A valószínűség a hétköznapi életben és a tudományban egyaránt nagy szerepet játszik, azonban vitatott kérdés, hogy hol húzódik a határ a fogalom pontos tudományos értelme és laza diszkurzív használata között. A problémát egy kitalált, de azért elég realisztikus részecskefizikai példán mutatom be.

Tegyük fel, hogy egy elméleti fizikus arra a következtetésre jut, hogy a neutron tartalmaz még három eleddig ismeretlen típusú kvarkot, amelynek két lehetséges változata van. A két változatot fizikusunk "fehér" és "fekete" kvarknak nevezte el, de nem tudta megmondani, hogy a három újfajta kvark között hány fehér és hány fekete van: Az elmélet mind a négy lehetőséget (0, 1, 2 vagy 3 fehér kvark) egyformán megengedte.

Sikerült azonban megmutatnia, hogy a fehér-fekete színmegoszlás kísérletileg vizsgálható. Amikor ugyanis elektronokkal bombázzuk a neutronokat, az új kvarkok egyike nagyon ritkán, véletlenszerűen, virtuális részecske formájában rövid időre kilép a neutronból, és az elektron szóródni tud rajta. A folyamatot az 1. ábra (természetesen fiktív) Feynman-gráfjai szimbolizálják. Az "elmélet" szerint a fehér és a fekete kvark különböző módon szórja az elektronokat (az egyik mondjuk "jobbra", a másik "balra"), ezért ebből a kísérletből meg lehet tudni, hány fehér és hány fekete kvark van a neutronokban.



1.ábra

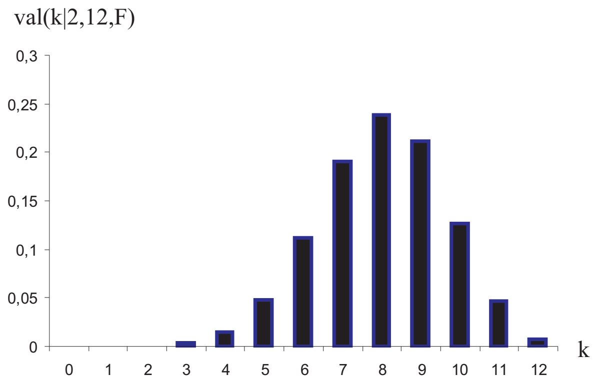
A kísérletnek ezt az aspektusát egy egyszerű urna-modell szemlélteti. Egy urnában van három golyó, amelyek csak a színükben különböznek egymástól:

lehetnek fehérek vagy feketék. Többször egymás után véletlenszerűen kiveszünk egy golyót úgy, hogy miután a golyó színét feljegyeztük, rögtön vissza is tesszük az urnába (és az urnát természetesen minden húzás előtt jól megrázzuk).

Mi a valószínűsége annak, hogy n próbálkozás során k fehér golyót húzunk? A válasz nyilván attól függ, milyen a golyók színmegoszlása. Jelöljük a fehér golyók számát M-mel ($M = 0, 1, 2$ vagy 3). A fekete golyók száma ekkor $(3 - M)$. Annak valószínűsége, hogy egy fehér golyót húzzunk ki, $M/3$-mal, annak valószínűsége pedig, hogy feketét, $(1 - M/3)$-mal egyenlő. A kérdésünkre — milyen valószínűséggel lesz n kihúzott golyó között k fehér, — a *binomiális eloszlás* adja meg a választ:

$$\text{val}(k|M, n, F) = \binom{n}{k} \left(\frac{M}{3}\right)^k \left(1 - \frac{M}{3}\right)^{n-k}. \tag{1}$$

Ebben a képletben és a továbbiakban is minden valószínűséget jelentő függvényt val(.|...) alakban írunk. A val() függvény argumentumát a | jel két részre osztja. A vonaltól balra találjuk azt a változót, amelynek a valószínűségéről szó van. A függőleges vonaltól jobbra található szimbólumok azokra a *feltételekre* utalnak, amelyek mellett a valószínűségi eloszlás érvényes. Az (1) képlet pl. akkor igaz, amikor az urnában M fehér golyó van és összesen n-szer húzunk. A biztonság kedvéért szerepel még egy F jel is, amely figyelmeztet rá, hogy mindig vannak specializálatlan feltételek, amelyek közül később egyet vagy többet esetleg explicite is meg kell majd adnunk. Mivel *minden* valószínűségre egyöntetűen a val() jelet használjuk, az argumentumból kell kiderülnie, minek a valószínűségéről van szó.



2.ábra

Az (1) mindhárom tényezőjének világos matematikai jelentése van. Az $(M/3)^k$ annak következménye, hogy — feltételezésünk szerint — k-szor húzunk fehér golyót, az $(1 - M/3)^{n-k}$-ra pedig azért van szükség, mert a maradék $(n - k)$ alkalommal fekete golyót húzunk ki. Az $\binom{n}{k}$ binomiális együttható veszi figyelembe, hány különböző sorrendben következhetnek egymás után a fehér és a fekete golyók. Ez az eloszlás helyesen van normálva, mert a binomiális tétel alapján

$$\sum_{k=0}^{n} \mathrm{val}(k|M, n, F) = \left[\frac{M}{3} + \left(1 - \frac{M}{3}\right)\right]^n = 1.$$

A 2. ábrán példaként a $\mathrm{val}(k|2, 12, F)$ eloszlást ábrázoltuk.

Az (1) arra jó, hogy ha M-et ismerjük, kiszámíthassuk, milyen valószínűséggel húzunk n próbálkozásból k-szor fehéret. Azonban nem ez az a valószínűség, amelyre szükségünk van. *Az M-et szeretnénk megtudni*, azt, hogy hány fehér kvark van a neutronban (hány fehér golyót tartalmaz az urna). Már említettük, hogy az 1. ábrán felrajzolt folyamatok, amelyek erről informálnak, nagyon ritkák, ezért a kísérlet nagyon drága (a szükséges gyorsító idő hosszú). Ezt figyelembe véve a költségvetésünket — elég rugalmasan — úgy állapították meg, hogy a kísérletet addig folytathatjuk, ameddig — mondjuk — 5 bennünket érdeklő folyamatot nem találunk, ami az urnamodellben $n = 5$ húzásnak felel meg.

Tegyük fel, hogy a kísérlet megtörtént és az 5 folyamatból 2 tartozott fehér, 3 pedig fekete kvarkhoz. Az urnamodell nyelvén ebben a helyzetben a következő kérdésre kell választ találnunk: Ha $n = 5$ húzásból $k = 2$ fehéret kaptunk, akkor hány fehér golyó van az urnában (mekkora az M)?

Biztosat nyilván nem mondhatunk, de esetleg kaphatunk M-re egy valószínűségi eloszlást, $\mathrm{val}(M|k, n, F)$-t, amely $\mathrm{val}(k|M, n, F)$-től csak abban különbözik, hogy a k és az M helyet cserélt egymással. Ez azonban egyáltalán nem lényegtelen különbség, mert a két jel a függőleges választóvonal különböző oldalán áll. A különbség illusztrálására forduljunk a kockadobáshoz. Annak $\mathrm{val}(6|ps)$ valószínűsége, hogy 6-ost dobjunk *feltéve*, hogy a dobásunk páros, 1/3-dal egyenlő, mert a kockán 3 páros érték található, és ezek egyenlően valószínűek. A $\mathrm{val}(ps|6)$ valószínűség viszont nyilván 1, hiszen ha 6-ost dobunk, azzal automatikusan a dobásunk páros is.

A valószínűségszámítás ismer egy képletet, a *Bayes-formulát*, amely kapcsolatot teremt a $\mathrm{val}(X|Y, F)$ és a $\mathrm{val}(Y|X, F)$ valószínűségek között. A képlet levezetéséhez azt kell felhasználnunk, hogy a $\mathrm{val}(X|Y,F)$ feltételes valószínűséget a

$$\mathrm{val}(X|Y,F) = \frac{\mathrm{val}(X \text{ és } Y|F)}{\mathrm{val}(Y|F)} \qquad (\mathrm{val}(Y|F) \neq 0) \tag{2}$$

képlet segítségével számíthatjuk ki, amelyben $\mathrm{val}(X \text{ és } Y|F)$ az X és az Y együttes bekövetkezésének a valószínűsége.

A (2) igaz marad, ha az X-et és az Y-t felcseréljük egymással:

$$\mathrm{val}(Y|X,F) = \frac{\mathrm{val}(Y \text{ és } X|F)}{\mathrm{val}(X|F)} \qquad (\mathrm{val}(X|F) \neq 0).$$

Az (X és Y) kijelentés (vagy esemény) természetesen azonos az (Y és X) kijelentéssel (eseménnyel). A két képletből ennek a kijelentésnek a valószínűségét kiküszöbölve jutunk el a

$$\mathrm{val}(Y|X,F) = \frac{\mathrm{val}(X|Y,F)\,\mathrm{val}((Y|F)}{\mathrm{val}(X|F)} \qquad (\mathrm{val}(X|F) \neq 0) \tag{3}$$

Bayes-formulához (*Bayes-tételhez*).

Mint látjuk, ez a képlet valóban kapcsolatot létesít a $\mathrm{val}(Y|X,F)$ és a $\mathrm{val}(X|Y,F)$ valószínűség között. A kockadobásos feladatban például $Y \to 6$ és $X \to ps$ azonosítással ez a reláció természetesen teljesül, mert — mint láttuk, — $\mathrm{val}(6|ps,F) = 1/3$, $\mathrm{val}(ps|6,F) = 1$, és $\mathrm{val}(6|F) = 1/6$, $\mathrm{val}(ps|F) = 1/2$.

Térjünk most vissza a részecskefizikai példához (és a neki megfelelő urnamodellhez), és a Bayes-tétel segítségével fejezzük ki $\mathrm{val}(M|k,n,F)$-t az (1) binomiális eloszláson keresztül:

$$\mathrm{val}(M|k,n,F) = \frac{\mathrm{val}(k|M,n,F)\,\mathrm{val}(M|n,F)}{\mathrm{val}(k|n,F)}. \tag{4}$$

Foglalkozzunk a jobboldalon szereplő valószínűségekkel. A nevezőbeli valószínűség nem tartalmaz M-et, ezért a

$$\sum_{M=0}^{3} \mathrm{val}(M|k,n,F) = 1 \tag{5}$$

normálási feltétel segítségével kifejezhető a számlálóban álló valószínűségeken keresztül:

$$\mathrm{val}(k|n,F) = \sum_{M=0}^{3} \mathrm{val}(k|M,n,F)\,\mathrm{val}(M|n,F). \tag{6}$$

A számláló első tényezője formailag azonos az (1) függvénnyel, de — az (1)-től eltérően, — itt nem k függvényeként kell érteni (fix M mellett), hanem M függvényeként rögzített k-nál. A két függvény tehát ugyanabban az értelemben különbözik egymástól, mint az x^n, amikor a független változó x (hatványfüggvény), és amikor a független változó n (exponenciális függvény). A $\mathrm{val}(k|M,n,F)$-nek is célszerű a két esetben különböző elnevezést adni. A (4)-ben, amikor a független változó M (rögzített k mellett), *likelihood-függvény* a neve (magyarul is!), míg az (1)-ben a k valószínűségi eloszlásának hívjuk.

A (4) számlálójának első tényezője tehát a likelihood-függvény, amit ismerünk. A második tényező mibenlétét kell még tisztáznunk. Ez a tényező — a baloldalhoz hasonlóan — valószínűségi eloszlás M-ben, ezért eleget kell tennie a

$$\sum_{M=0}^{3} \mathrm{val}(M|n,F) = 1 \tag{7}$$

normálási feltételnek. A (4)-ben a baloldali és a számlálóbeli M-eloszlás között az a különbség, hogy az előbbiben már figyelembe vettük a megfigyelés (szóráskísérlet, urnakísérlet) eredményét (*poszterior valószínűség*), míg az utóbbi csak azokat az ismereteket tartalmazza, amelyekkel a kísérlet elvégzése előtt is rendelkeztünk (*prior valószínűség*). A Bayes-formula tehát a kísérletezés lényegét fejezi ki tömör matematikai formában: Egy kísérlet értelme ugyanis az, hogy a meglévő ismereteinket korrigálja.

Mint már mondottuk, az elképzelt kísérletünkben 5 próbálkozásból 2 golyó bizonyult fehérnek. A likelihood-függvényünk ekkor

$$\mathrm{val}(2|M,5,F) = \frac{1}{3^5}\binom{5}{2} M^2(3-M)^3.$$

Mi lesz a prior? A kísérlet előtt fogalmunk sincs róla, milyen a színeloszlás. Ezt a teljes tudatlanságot valószínűleg akkor fejezzük ki megfelelő módon, ha kezdetben mind a négy lehetséges M-et egyenlően valószínűnek tekintjük:

$$\mathrm{val}(M|n,F) = \frac{1}{4}$$

(ez a valószínűség semmiképpen sem függhet attól, hogy a később elvégzendő kísérletben mekkora lesz az n).

Mindezeket (4)-be helyettesítve a keresett valószínűségre a

$$\mathrm{val}(M|2,5,F) = K \cdot M^2(3-M)^3 \tag{8}$$

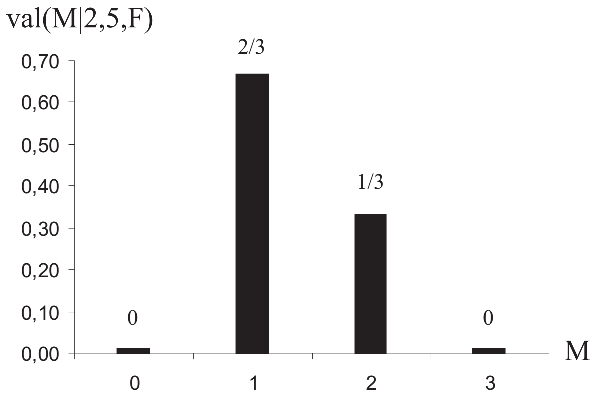
képletet kapjuk. A K-ra az (5) normálási feltételből a

$$K = \frac{1}{1^2 \cdot 2^3 + 2^2 \cdot 1^3} = \frac{1}{12}$$

érték adódik. Így végül

$$\mathrm{val}(M|2,5,F) = \frac{1}{12} M^2 (3-M)^3. \tag{9}$$

Ezt a valószínűségi eloszlást a 3. ábrán láthatjuk. Mivel fehér golyót is, feketét is húztunk, ezért bizonyos, hogy M nem lehet egyenlő se nullával, se hárommal, és a kísérlet szerint az urna kétszer olyan valószínűséggel tartalmaz egy fehér golyót, mint kettőt. Ez ugyan távol van a bizonyosságtól, de részecskefizikusunk számára jelenthet értékes információt.



3.ábra

***

Ez az a pont, ahol visszatérhetünk a legelső bekezdésben felvetett problémához. A helyzet ugyanis az, hogy a valószínűségszámítás művelőinek óriási többsége szerint a (9) képlet *tökéletesen értelmetlen*, mert annak, hogy a neutronban hány fehér kvark van, *nincs valószínűsége*, hiszen ez a szám nem a véletlenen múlik, hanem határozottan nulla, vagy egy, vagy kettő, vagy három. Ugyanez a helyzet természetesen az urnával is. Bizonytalanságról és valószínűségről itt csak a színeloszlásra vonatkozó *tudásunk* kapcsán lehet szó, ami azonban esetleges, szubjektív dolog (*szubjektív valószínűség*). Az effajta szubjektív elvárást szigorúan véve nem is lenne jogos valószínűségnek hívni, mert egyáltalán nem lehetünk biztosak benne, hogy alkalmazhatók rá a valószínűségszámítás tételei, konkréten például a Bayes-formula.

A (9)-cel ellentétben — folytatódik az érvelés, — az (1) binomiális eloszlás a véletlenen múló objektív valószínűséget fejez ki, hiszen ugyanazt a kísérletet sokszor megismételve ellenőrizhető a helyessége. A legegyszerűbb esetben például, amikor csak egyszer próbálkozunk ($n = 1$), a képlet szerint $\mathrm{val}(1|M, 1, F) = M/3$ valószínűséggel húzunk fehér golyót. Az M itt természetesen tökéletesen határozott szám, ha az urnában mondjuk csak 1 fehér golyó van, akkor 1/3. A kísérletet akárhányszor megismételhetjük és empirikusan meghatározhatjuk a fehér golyók előfordulásának *relatív gyakoriságát*. Azt fogjuk találni, hogy a húzások számának a növelésekor a relatív gyakoriság 1/3-hoz tart, ahogy ezt a valószínűség intuitív jelentése alapján várjuk is, és ez teljesen független attól, mit tudunk vagy gondolunk a színmegoszlásról az urnában.

A relatív gyakoriságként értelmezett valószínűségekről (*objektív valószínűség*) nagyon könnyű belátni, hogy eleget tesznek a valószínűségszámítás axiómáinak, és ennek következtében természetesen a Bayes-formulának is. Annak $\mathrm{val}(ps|F)$ valószínűsége például, hogy egy kockával párost dobunk, vagy az a $\mathrm{val}(6|F)$, hogy 6-ost dobunk vele, nyilvánvalóan objektív valószínűségek. Ezért nem meglepő, hogy a Bayes-formula érvényes rájuk. A $\mathrm{val}(k|M, n, F)$-fel és a $\mathrm{val}(M|k, n, F)$-fel azonban nem ez a helyzet, mert közülük csak az első objektív.

Lehet-e bármit is felhozni ezzel az objektív ("frekventista") nézőponttal szemben a szubjektív ("bayesi") felfogás védelmében? Meg lehet például jegyezni, hogy a relatív frekvenciát csak végtelen hosszú kísérletsorozatban lehet pontosan azonosítani a valószínűséggel. Ilyen sorozatok azonban csak a képzeletünkben léteznek, ezért a frekventista felfogás is tartalmaz szubjektív elemet. Elképzelhető azonban ennél konstruktívabb ellenvetés is: Az, ha sikerülne megmutatni, hogy *a szubjektív valószínűség is rendelkezik azokkal a tulajdonságokkal, amelyek elégségesek a Bayes-tétel igazolásához.*

Több olyan gondolatmenet is ismeretes, amelyek elég meggyőzően demonstrálják, hogy valóban ez a helyzet. Az alábbiakban — befejezésül, — röviden kitérek az egyikre (azért csak röviden és vázlatosan, mert a gondolatmenet nehéz, és részletes ismertetése messze meghaladná ennek az írásnak a kereteit). Előbb azonban tisztázni szeretnék egy lehetséges félreértést. Az, hogy valami igazi valószínűség-e vagy csupán "amorf elvárás", nem azon múlik, hogy érvényes-e vagy sem. Ha pl. egy kockáról feltételezem, hogy egyenlő valószínűséggel esik mind a hat oldalára, utóbb pedig kiderül, hogy ez nem igaz, mert a kocka cinkelt, az eredeti hibás várakozásom ettől még valószínűségi természetű marad. Csak az érvényessége az, ami elvész. Ugyanígy, a $\mathrm{val}(M|n,F) = 1/4$ szubjektív prior az urnapéldában biztosan nem érvényes (nem igaz), de nem ezen múlik, hogy tekinthető-e igazi valószínűségnek, vagy sem.

1946-ban *Richard Cox* igényes bizonyítást adott arra, hogy a Bayes-tétel a szubjektív valószínűségekre is alkalmazható[1]. A bizonyítás a következő három feltevésen alapult:

1) A szubjektív elvárások[2] a mértékük szerint sorba rendezhetők: Ha X teljesülésére inkább számítunk, mint Y-éra, Y teljesülésében pedig jobban bízunk, mint Z megvalósulásában, akkor X teljesülésére inkább számítunk, mint Z-ére. Ezt a feltevést matematikusabb formában is megfogalmazhatjuk, ha az X, az Y, a Z stb. kijelentésekhez úgy rendelünk (egyébként önkényesen) $\mathrm{elv}(X|F)$, $\mathrm{elv}(Y|F)$, $\mathrm{elv}(Z|F)$ valós számokat, hogy teljesüljenek az

$$\mathrm{elv}(X|F) > \mathrm{elv}(Y|F) > \mathrm{elv}(Z|F)$$

egyenlőtlenségek. Az $\mathrm{elv}(.|F)$ függvény természetesen az elvárás mértékére utal az F nem specifikált feltételek teljesülése mellett. Az F-re való utalás biztosítja, hogy a jelentős mértékben eltérő kontextushoz tartozó elvárásoknak ne kelljen egymással összehasonlíthatóknak lenniük. Nehéz elképzelni olyan körülményeket, amelyek között össze kellene hasonlítanunk mondjuk annak az elvárását, hogy holnap földrengés lesz Lisszabonban azzal, hogy a Voyager 2 űrszonda rádióadója holnap végleg felmondja a szolgálatot.

2) Ha X bekövetkeztére van elvárásunk, akkor X elmaradására is automatikusan rendelkezünk elvárással.

3) Ha van elvárásunk Y teljesüléséről, valamint arról, hogy X teljesül feltéve, hogy Y teljesül, akkor arról is van elvárásunk, hogy X és Y együttesen teljesül.

[1] *Am. J. Phys*, **14**, 1-13 (1946)

[2] Az angol *belief* szó pontosabban fejezi ki, hogy mire kell gondolni. A *belief* magyar jelentései azonban (hit, meggyőződés, bizalom) olyan érzelmi töltettel rendelkeznek, amelyek alkalmatlanná teszik őket egy tisztán valószínűségszámítási fogalom megnevezésére. Ezért választottam — jobb híján — az *elvárás* nevet.

A gondolatmenetében Cox még azt is megkövetelte, hogy ha egy adott információt több különböző módon tudunk felhasználni, mindig ugyanazokhoz a következtetésekhez jussunk el, akármelyik lehetséges eljárást válasszuk is az analízishez. Azt bizonyította be, hogy ha a három feltétel teljesül, akkor az $\mathrm{elv}(X|F)$ elvárások eleget tesznek a valószínűségszámítás axiómáinak. Így speciálisan

$$\mathrm{elv}(X|F) + \mathrm{elv}(\overline{X}|F) = 1 \qquad (\overline{X} \text{ a "nem } X\text{" jele}),$$

$$\mathrm{elv}(X \text{ és } Y|F) = \mathrm{elv}(X|Y,F) \cdot \mathrm{elv}(Y|F)$$

összefüggéseknek, és ha I bizonyosan igaz, H pedig bizonyosan hamis, akkor

$$\mathrm{elv}(I|F) = 1, \qquad \mathrm{elv}(H|F) = 0.$$

Ezek az összefüggések a *valószínűségek* legalapvetőbb tulajdonságait fejezik ki, közöttük azt is, amelyik a Bayes-tétel igazolásához szükséges, ezért az elv() függvényeket helyettesíthetjük bennük val() függvényekkel. Cox bizonyítása alapján ezért konzisztens elvárásainkat valóban tekinthetjük valószínűségeknek.

***

A "bayesi módszerrel" egyre gyakrabban lehet találkozni az olyan publikációkban, amelyek mérések tervezésével és az eredmények kiértékelésével foglalkoznak. Ez a módszer ugyanis jelentős mértékben kitágítja az analízis lehetőségeit, mert a problémák sokkal szélesebb körében teszi lehetővé a valószínűségszámítás részletesen kidolgozott, hatékony fogalmi és matematikai apparátusának alkalmazását. Ez az ismertetés azért született meg, mert *szubjektíve* nagyon *valószínűnek* tartom, hogy a szubjektív valószínűségek Bayes-tételen alapuló bevezetését egyre szélesebb körben fogják megengedett eljárásnak tekinteni és alkalmazni.

Függelék

A $\mathrm{val}(M|k,n,F)$ poszterior valószínűség — mint láttuk — szubjektív, de matematikailag teljesen határozott kifejezés, amelynek a tulajdonságait meg lehet vizsgálni. A függelékben két olyan fontos matematikai tulajdonságára mutatok rá, amelyek megerősítenek abban, hogy nincs okunk idegenkedni a szubjektív valószínűségektől.

A $\mathrm{val}(M|k,n,F)$ első argumentuma, az M változó, a 0, 1, 2, 3 értékeket veheti fel, és a függvény értéke általában egyik M-nél sem nulla. Ezek közül csak az egyik egyezik meg a fehér golyók valóságos számával az urnában. Jelöljük ezt a számot $\mathcal{M}$-mel. A $\mathrm{val}(M|k,n,F)$ függvény első tulajdonsága, amelyet igazolni is fogunk az, hogy amikor a próbálkozások n száma egyre nagyobb és nagyobb, a $\mathrm{val}(M|k,n,F)$ értéke $M=\mathcal{M}$-nél 1-hez, a többi M értéknél pedig nullához tart. Tömören ezt így fejezhetjük ki:

$$\lim_{n\to\infty} \mathrm{val}(M|k,n,F) = \begin{cases} 1 & \text{ha } M=\mathcal{M}, \\ 0 & \text{ha } M\neq\mathcal{M}. \end{cases} \tag{10}$$

Ennek a tulajdonságnak az ismeretében válaszolni tudunk a következő kérdésre: n előzetes húzás eredményének (a k-nak) az ismeretében mi a $\mathrm{val}(1|k,n,I)$ valószínűsége[3] annak, hogy a következő, $(n+1)$-edik, próbálkozásnál fehér golyót húzunk. A választ a

$$\mathrm{val}(1|k,n,I) = \sum_{M=0}^{3} \frac{M}{3}\cdot \mathrm{val}(M|k,n,F) \tag{11}$$

összeg adja meg. A fehér golyó valószínűsége ugyanis $M/3$-mal egyenlő, ezt kell súlyozni a lehetséges M-ek valószínűségével. A (10) figyelembevételével nagy n-nél ez az összeg az $\mathcal{M}/3$ értékhez tart. Egy hosszú sorozatban tehát a fehér golyók gyakoriságára a $\mathrm{val}(M|k,n,F)$ szubjektív poszterior valószínűség ugyanazt az objektív értéket szolgáltatja, mint a frekventista felfogás — teljesen elmosódik a különbség a két megközelítés között[4].

[3] A fehér golyóhoz az 1, a feketéhez a 0 számot rendeljük.

[4] Ilyen természetű problémával *Laplace* foglalkozott először. Azt a kérdést vizsgálta, hogy ha egy ismeretlen tulajdonságú pénzérmét n-szer feldobunk és a sorozatban k fejet találunk, akkor mi lesz a fej valószínűsége az $(n+1)$-edik dobásban. A pénzfeldobást is az (1) binomiális eloszlás írja le, ha benne $M/3$-t az érmét jellemző p paraméterrel helyettesítjük, amely a $(0,1)$ intervallumban bármilyen értéket felvehet. Ha ezt a p-t ismerjük, a kérdés könnyen megválaszolható: A fej valószínűsége minden egyes dobásban p-vel egyenlő. A p-t azonban általában nem ismerjük, erre a nagyon is valóságos esetre vonatkozott Laplace kérdése. Azt találta, hogy a keresett valószínűség $(k+1)/(n+2)$-vel egyenlő. Ez a képlet tetszőleges n-nél ($n=0$-nál is) érvényes. Ebből vonható le az a következtetés, hogy az ismeretlen p egy hosszú sorozat relatív frekvenciájával egyenlő.

Laplace ismerte a Bayes-tételt és a jelentőségével is tisztában volt, de ebben a bizonyításban nem használta fel. A Bayes-tétel kihasználásával a bizonyítás nagyon leegyszerűsödik és kevesebb feltevést igényel, mint amennyire Laplace-nak szüksége volt.

A (11) képlet azonban tetszőleges számú próbálkozásra érvényes, sőt, ha a val$(M|k,n,F)$-t a prior valószínűséggel helyettesítjük benne, azt a valószínűséget adja meg, amellyel a *legelső* próbálkozásnál húzunk fehér golyót. Mivel a prior valószínűségi eloszlás 1/4-del egyenlő, erre a valószínűségre — ahogy várható, — az 1/2 értéket kapjuk. Ha pedig az első húzás fehér golyót eredményezett, akkor — mivel ezen az egy tényen kívül semmiféle más információnk sincs a golyók színeloszlásáról, — a második húzásban 1/2-nél *nagyobb* valószínűséggel várhatjuk (szubjektív valószínűség!), hogy újra fehéret húzzunk. A (11) képlet erre a megnövekedett valószínűségre a 7/9 értéket szolgáltatja, mert benne val$(M|1,1,F) = M/6$-tal egyenlő. Persze ha ismernénk $\mathcal{M}$-et (a fehér kvarkok számát a neutronban), akkor mindkét valószínűség ugyanazzal az $\mathcal{M}/3$-mal lenne egyenlő. De most még csak a kezdeti lépéseket tesszük $\mathcal{M}$ megismerése felé: Ezért változik egyik lépésről a másikra a fehér golyó találati valószínűsége.

A (10) bizonyítására rátérve mindenekelőtt azt kell tisztázni, hogyan viselkedik k, amikor n a végtelenhez tart. Az $\mathcal{M} = 0$ esetben a k végig zérus, mert ekkor az urnában egyáltalán nincs fehér golyó. Amikor pedig $\mathcal{M} = 3$, az urna csak fehér golyót tartalmaz, ezért k folyamatosan n-nel egyenlő. A két közbülső esetben k viselkedését a Nagy Számok Törvénye határozza meg, amely szerint hosszú húzássorozatban a k/n arány ahhoz a valószínűséghez tart, amellyel egyszeri próbálkozásnál fehér golyót húzunk: $\mathcal{M} = 1$-nél ez a valószínűség 1/3-dal, $\mathcal{M} = 2$-nél 2/3-dal egyenlő.

A (8)-ra vezető gondolatmenet alapján a poszterior valószínűségre tetszőleges k, n párnál a

$$\mathrm{val}(M|k,n,F) = K_{kn} \cdot M^k (3-M)^{n-k} \tag{12}$$

képletet kapjuk, amelyben a normálási tényező

$$K_{kn} = \frac{1}{\sum_{m=0}^{3} m^k (3-m)^{n-k}} \tag{13}$$

-val egyenlő (a szumma alatti M-t azért változtattuk m-re, hogy nehezebb legyen összetéveszteni az M változóval).

Tárgyaljuk először a $k = 0$ esetet, amikor az n próbálkozás egyikében sem húzunk ki fehér golyót. Ha $\mathcal{M} = 0$, akkor ez biztosan így van, de $\mathcal{M} = 1$-nél és $\mathcal{M} = 2$-nél is előfordulhat. A Nagy Számok Törvénye alapján azonban ez utóbbi két esetben egy $k = 0$-s sorozat valószínűsége $n \to \infty$-nél nulla. Amikor $k = 0$, a (12)-ben $M = 0$-nál a 0^0 kifejezés jelenik meg, amely határozatlan: $\lim_{x\to 0} x^0 = 1$, míg $\lim_{x\to 0} 0^x = 0$. A (12)-ben ezt a kifejezést 1-nek kell tekinteni, mert a

$$\mathrm{val}(k|0,n,F) = \binom{n}{k} \cdot 0^k \cdot (1-0)^{n-k}$$

likelihood-függvényből származik, amely — mint a k valószínűségi eloszlása — 1-re van normálva:

$$\sum_{k=0}^{n} \binom{n}{k} \cdot 0^k \cdot (1-0)^{n-k} = [0 + (1-0)]^n = 1.$$

Az összeg $k > 0$ tagjai azonban mind nullák, ezért a $k = 0$-s tag az, ami 1-gyel egyenlő. Ebből látható, hogy az ebben a tagban szereplő 0^0 kifejezés 1-nek tekintendő.

Ennek ismeretében a (12)-ből nyilvánvaló, hogy

$$\mathrm{val}(M|0,n,F) = \frac{(3-M)^n}{1^n + 2^n + 3^n}.$$

Amikor $n \to \infty$, a nevező első két tagja elhanyagolható a 3^n mellett, ezért ebben a határesetben az $M = 0$-hoz tartozó valószínűség 1-hez, a többi nullához tart. A most vizsgált esetben ($\mathcal{M} = 0$) a (10) szerint valóban ennek kell történnie. Teljesen hasonlóan igazolható (10) helyessége akkor is, amikor $k = n$.

A közbenső tartományban, amikor $0 < k < n$, (12) szerint

$$\mathrm{val}(M|k,n,F) = \begin{cases} \dfrac{2^{n-k}}{2^{n-k} + 2^k} & \text{ha } M = 1, \\ \dfrac{2^k}{2^{n-k} + 2^k} & \text{ha } M = 2, \\ 0 & \text{ha } M = 0 \text{ vagy } M = 3. \end{cases} \tag{14}$$

Most, amikor k nem egyenlő se 0-val, se n-nel, az n kihúzott golyó között fehérek és feketék is vannak, ezért $\mathcal{M}$ nem lehet se 0, se 3. A (14) harmadik sora mutatja, hogy ebben a tekintetben összhangban vagyunk (10)-zel. Véges n-nél a (14)-beli két tört egyike se nulla, de az $n \to \infty$ határesetben, a (10)-zel ugyancsak összhangban, az első tört 1-hez, a második nullához tart, $\mathcal{M} = 2$-nél pedig pont megfordítva. Ez annak a következménye, hogy a Nagy Számok Törvénye szerint $\mathcal{M} = 1$-nél k helyébe $n/3$-t, $\mathcal{M} = 2$-nél pedig $2n/3$-t kell írnunk. A (10) érvényességét ezzel minden lehetséges esetre igazoltuk.

A $\mathrm{val}(M|k,n,F)$ másik kedvező matematikai tulajdonsága, amelyre a függelék elején utaltunk az, hogy a (10) képlet érvényes marad akkor is, amikor $\mathrm{val}(M|n,F) = 1/4$ helyett másik prior valószínűséget választunk. Ez a prior lényegében tetszőleges lehet. Az egyetlen kikötés az, hogy egyik M-nél se legyen pontosan nulla. Ez a tulajdonság, amelyet numerikus szimuláció segítségével lehet demonstrálni, azért nagyon fontos, mert a szubjektív valószínűség kritikusainak gyakori érve, hogy a prior valószínűség megadásának a szükségessége elfogadhatatlan önkényességet visz bele a számításokba. Amikor azonban n elég nagy, a függés a prior valószínűségtől — mint látjuk, — gyenge. Abban a nagyon gyakori esetben pedig, amikor az n értéke korlátozott, a poszterior valószínűséget *javíthatjuk* azzal, hogy a prior megválasztásánál maximálisan kihasználjuk az előzetes ismereteinket és a megalapozott elvárásainkat. Így például egy kísérlet megismétlésénél a korábbi kísérlet poszterior valószínűségét választhatjuk priornak.