

WÁGER JÁNOS

Néhány gondolat a Bayes-analízisről

A jogász természetéből ered, hogy idegenkedik a matematikától, jogászörökben a többjegyű természetes számokkal végezendő alapműveletek gyakran olyan speciális szakértelmet igénylő kérdésekké alakulnak, amelyek szakértő kirendelését – jobb esetben csak számológép formájában – teszik indokolttá. Miben segíthet akkor jogászoknak egy tizenharmadik századi matematikus? Eleve ki volt, és mit csinált ez a Bayes?

A büntetőeljárás kulcskérdése: bűnös vagy ártatlan a vádlott. Bizonyára sok bíró boldogan venne olyan módszertant, amellyel ez biztosan megállapítható. Talán ilyen a Bayes-analízis? Nos, sajnos nem. A bűnösség olyan fogalom, amelynek a matematikában semmi keresnivalója, az ugyanis *Földvári József* megfogalmazásában az elkövető és a cselekmény közötti pszichikai viszony, amelynek alapján a terhére rójuk a cselekményt.

Annak a kérdésnek a megválaszolásában azonban, hogy a vádlott követte-e el a terhére rótt cselekményt, vagy sem, a matematika igenis segítségünkre lehet. Természetesen azon felül, hogy a vádlott követte el a cselekményt, számos egyéb feltétele van annak, hogy az elkövetőt bűnösnek nyilvánítsa, és vele szemben büntetést szabjon ki a bíróság. Ezekkel a feltételekkel munkámban nem foglalkozom, arra helyezem a hangsúlyt, hogy milyen módon lehet matematikai eszközökkel alátámasztani vagy cáfolni azt, hogy a vádlott követte el a terhére rótt bűncselekményt. Azaz: hogyan lehet az elkövetés esélyét/valószínűségét meghatározni?

Ennyi bevezető után térjünk rá arra, hogy ki is volt *Thomas Bayes*. Bayes egy tizenharmadik századi angol matematikus és lelkész, aki ismertségét annak köszönheti, hogy halála után az egyik barátja, *Richard Price* megtalálta az iratai között az *Essay Towards Solving a Problem in the Doctrine of Chances* című munkáját, amelyet a Royal Society of London kiadott. Az esszé a Bayes-tételként ismert képlet egy speciális esetét tartalmazza.¹ Bayes az esszéjét nem büntetőjogi esetek megoldására dolgozta ki, hanem a valószínűség-számítás egy általános változatát határozta meg, amely eltér a halmazelméleten alapuló valószínűség-számítástól.

¹ http://de.wikipedia.org/wiki/Thomas_Bayes

A Bayes után bayesi valószínűség-fogalomként (angolul: *Bayesianism*) elnevezett meghatározás a valószínűséget a meggyőződés fokaként (*degree of belief*) értelmezi. Ez az értelmezés lehetővé teszi, hogy az állítás valószínűségét új ismeretek fényében újra meghatározzuk, így lehetőség van arra, hogy állást foglaljunk abban a kérdésben, hogy az új ismeret növelte vagy csökkentette-e azt. *Laplace* Bayestől függetlenül felismerte ezt az összefüggést, és a csillagászatban, az egészségügyben, sőt az igazságszolgáltatásban is alkalmazta.² A bayesi valószínűség-fogalmat elsősorban Angliában alkalmazzák, vezető képviselői ennek az irányzatnak *Harold Jeffreys* és *Frank Plumpton Ramsey*. A Bayes módszertanán alapuló eljárások különösen a játékelméletben, illetve a sportfogadások (oddsok meghatározása) esetén elterjedtek.³

A kriminalisztikában a Bayes-képlet a következő formában terjedt el:

$$E_{a\text{-posteriori}} = P_b/p_a * E_{a\text{-priori}}$$

$E_{a\text{-posteriori}}$: A terhelt bűnösségének esélye a vizsgált bizonyíték alapján (a posteriori esély)

p_b : A vizsgált bizonyíték keletkezésének a valószínűsége abban az esetben, ha a terhelt követte el a cselekményt

p_a : A vizsgált bizonyíték keletkezésének a valószínűsége abban az esetben, ha a terhelt ártatlan

$E_{a\text{-priori}}$: A terhelt bűnösségének esélye a bizonyíték vizsgálat előtt (a posteriori esély)⁴

A Bayes-képlet megmutatja, hogy hányszor esélyesebb a terhelt arra, hogy ő az elkövető, mint hogy nem ő az. Úgy is fogalmazhatunk, hogy az ártatlan elítélés szorzóját adja meg⁵. Amennyiben eredményként például ezret kapunk, azt jelenti, hogy ilyen bizonyítékok mellett ezer az egyhez annak az esélye, hogy tévesen kerül sor elítélésre.

A Bayes-analízis alkalmazásakor a kulcskérdés, ahogy arra *Tatár László* helyesen rámutat, az a priori valószínűségek meghatározása, ennek az értéknek a megválasztása ugyanis döntően kihat a későbbi számításokra.⁶ Az a

² http://de.wikipedia.org/wiki/Bayesscher_Wahrscheinlichkeitsbegriff

³ Uo.

⁴ Tremmel Flórián – Fenyvesi Csaba – Herke Csongor: *Kriminalisztika. Tankönyv és atlasz. Dialóg Campus Kiadó, Budapest–Pécs, 2005*

⁵ Sportnyelven fogalmazva: annak az oddsát, hogy az esélyes csapat (az ártatlan személyről) elveszti a meccset (a bíróság megállapítja, hogy ő követte el a terhére rótt cselekményt).

⁶ http://www.drtatar.hu/bayes_de.htm

priori esély és az a priori valószínűség eltérő fogalmak azonban ezek egymásból származtathatók.⁷ A legrégebbi módszert az a priori érték meghatározására Laplace alkotta meg. E szerint amennyiben nincsen annak konkrét oka, hogy mást feltételezzünk, akkor valamennyi elementáris kimenetnek a valószínűségét azonosnak kell tekintetni.⁸ Lefordítva ezt a szabályt a büntetőjog nyelvére: valamennyi számításba jövő személy esetén azonosnak kell tekinteni az elkövetés esélyét.

A kiindulási esélyt (mielőtt bármilyen bizonyíték értékelésére sor került volna) a gyakorlatban a számításba jöhető személyek száma alapján szokás meghatározni, azaz a bűnösök és az ártatlanok arányát vesszük. Tehát egy olyan hányadost tekintünk a priori esélynek, amelynek a számlálójában 1, a nevezőjében a szóba jöhető személyek –1, azaz a szóba jöhető személyek közül az ártatlanok szerepelnek. Tegyük fel, hogy biztosak vagyunk abban, hogy a cselekményt egy személy követte el, és a lehetséges elkövetők száma három. Ilyen probléma állhat elő egy kocsmai szurkálásnál, hárman ütlegetik a sértettet, majd egyszer csak előkerül egy kés, amely a sértett szívében köt ki. A kérdés az, ki szúrta. Kétséget kizáróan valamelyik a három közül. Ebben az esetben annak az a priori esélye, hogy X. Y. követte el a cselekményt $1 : 2$, egy személy szúrta, kettő nem.⁹

A bizonyíték keletkezésének a valószínűsége abban az esetben, ha a terhelt az elkövető, illetve abban az esetben, ha ártatlan, első hallásra nehezen értelmezhető. A könnyebb megértés érdekében, az előző példánál maradva, tegyük fel, hogy biztosan tudjuk, hogy aki szúrta, A vércsoportú. (Ennek a vércsoportnak a gyakorisága Magyarországon kb. ötven százalék.) Tekintettel arra, hogy a feltételezés szerint biztosan tudjuk, hogy A vércsoportú, aki szúrta, a bizonyíték (A vércsoport) keletkezésének a valószínűsége az elkövető esetén száz százalék (azaz egy). Mivel a lakosság fele A vércsoportú, így annak az esélye, hogy egy ártatlan személy tartozzon ebbe a vércsoportba,

⁷ Például ha valaminek az esélye $5 : 1$, akkor ötször ritkábban kerül sor az adott kimenetre (például 6-ost dobunk egy dobókockával), mint annak az elmaradására (1-est, 2-est, 3-ast, 4-est vagy 5-öst dobunk), azaz ötször gyakrabban marad el a kívánatos kimenet, mint ahogy megvalósul. Tehát átlagosan minden hatodik esetben dobunk 6-ost. Már itt hangsúlyozom, hogy átlagosan minden hatodik esetben kerül sor arra, hogy 6-ost dobunk. Nyilván mindenki tudja saját társasjáték-tapasztalatai alapján, hogy vannak olyan peches napok, amikor sehogy sem sikerül 6-ost dobni, a többiek már a cél közelében járnak, míg mi a startmezőn toporgunk. Máskor meg egymás után csak úgy szórjuk a 6-osokat.

⁸ <http://de.wikipedia.org/wiki/A-priori-Wahrscheinlichkeit>

⁹ Utóbbi két személy cselekményének büntetőjogi értékelésétől tekintsünk el, jelen esetben csak arra keressük a választ, hogy ki szúrta.

1 : 2, azaz a valószínűség ötven százalék, így ötvenszázalékos valószínűséggel keletkezik a bizonyíték.

Ha az előbbi adatokat behelyettesítjük a képletünkbe, a következőképpen alakul az a posteriori esély:

$$E_{a\text{-posteriori}} = p_b/p_a * E_{a\text{-priori}} = 1/0,5 * 1/2 = 1$$

A későbbiekben bizonyítani fogom, hogy ez az eredmény nem helyes, egyelőre azonban maradjuk a kapott eredmény értelmezésénél.

Az a posteriori esélye annak, hogy a vádlott az elkövető 1 : 1, azaz ugyanakkora eséllyel ő az elkövető, mint amekkora eséllyel nem ő.

Alakítsuk át némileg a képletet, annak érdekében, hogy könnyebben felismerjük azokat a tényezőket, amelyek az eredményt befolyásolják. A bizonyíték keletkezésének a valószínűségét ártatlanság esetén (p_a) az adott tulajdonság megjelenési arányaként értelmezhetjük. Ezt tört formájában a következőképpen fejezhetjük ki: 1/gyakoriság

Gyakoriságon jelen esetben azt értem, hogy a kérdéses tulajdonság átlagosan hány személyenként jelenik meg egyszer.

Legyünk továbbra is biztosak abban, hogy az elkövetőt jellemzi az adott tulajdonság, így a következőket kapjuk:

$$p_b/p_a = 1/1/\text{gyakoriság} = 1 * \text{gyakoriság}/1 = \text{gyakoriság}$$

Behelyettesítve ezt a Bayes-képletbe:

$$E_{a\text{-posteriori}} = \text{gyakoriság} * E_{a\text{-priori}}$$

Tegyük fel továbbra is, hogy a cselekményt egy személy követte el, azaz:

$$E_{a\text{-priori}} = 1/(\text{szóba jöhető személyek}-1).$$

Az előbbiek alapján tehát a Bayes-képlet ekként egyszerűsödik:

$$E_{a\text{-posteriori}} = \text{gyakoriság}/(\text{szóba jöhető személyek}-1)$$

Eljutottunk tehát ahhoz, amit matematika nélkül is feltételezhettünk, annak esélyét, hogy a terhelt az elkövető, két tényező befolyásolja:

1. Milyen gyakran jelennek meg ártatlan személyek esetén az elkövetőnél azonosított tulajdonságok, azaz mennyire ritka tulajdonságot sikerült azonosítani.
2. Mekkora a szóba jöhető személyek köre.

Milyen módon tudja tehát a nyomozó hatóság növelni annak az esélyét (bizonyosságát), hogy a terhelt követte el a bűncselekményt?

Ennek egyik módja, hogy a gyakoriságot növeli, azaz olyan minőségű és mennyiségű bizonyítékot kutat fel, amelyek együtteséből megnyugtató bizonyossággal¹⁰ lehet arra következtetni, hogy a vádlott elkövette a terhére rótt cselekményt.

A másik módja az lenne, hogy csökkenti az a priori valószínűséget. Feltételezve, hogy az elkövetők számára való feltételezés nem változik, azaz a számláló állandó, ennek az egyetlen módja, hogy a nevezőt csökkenti. Mint ahogy korábban kifejtettem, a nevező az ártatlan személyek számát tartalmazza, azaz egy elkövető esetén az összes szóba jöhető személy–1.

És ezzel el is jutottunk a Bayes-analízis egyik kulcsproblémájához: ki az összes szóba jöhető személy?

Ezt meghatározni a nyomozó kompetenciája. Általános recept nem adható, mindig az adott bűncselekmény függvényében kell ezt meghatározni. Például nyilván csak olyan személy követhette el a bűncselekményt, aki számára az elkövetés helye elérhető közelségben van. Az elérhető közelség teljesen más egy besurranásos lopás és egy bérgyilkosság esetén, előbbi esetben aligha utazik az elkövető több száz kilométert, míg utóbbinál az sem elképzelhetetlen, hogy a bérgyilkos felszáll egy repülőre, elrepül az adott városba, ott elvégzi a melót, majd hazarepül, ez pedig akár más kontinens is lehet. Természetesen a besurranásos lopás esetén elvileg bárki a helyszínen lehet, például a Balaton környékén egy turista is „dolgozhat”, de az megbecsülhető, hogy adott napon hányan tartózkodtak a környéken. Az analízis alkalmazásához ugyanis nem a szóba jöhető személyek kilétét, hanem a számát kell ismerni, azaz a nyomozó tapasztalata, intuíciói alapján meghatározza az alaphalmaz nagyságát, amelybe nem sorolja bele azon személyeket, akik szerint semmi-

¹⁰ Megítélésem szerint százszázalékos valószínűség mellett nem állítható, hogy valaki elkövette a terhére rótt cselekményt, ugyanis minden egyes bizonyíték esetén (legalább elvi szinten) el kell ismerni annak a lehetőségét, hogy az keletkezhetett akkor is, ha a vádlott ártatlan. Bármilyen csekély is ennek a valószínűsége az egyes bizonyítékok esetén, és bármilyen bizonyíték áll rendelkezésre, végeredményben mindig megmarad annak a lehetősége (még ha adott esetben ennek aránya egy a több száz milliárdhoz is), hogy a vádlott ártatlan.

képpen sem követhették el a cselekményt. Az alaphalmaz meghatározása bizonyos mértékben szubjektív, éppen ettől a szubjektív elemtől válik a valószínűség Bayes értelmezésében a meggyőződés fokává.

Az a priori érték, megítélésem szerint, a nyomozás során nem csökkenthető, ugyanis a kezdeti benyomás után érkező minden információ már a bizonyítékok értékelése körében veendő figyelembe. Például egy tanú azt állítja, hogy látta az elkövetőt, aki 30-40 éves 185-190 cm magas férfi. Ekkor azt kell meghatározni, hogy a nyomozó által kijelölt alaphalmazban mekkora gyakorisággal fordulnak elő ilyen személyek. Fontos hangsúlyozni, hogy nem a teljes populációra vetítve, hanem az alaphalmazra vetítve kell meghatározni az adott tulajdonság gyakoriságát. Tegyük fel, hogy a nyomozó ennek a bűncselekmények a kapcsán feltételezte, hogy az elkövető 15 és 60 év között van, így a koruk alapján bizonyos személyeket már kizárt, a 15 és 60 év közöttiek körében pedig nyilvánvalóan magasabb a 30 és 40 év közöttiek aránya, mint a teljes népesség esetén.

Az a priori értéket bizonyos esetben az új információk birtokában még növelni is kell. Az ártatlan vádlottra nézve éppen abban rejlik a legnagyobb veszély, hogy a nyomozó túl szűken határozza meg az alaphalmazt, és ezáltal a tényleges elkövető a halmazon kívül marad.¹¹ Azaz a nyomozó abban az esetben jár el helyesen, ha az alaphalmaz meghatározásakor csak olyan személyeket vesz ki, akik megítélése szerint biztosan nem követhették el a cselekményt, azok a személyek, akiknek esetén ebben nem biztos, nem zárhatók ki.

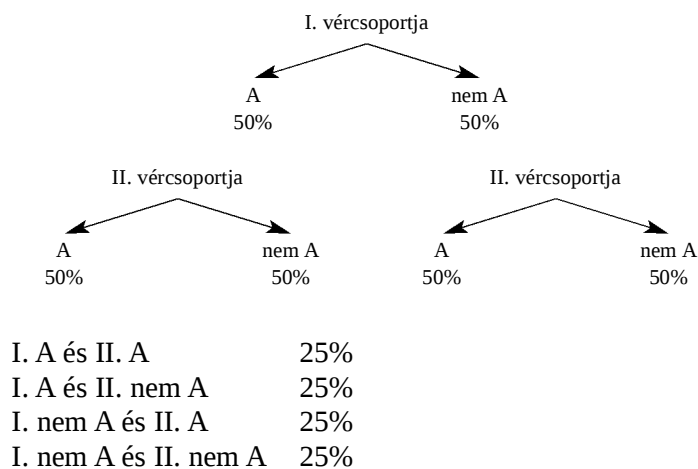
Ha biztosra akarunk menni, hogy semmi esetre sem rekesztjük ki a tényleges elkövetőt, több tízmilliós, százmilliós alaphalmazok is lehetnek, feltételezve, hogy valamennyien azonos a priori eséllyel pályáznak az elkövetésre, még a legmeggyőzőbb bizonyítékok esetén is elenyésző lenne annak a valószínűsége, hogy a vádlott követte el a cselekményt. De miért kellene azonosnak tekinteni az elkövetés valószínűségét? Az, hogy nem zárhatunk ki bizonyos személyeket, nem jelenti azt, hogy azonos valószínűséggel követték el a cselekményt, mint mások. Laplace szerint, addig kell minden kimenetnek (minden embernek) azonos esélyt adni, amíg nincsen annak konkrét oka, hogy mást feltételezzünk. Tekintettel arra, hogy a földrajzi távolság, a kor stb. olyan tényezők, amelyek alapján megalapozottan feltételezhető, hogy nem azonos eséllyel követték el az egyes emberek a bűncselekményt, az a priori érték korrigálható e tények alapján. Képzeljük el, hogy dobókockánkat (amelynél min-

¹¹ Bizonyára mindannyian tudunk olyan esetekről, hogy az újabb lehetséges elkövető csak az után került képbe, hogy az első eljárás már befejeződött, gondoljunk csak a móri ügyre.

den számnak azonos esélye van alulra kerülni) lecseréljük egy olyan testre, amelynek számos oldala van, és ezek eltérő méretűek, vannak egészen aprók és óriások is, így nyilván az egészen pici oldal nagyon ritkán, míg a hatalmas sokkal gyakrabban kerül alulra. Azaz nem zártuk ki senkinek az „esélyét”, hogy ő legyen az elkövető, de figyelemmel voltunk az eltérő tényezőkre.

Az előbbiekből láthatjuk, hogy a helyes a priori érték meghatározása nagy nehézséget okoz, sajnos hasonlóan bonyolult a valószínűségek helyes értelmezése is. Mint említettem, a gyakoriság meghatározása kapcsán figyelemmel kell lenni az a priori halmazba tartozó elemeknek a teljes népességtől való eltérésére. A helyzetet azonban tovább bonyolítja, hogy az a priori halmazba tartozó elemek kölcsönhatásban vannak egymással. Tehát, maradva a kocsmai szurkálós példánál, az igaz ugyan, hogy egy embernek az esélye arra, hogy A vércsoportja legyen, ötven százalék, de két embernek nem kétszer ötven százalék, azaz nem száz százalék az esélye arra, hogy közülük legyen egynek A vércsoportja. A problémát még bonyolítja, hogy két ember esetén elképzelhető az is, hogy egyikük sem, vagy mindkettőjük A vércsoportú.

Foglaljuk össze egy gráfban, hogy milyen két ember (I. és II.) esetén milyen lehetséges kimenetek vannak, és ezeknek mekkora a valószínűségük: Tehát a következő kimeneteket és valószínűségeket kapjuk:



Huszonöt százalék a valószínűsége, hogy a három szóba jöhető személy közül csak X. Y. A vércsoportú, ötven százalék a valószínűsége, hogy a két másik személy közül egy A vércsoportú, és huszonöt százalék a valószínűsége, hogy mindhárman A vércsoportúak. Ebből pedig azt a következtetést vonhat-

juk le, hogy 25 százalék a valószínűsége (1 : 3 az esélye), hogy biztosan X. Y. követte el a cselekményt, és 58,33 százalék¹² annak a valószínűsége (1,4 : 1 az esélye), hogy X. Y. követte el a cselekményt.

Az előbbi számítás alapján két jelentős következtetést vonhatunk le:

1. Lényegesen eltér annak a valószínűsége, hogy valaki elkövette a cselekményt annak a valószínűségétől, hogy csak ő követhette el a cselekményt.
2. A kapott eredmény eltér a Bayes-képletből kapott eredménytől. Ezek szerint akkor valamelyik számítás hibás?

Az eltérést a következők okozzák. Feltételezhetjük, hogy az alaphalmazba kerülés esélye független a vércsoporttól, így a vércsoport szempontjából véletlen mintavételnek tekinthetjük a három személy kiválasztását, akik részt vettek az incidensben. Ha az alaphalmazba bevont személyek számát minden határon felül növelnénk, akkor az A vércsoport relatív gyakorisága közelítene a 0,5-hez, ezzel érvényesülne a nagy számok törvénye¹³. Csakhogy a terhelten kívüli két személy annyira kevés, hogy nem képes leképezni a teljes népességre jellemző eloszlást, és ezért mutatkoznak jelentős eltérések az egyes konkrét minták esetén az elvárt egy A vércsoportú személytől. A teljes megértés kedvéért ismét egy dobókocka-analógia: tegyük fel, hogy 600 dobásos sorozatokat produkálunk egy kockával, a 6-os dobások száma szinte soha nem lenne legalább 200, de 6-os sorozat esetén bizonyos idő után biztosan sikerülne olyan sorozatot produkálni, amelyben legalább két 6-os van.¹⁴

A Bayes-képletben alapul vett gyakoriságok csak nagy számok esetén érvényesülnek nagy relatív pontossággal, így csak nagy számok esetén várhatunk pontos eredményt. Feltehetően ebből is ered az a tény, hogy az angolszász joggyakorlatban 100-as esély alatt, mint nem releváns, kizárják a bizonyítékot¹⁵.

Nagy a posteriori eredmény esetén¹⁶ a Bayes-képlet kellő pontossággal határozza meg az elkövetés esélyét/valószínűségét. Alacsony a posteriori

12 $58,33\% = 25\% + 50\% \cdot 50\% + 33,33\% \cdot 25\%$. Huszonöt százalék annak a valószínűsége, hogy csak X. Y.-nak van ilyen vércsoportja, így ebben az esetben ő az elkövető, ötven százalék a valószínűsége, hogy még egy személynek A a vércsoportja, így X. Y. ötven százalék valószínűséggel az elkövető, és huszonöt százalék annak a valószínűsége, hogy mindhármuknak A vércsoportja van, ebben az esetben X. Y. elkövetési valószínűsége 33,33 százalék, azaz összesen 58,33 százalék.

13 A nagy számok törvényének a megfogalmazását lásd Varga József: Mérték- és valószínűségelméleti alapok. Janus Pannonius Tudomány Egyetem Közgazdaságtudományi Kar, Pécs, 1995, 172–173. o.

14 Utóbbi valószínűsége nagyjából 26,5 százalék.

15 http://www.drtatar.hu/bayes_de.htm

16 Megítélésem szerint ilyennek tekinthető a 100-at meghaladó szám, ebben az esetben a Bayes-képlet már elenyésző tévedést mutat. Persze más kérdés, hogy a 99 százalékos elkövetési valószínűség elegendő-e az elítéléshez, de ez már nem a matematika területe.

eredmény esetén pedig úgyis a vádlott felmentésének van helye bizonyított-ság hiánya miatt, így a végeredményt nem befolyásolja a téves valószínűség.

Végül még egy kérdéskörrel, a hibás pozitív azonosítás kérdésével foglal-kozom. A probléma lényegének szemléltetéséhez tegyük fel, sikerült megál-lapítani, hogy az elkövető kétséget kizáróan olyan tulajdonságkombinációval bír, amelynek előfordulási valószínűsége 100 000 000 : 1 (azaz százmillió az egyhez).¹⁷ Megvizsgálva a vádlottat, a szakértő megállapította, hogy ő ezzel a tulajdonságkombinációval bír. Ezek után a bíró feltehetően gondolatban már az ítéletet fogalmazza, az ügyész gondolatban már koccint a kollégákkal, hogy sikerült fokozni a 99 százalékos vádhatékonyt. De mit tegyen a védő? Búcsúzzon el szép sorban azoktól a dolgoktól, amelyeket a felmentés esetére kikötött sikerdíjból tervezett beszerezni?

Mielőtt ezt megtenné, javaslom, hogy még egy kérdést tegyen fel a szak-értőnek: elképzelhetőnek tartja-e a szakértő, hogy egymillió esetből egyszer téved a teszt, azaz hibásan határozza meg, hogy az adott személyt jellemzik-e az adott tulajdonságok? (Továbbra is csak azt feltételezzük, hogy a tévedés a vádlott vizsgálata során merülhetett fel, az elkövető biztosan bír ezzel a tulaj-donságkombinációval.)

Vélhetően a szakértő elismeri, hogy egymillió esetből egyszer tévedhet. De mit ér ez? Ez olyan kicsi arány, hogy ennek alapján nem ingatható meg a szakvélemény. Vagy mégis? Nos a hatás szemléltetésére hívjuk segítségül a Bayes-tételként elterjedt képletet:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

$P(A)$ A esemény a priori valószínűsége

$P(B)$ B esemény a priori valószínűsége

$P(B|A)$ B esemény valószínűsége feltételezve, hogy A esemény bekövetkezett

$P(A|B)$ A esemény valószínűsége feltételezve, hogy B esemény bekövetkezett

Értelmezzük a korábbi adatokat!

$P(A)$ annak a valószínűsége, hogy egy emberre jellemző az adott tulajdonság, jelen esetben 0,00000001. B jelöli azt az eseményt, hogy a teszt egy adott embernél pozitív lett, mivel nem tesztelünk valamennyi embert, nem tudjuk pontosan, hogy B hány esetben következett be. $P(B|A)$ annak a valószínűsége,

¹⁷ Ilyen arány például DNS-teszt esetén elképzelhető. A DNS-vizsgálat azonban nem a példában feltéte-lezett módon történik.

ge, hogy a teszt a tulajdonságot felismeri, azaz 0,999999, ennek a komplementere 0,000001, ezt jelölje $P(B|A^c)$. Szükségünk van még $P(A)$ komplementerére, ezt jelölje $P(A^c)$. Felhasználva az alábbi összefüggést:

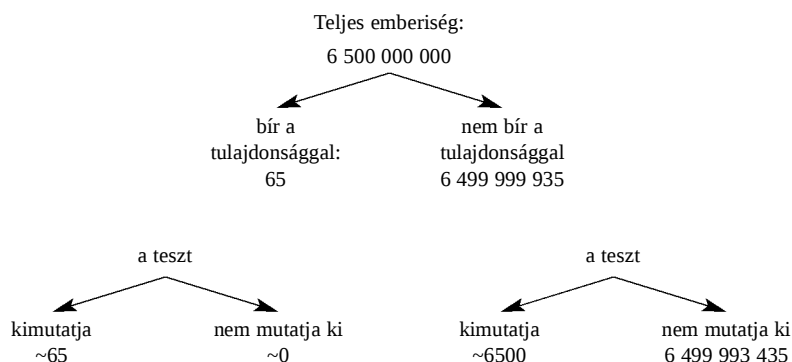
$$P(B) = P(B) \cdot 1 = P(B|A^c) \cdot P(A^c) + P(B|A) \cdot P(A)$$

Ezek alapján a Bayes-tételt átalakítva a következőt kapjuk:

$$P(A|B) = P(B|A) \cdot P(A) / ((P(B|A^c) \cdot P(A^c) + P(B|A) \cdot P(A)) = 0,999999 \times 0,00000001 / (0,000001 \times 0,99999999 + 0,999999 \times 0,00000001) = 0,0099$$

Tehát annak a valószínűsége, hogy az adott ember tényleg bír a tulajdonsággal, kevesebb mint egy százalék.¹⁸

Más megközelítésben ugyanez a probléma:



Azaz a teszt szerint összesen 6565 ember bír az adott tulajdonságkombinációval, míg valójában csak 65. Tehát annak a valószínűsége, hogy az, akiről a teszt kimutatta, hogy bír a tulajdonságkombinációval, ténylegesen is bír vele:

$$65/6565 = 0,0099,$$

tehát kevesebb mint egy százalék. Így a védőnek már csak arra kell hivatkoznia, hogy a kétséget kizáróan nem bizonyított tény nem értékelhető a vádlott terhére, azaz úgy kell tekinteni, mintha a vádlott nem bírna az adott tulajdonsággal. Így máris közelebb jutottunk a sikerdíjhoz.

Ha végiggondoljuk előbbi példánkat úgy, hogy a tulajdonságkombináció megjelenési aránya 10 000 000 : 1, illetve 1 000 000 : 1, megállapíthatjuk,

¹⁸ A képletet és egy másik példát lásd: <http://de.wikipedia.org/wiki/Bayestheorem>

hogy előbbi esetben kb. kilenc, utóbbiban ötven százalék a valószínűsége annak, hogy a pozitív teszteredmény helyes.

Mindezekből levonható az a következtetés, hogy adott megbízhatósági szint mellett minél ritkább tulajdonságot azonosítunk, annál nagyobb a téves pozitív azonosítás valószínűsége. Ez vonatkozik valamennyi bizonyítási eszközre.

A rendkívül ritka tulajdonságok így, megítélésem szerint, a pozitív azonosításra csak korlátozottan alkalmasak, a kizárásra azonban jól használhatók, hiszen az előbbieken láttuk, hogy szinte senkit sem zár ki a teszt azok közül, akikre jellemző a tulajdonság.

Ennyi matek után mégis milyen konklúziót vonhatunk le? Teljes bizonyosságot (százszázalékos valószínűséget) matematikai alapon soha sem érhetünk el abban a kérdésben, hogy a terhelt elkövette-e a terhére rótt bűncselekményt. Végül soron a bírónak a személyes meggyőződése alapján kell abban a kérdésben állást foglalnia, hogy a vádlott elkövette-e a cselekményt, vagy sem, ezt a felelősséget a matematika nem vállalhatja át. A bírót nem helyettesíthetők számítógépekkel, amelyek bonyolult valószínűség-számítás után ítéletet mondanak arról, hogy a terhelt elkövette-e a cselekményt, vagy sem. A matematika arra alkalmas, hogy rávilágítson bizonyos összefüggésekre, segíthet abban, hogy az egyes bizonyítékokat a helyükön kezeljük (a Bayes-tétel különösen a hibás pozitív azonosítás súlyos veszélyére hívhatja fel a figyelmünket), a döntést azonban nem vállalhatja át a bírótól. Azaz a matematika segítheti, de nem helyettesítheti a bírót a döntéshozatalban.