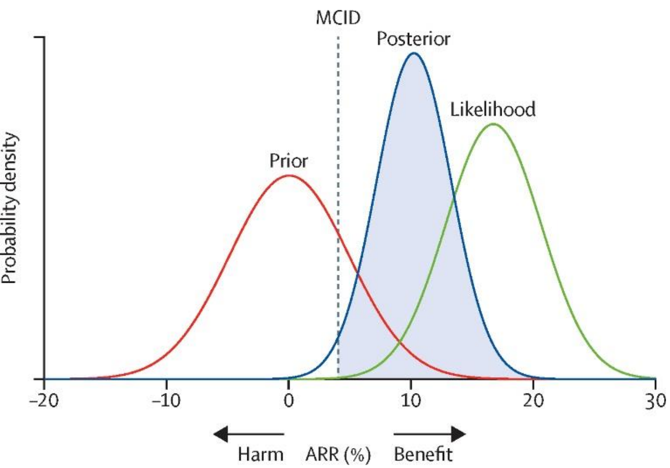
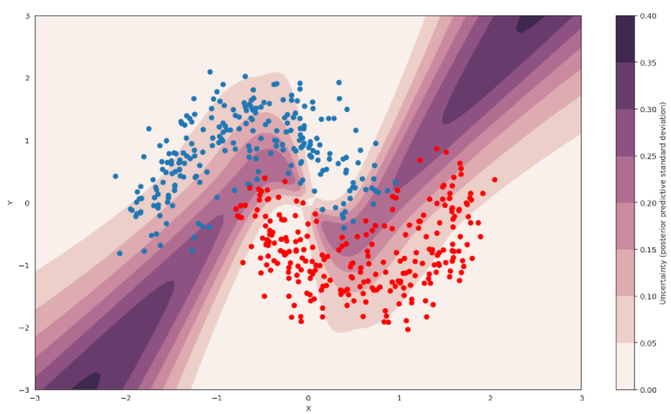
Bayes-statisztika

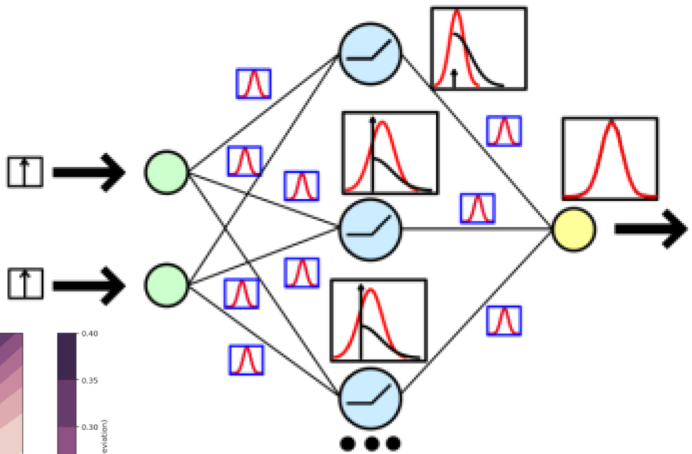
Hogyan tudjuk az előzetes ismereteinket, *várakozásainkat* beépíteni a döntéseinkbe?

avagy

Hogyan lehet egy kérdést bizonytalan adatokból megválaszolni?







Schay Gusztáv

A dichotóm (I/N) döntési fa

Kérdés

Eldöntendő kérdéssé alakítás ha szükséges

H0 megfogalmazása

Az adataink H0-hoz tartozása számítható kell hogy legyen, ezért H0 többnyire egy ismert, referencia-helyzetet mint előzetes feltevést jelent.

Eldöntjük mennyire ragaszkodunk H0-hoz, azaz mekkora az a valószínűség ami alatt már nem tartjuk meg az előfeltevést. -> ezt hívjuk **szignifikancia szint**nek

α vagy α_{max}

Kiszámítjuk a valószínűségét annak, hogy véletlenszerűen *legalább ennyire* eltérjen egy eredmény a H0-tól mint amennyire a mi adataink eltérnek.
-> Ez valójában egy feltételes valószínűség: P(legalább ekkora eltérés | H0 igaz)

Döntés

P < szign.

P >= szign.

H0-t elvetjük

H0-t megtartjuk

Azaz nem vetjük el!

A döntés csupán valószínűségi!

Mindig fennáll a tévedés valószínűsége, de jobb megoldás nincsen. Ha ezzel a módszerrel hozunk döntést, akkor legalább abban biztosak lehetünk hogy *nagy átlagban* a lehető legkevesebbet fogunk hibázni.

Döntésünk

Valóság (ezt vagy megtudjuk utólag, vagy nem)	**H0-t megtartjuk**	**H0-t elvetjük**
Valójában H0 igaz	**Helyes döntés**	Első-fajú hiba α
Valójában H0 hamis	Másod-fajú hiba β	**Helyes döntés**

„A legjobb orvos a patológus, mert mindent tud. Csak általában már későn…”

A szignifikancia-szinttel azt határozzuk meg, hogy mekkora α-t vagyunk hajlandóak tolerálni.

Próba ereje: Egy ismert alternatív hipotézis igaz volta esetén esetén mekkora valószínűséggel veti el H0-t.

Valójában minden *állításunk és valószínűség* feltételes:
„hypo thesis” = „az alulra állított”

$P(A|H_0,C)$: nemcsak a nullhipotézis, de a (kísérleti) körülmények is feltételt jelentenek.

C azonban nehezen meghatározható!
-> szükségünk van az *interszubjektív konszenzusra*
(bizonyos előfeltételeket, axiómákat mindenki el kell hogy fogadjon, másképp tudomány sincsen)

De: **C**-ben benne van a szubjektum is!

P(eső | **C**) vagy P(jól sikerült a műtét | **C**)

Sokszor használunk intuitív, *frekventista módon* nem kezelhető „valószínűség” fogalmakat.

„Szerintem ma nem fog esni.”
„Nyugodjon meg, a műtét jól sikerült.”

-> ezeket is be kellene építeni valahogyan a döntési mechanizmusokba!

Frekventista valószínűség: $P_A = \lim_{N\to\infty} k_A$

Vigyázat, valójában itt is P(A|C) és k_A|C van!

Szubjektivista valószínűség:
P_A = „egy esemény (vagy más!) bekövetkeztében való *meggyőződés mértéke*”

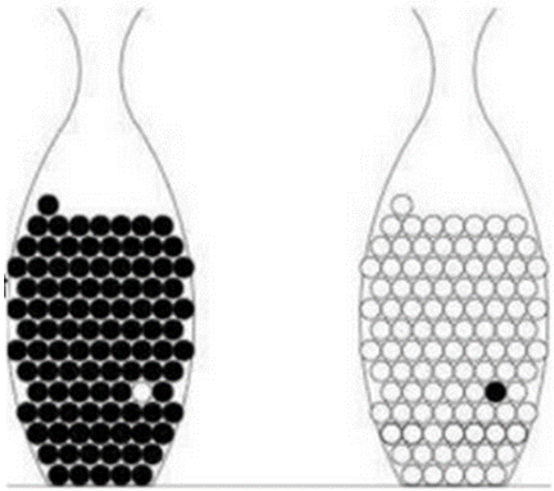
-> de az interszubjektív konszenzus továbbra is feltétel!
Akkor lehet értelmes P-t definiálni, ha
P(A|C)-t *mindenki azonosnak ítéli meg aki C-nek birtokában van*.

-> ugyanazt a leletet olvasva, ugyanazzal a szaktudással rendelkező kollégák az Ω eseménytér (pl. lehetséges diagnózisok) valószínűségeit egyformán adják meg.

Akkor minden relatív?

A közös nevező az **urnamodell**.

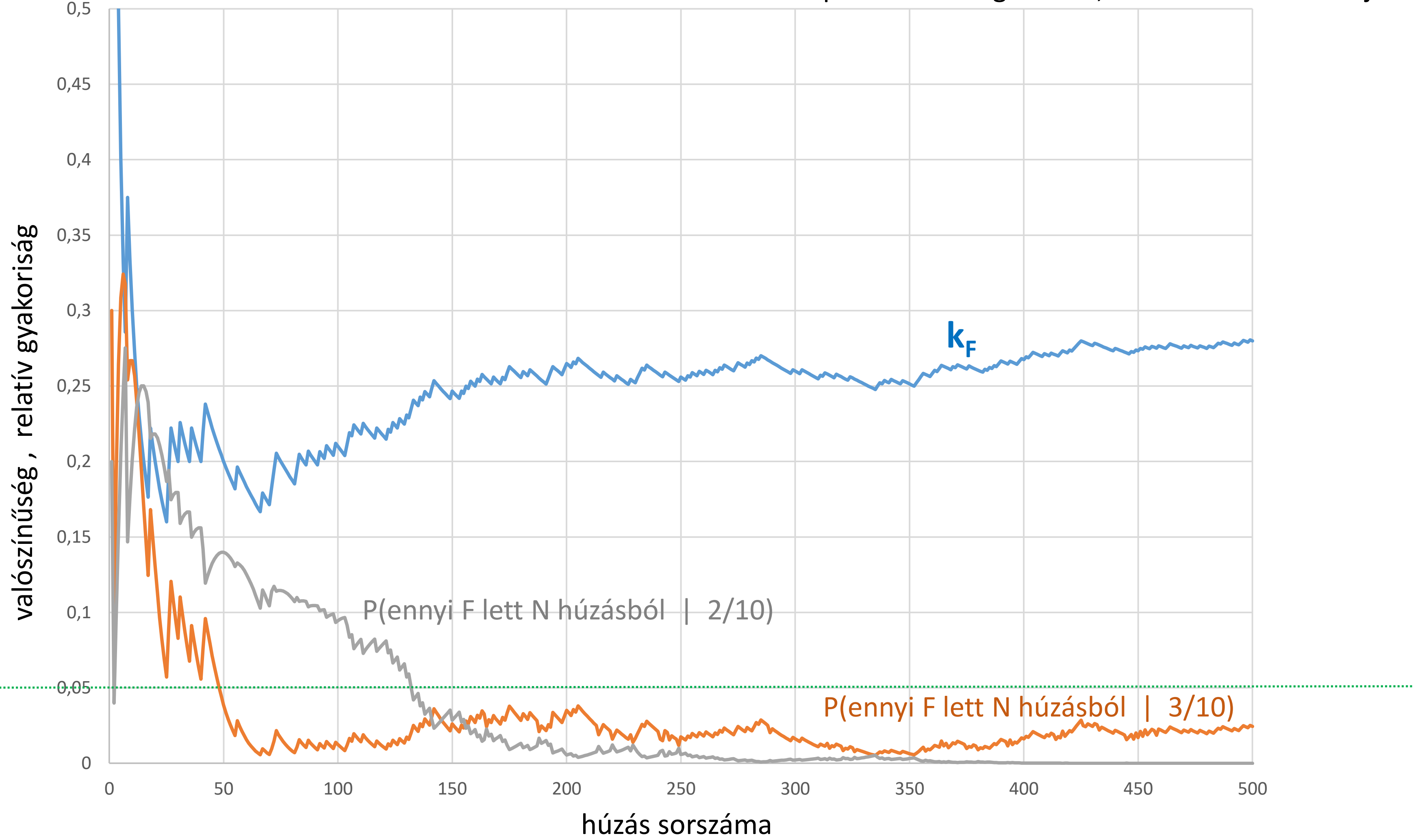
A valószínűséget megadhatjuk egy urnával



P(A|C) = P(F golyó | adott urnával játszunk)

A dichotóm döntési fa is gondokat okoz.

-> néha több nullhipotézist is megtartunk, máskor mindet elvetjük…



Valóság: 3/10-es urnából húztunk.

A szubjektív valószínűség definiálható axiomatikusan is.
Ehhez logikai szabályokat követve nemcsak eseményekre, de állításokra vonatkozó algebra is kialakítható.

$$\sim\sim a = a, \qquad (1)$$

$$a \cdot b = b \cdot a, \quad (2) \qquad a \vee b = b \vee a, \quad (2')$$

$$a \cdot a = a, \quad (3) \qquad a \vee a = a, \quad (3')$$

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c = a \cdot b \cdot c, \qquad (4)$$

$$a \vee (b \vee c) = (a \vee b) \vee c = a \vee b \vee c, \qquad (4')$$

$$\sim(a \cdot b) = \sim a \vee \sim b, \qquad (5)$$

$$\sim(a \vee b) = \sim a \cdot \sim b, \qquad (5')$$

$$a \cdot (a \vee b) = a, \quad (6) \qquad a \vee (a \cdot b) = a. \quad (6')$$

~ = nem
. = és
v = vagy

(1) És még 5 szabály független egymástól, a többi levezethető.

NEM kell megtanulni!
Az a lényeg hogy VAN matematikai alapja.

I. $(cb)|a = F(c|ba\ ,\ b|a)$

II. $\sim b|a = S(b|a)$

$F(x,y)$ és $S(x)$ meghatározandó függvények

-> ha a szabályokat követjük, és keressük a legegyszerűbb megoldást, akkor $F(x,y) = F(x)*F(y)$ és $F(x)=x$, valamint $S(x) = 1-x$.

-> Ez viszont megegyezik a frekventista Kolmogorov-féle számolással!

$b|a + \sim b|a = 1$

és

$a|a + \sim a|a = 1$

biztos és lehetetlen események

NEM kell megtanulni!

Az a lényeg hogy VAN matematikai alapja.

A Bayes-féle szubjektivista valószínűségfogalom tág:
nem követeli meg a végtelen ismételhetőséget, és a relatív gyakoriság létezését
DE
Abban a határesetben amikor létezik a frekventista valószínűség, akkor a bayes-értelmezésű valószínűség ezzel egyező értéket fog adni.
(interszubjektív konszenzus követelménye)
Számolni úgyanúgy kell mint eddig!

(1) Minden p ***kijelentésre*** *igaz hogy*, $0 \leq P(p) \leq 1$
Ha p biztosan igaz akkor , $P(p) = 1$
Ha p biztosan nem igaz akkor $P(p)=0$
(1) Ha p és q egymást kizárják akkor, $P(p \text{ vagy } q) = P(p) + P(q)$

$P(\text{nem-}p) = 1 - P(p)$

A meggyőződésünk mértékét is standardizálhatjuk:
Minden fogadás

Egy fogadás, szerencsejáték *fair módon* annyit ér amennyi a várható értéke.

Frekventista áthallás (és interszubjektív konszenzus lehetőség)
A fair játékban *átlagban* egyik fél sem veszít

A meggyőződésünk mértékét értelmes viselkedés modelljébe csomagoljuk:
Mekkora *értelmes* fogadást vagyunk hajlandóak kötni?
(tehát ne legyen rajta biztos vesztesége senkinek, csak a játék öröméért fogadunk)

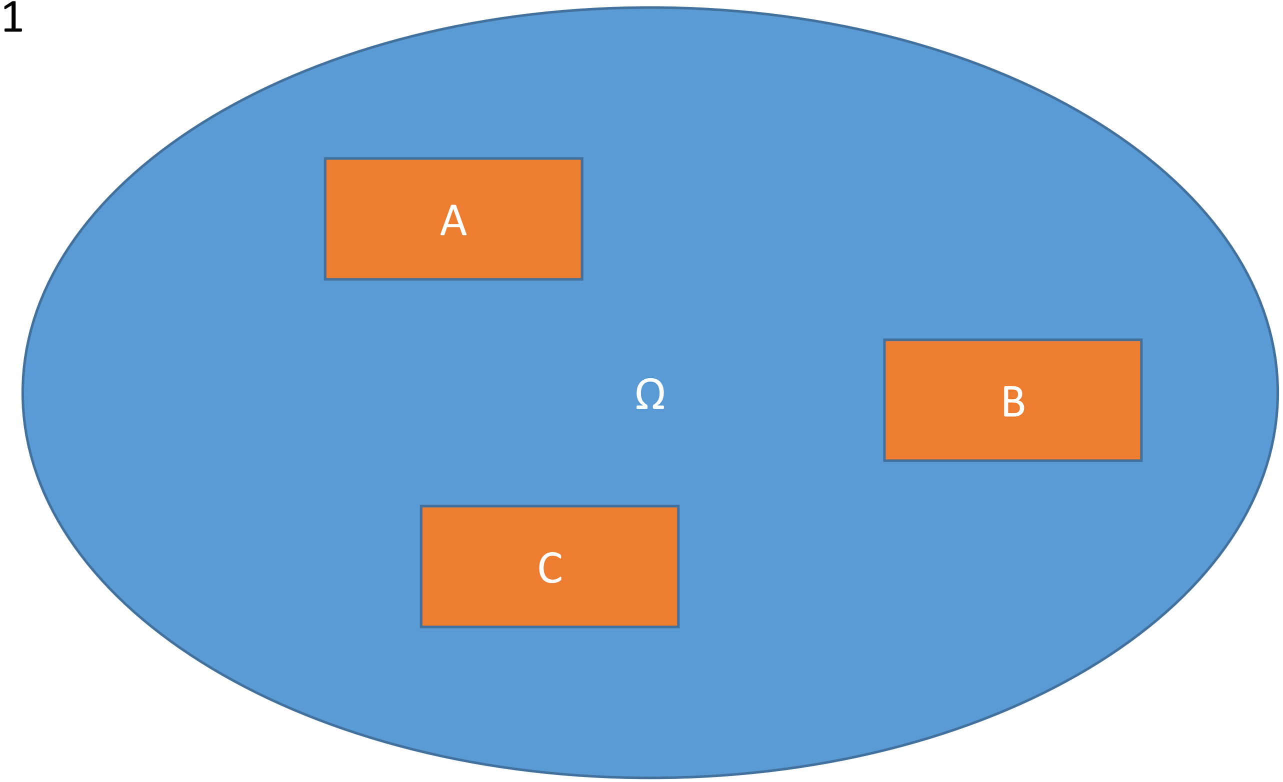
-> az urnához hasonló, de közvetlenül alkalmas a vélekedésel összehasonlítására:
Pl. 1:100000 hogy súlyos mellékhatás alakul ki az Ön esetében.

1:x-hez fogadás azt jelenti, hogy P=1/(1+x) valószínűségű a fogadott esemény, állítás, kijelentés

Becsléseinknek koherenseknek is kell lenniük:

$P(A)+P(B)+P(C) <= 1$

$P(\Omega)=1$



ld.: „Dutch book argument” -> ha nem koherens a valószínűségek kiosztása akkor egy bookmaker rendszerben valahol biztos nyereség, máshol biztos veszteség jelentkezik, ami ellentétes a racionális fogadás előbbi elvével.

Mire jó mindez?

Hipotézisek valószínűsége is megadható.

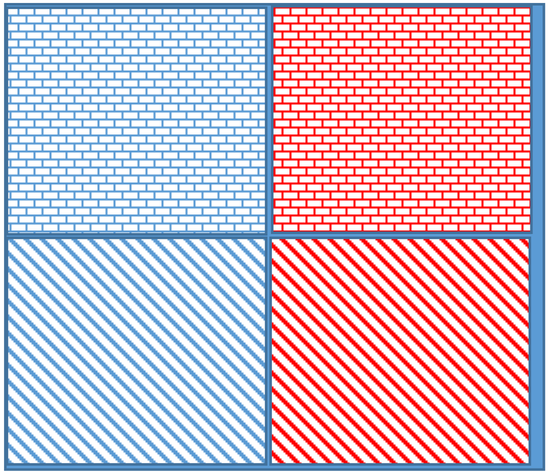
Ha kapunk egy leletet, akkor megmondhatjuk hogy az egyes diagnózis lehetőségek mekkora valószínűségűek.

Megmondhatjuk azt is hogy egy diagnosztikus eredmény mennyi bit információ értékű

Feltételes valószínűség

P(A | B) = annak a valószínűsége, hogy A esemény bekövetkezik, ***feltéve hogy*** B bekövetkezett.

pl. : lázas-e a páciens, *feltéve hogy* tudjuk hogy COVID-19 fertőzött.
5-öst kapok statisztikából *feltéve hogy* eddig minden vizsgám jeles volt.



Ω egy része érdekel minket csupán, a kérdésünk az hogy azon belül milyen az előfordulási valószínűség.

P(kék | csíkos) = kékek valószínűsége a csíkosak között. Itt két csíkos van, abból egy a kék, tehát P(kék | csíkos) = ½

Megjegyzések:

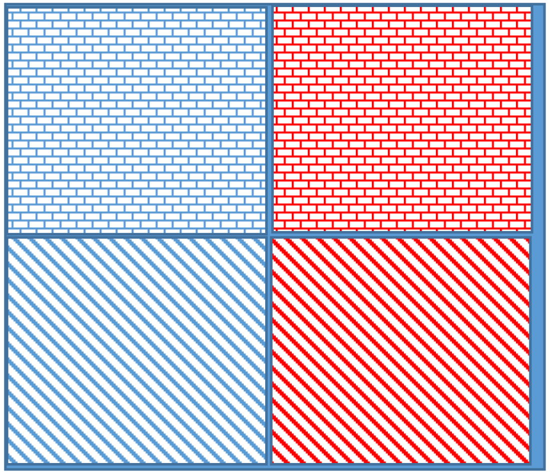
1. független eseményekre P(A|B=P(A)
2. Két eseményre P(AB) = P(A|B)*P(B) ez a „szorzási szabály” vagy Bayes-tétel

Bayes tétel

feltételes valószínűség

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

Két „eseményre" : A,H
H lehet egy hipotézis is.



P(AH) = P(kék ÉS csíkos)
P(H) = P(csíkos)

P(A|H) = P(kék, feltéve hogy tudjuk hogy csíkos)

Átrendezve: P(H) = P(AH)/P(A|H)
-> H valószínűsége kiszámítható az együttes és a feltételes valószínűségekből.

Bayes tétel

feltételes valószínűség $P(A|H) = \frac{P(A \cdot H)}{P(H)}$

Két „eseményre” : A,H
H lehet egy hipotézis is.

Az együttes valószínűség kétféleképpen is megadható, ha P(A) és P(H) is létezik:

$$\mathrm{P}(\mathrm{A} \cdot H) = P(A|H) * P(H) = P(H|A) * P(A)$$

Amiből: $P(H|A) = \frac{P(A|H) * P(H)}{P(A)}$

Vagyis „megfordítottuk” a feltételes valószínűséget.

Ez azt is jelenti, hogy ha H a hipotézis, akkor ki tudjuk számolni éppen azt, hogy A esemény bekövetkezte **után** mekkora H valószínűsége!

Vigyázat: P(A) meghatározásakor minden lehetséges bekövetkezési utat figyelembe kell venni!

Példa: (A számolást magát *nem* kell tudni megcsinálni!)

$$P(H|A) = \frac{P(A|H) * P(H)}{P(A)}$$

Legyen annak valószínűsége (azaz meggyőződésünk mértéke!) az anamnézis alapján hogy a páciensünknek gyomorfekélye van P(H)=15%.

Elvégzünk egy helicobacter pylori PCR tesztet ami a gyomorfekéllyel diagnosztizáltak esetében 92%-ban pozitív. A teszt eredménye pozitív.

Azt is tudjuk, hogy általában ennek a tesztnek a szenzitivitása 0.95, specificitása 0.91, míg a h.pylori becsült prevalenciája 2%.

	+	-	
h.pylori fertőzött	1.90%	0.10%	2%
nem fertőzött	8.82%	89.18%	98%

10.72% a + teszt valószínűsége mindentől függetlenül

P(A) azaz a **+** eredmény két egymást kizáró úton következhet be:

- H igaz
- H hamis

$P(A)=P(AH)+P(A\overline{H})$

Tehát $P(A) = P(H)*P(A|H) + P(\overline{H})*P(A|\overline{H})$

Ha az illető nem gyomorfekélyes akkor a PCR teszt általános adataiból lehet kiindulni.

$$P(H|A) = \frac{0.92 \ * 0.15}{0.15 * 0.92 + (1 - 0.15) * 0.1072} = 0.602 = 60.2\%$$

Vagyis a pozitív teszt után a *meggyőződésünk mértéke* jócskán megnőtt.

-> nagyon hasonló a számolás a diagnosztikus tesztek PPV és NPV számolásához.

-> a prevalencia most is sokat számít

Ha csak két lehetőségünk van (Beteg vagy nem) és csak kétféle úton valósulhat meg mindegyik eredmény (+/-) akkor a számolás relatíve egyszerű

$$LR_{+} = \frac{P(+|Beteg)}{P(+|Egészséges)} \quad \text{és} \quad O_{Betegség} = \frac{P(Beteg)}{P(nem-Beteg)} = \frac{P(Beteg)}{P(Egészséges)}$$

$$O_{Betegség|+} = \frac{P(Beteg|+)}{P(Egészséges|+)} = \frac{\frac{P(Beteg\ \text{és}\ +)}{P(+)}}{\frac{P(Egészséges\ \text{és} +)}{P(+)}} = \frac{P(Beteg\ \text{és}\ +)}{P(Egészséges\ \text{és} +)} =$$

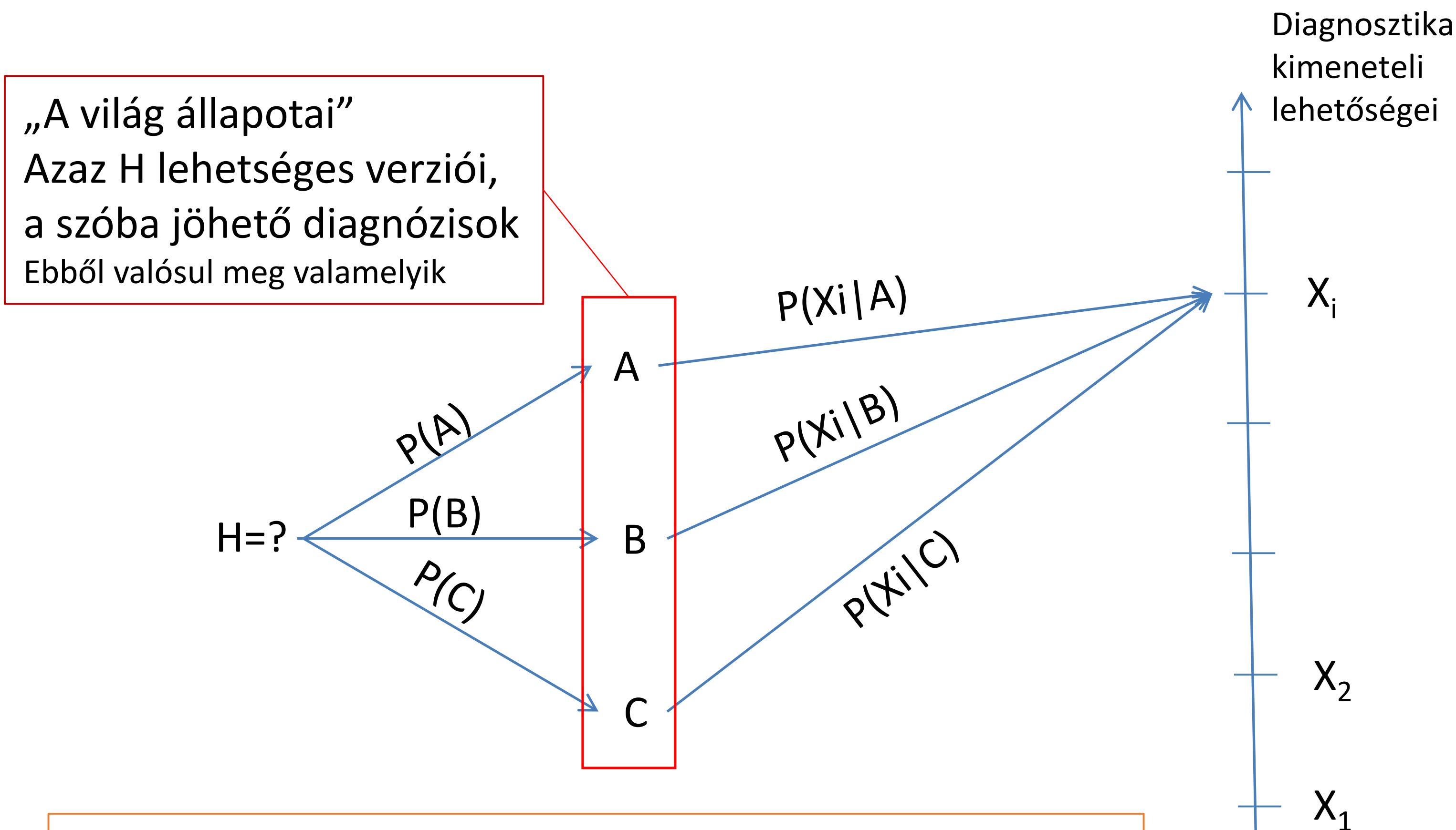
$$\frac{P(+|Beteg) * P(Beteg)}{P(+|Egészséges) * P(Egészséges)} = LR_{+} * O_{Betegség}$$

Hasonlóan megy az LR_{-}

$$P(Beteg|+) * P(+) = P(Beteg\ \text{és}\ +) = P(+|Beteg) * P(Beteg)$$

<table>
<tr><td>prevalence = D/ALL</td><td></td><td></td><td colspan="3">incidence = NEW cases over t time / number at risk</td></tr>
<tr><td>se</td><td>+|D</td><td>prevalence indep</td><td colspan="3">incidence RATE = incidence / t time</td></tr>
<tr><td>sp</td><td>-|H</td><td></td><td>RR_D</td><td>(D|R+)/(D|R-)</td><td>PPV/(1-NPV)</td></tr>
<tr><td>false neg rate
1-se</td><td>-|D</td><td></td><td>RR_H</td><td>(H|R+)/(H|R-)</td><td>(1-PPV)/NPV</td></tr>
<tr><td>false pos rate
1-sp</td><td>+|H</td><td></td><td>LR+</td><td>+|D / +|H</td><td>se/(1-sp)</td></tr>
<tr><td>PPV</td><td>D|+</td><td>prevalence DEP</td><td>LR-</td><td>-|D / -|H</td><td>(1-se)/sp</td></tr>
<tr><td>NPV</td><td>H|-</td><td></td><td>OR_D</td><td>(O_D|R+) / (O_D|R-)</td><td>RR_D / RR_H</td></tr>
<tr><td>false alarm rate
1-PPV</td><td>H|+</td><td></td><td>OR_H</td><td>O_H|R+ / O_H|R-</td><td>RR_H / RR_D</td></tr>
<tr><td>false reassurance rate
1-NPV</td><td>D|-</td><td></td><td>O_D = D/H</td><td>O_D_post = O*LR</td><td></td></tr>
<tr><td>D: diseased</td><td></td><td></td><td colspan="2">R+ risk factor present</td><td>O = p/1-p</td></tr>
<tr><td>H: healthy</td><td></td><td></td><td colspan="2">R- risk factor NOT present</td><td>p = O/1+O</td></tr>
<tr><td>+/- Test result</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="6">diagnostic test cheat sheet (c) G.Schay</td></tr>
</table>

A Bayes tétellel visszafele tudunk következtetni



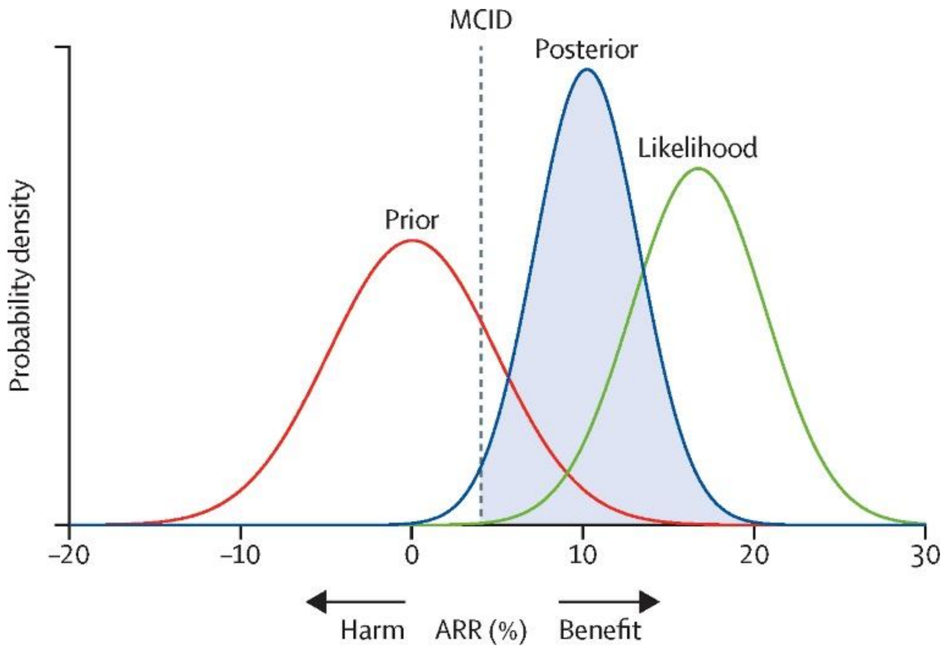
$$P(A|X = Xi) = \frac{P(A) * P(Xi|A)}{\sum_{k=A,B,C} P(k) * P(Xi|k)}$$

Folytonos „Hipotézis-számegyenes” esetén is lehet számolni

$$P(A|X = Xi) = \frac{P(A) * P(Xi|A)}{\sum_{k=A,B,C} P(k) * P(Xi|k)}$$

$$f(h|x) = \frac{f(h) * P(x|h)}{\int_h f(h)P(x|h)dh}$$

Ezzel a számolással a kísérlet, diagnosztika előtt meglevő meggyőződésünk (priori eloszlás, vagy prior) javítható az új ismeretek figyelembe vételével, így kapjuk a **poszteriori** eloszlást. A számoláshoz az „előrefele” mutató feltételes valószínűségeket (Likelihood) ismerni kell.

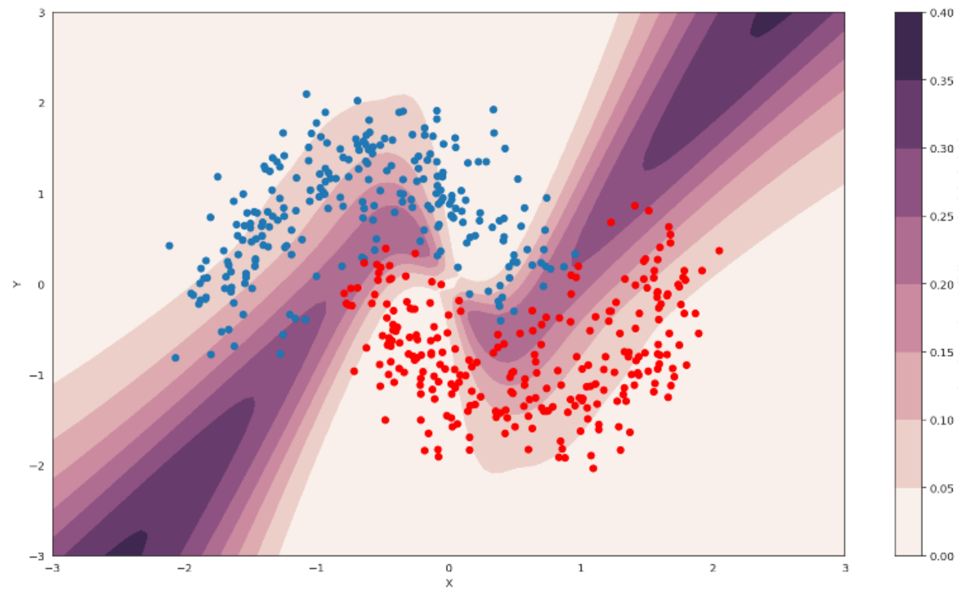


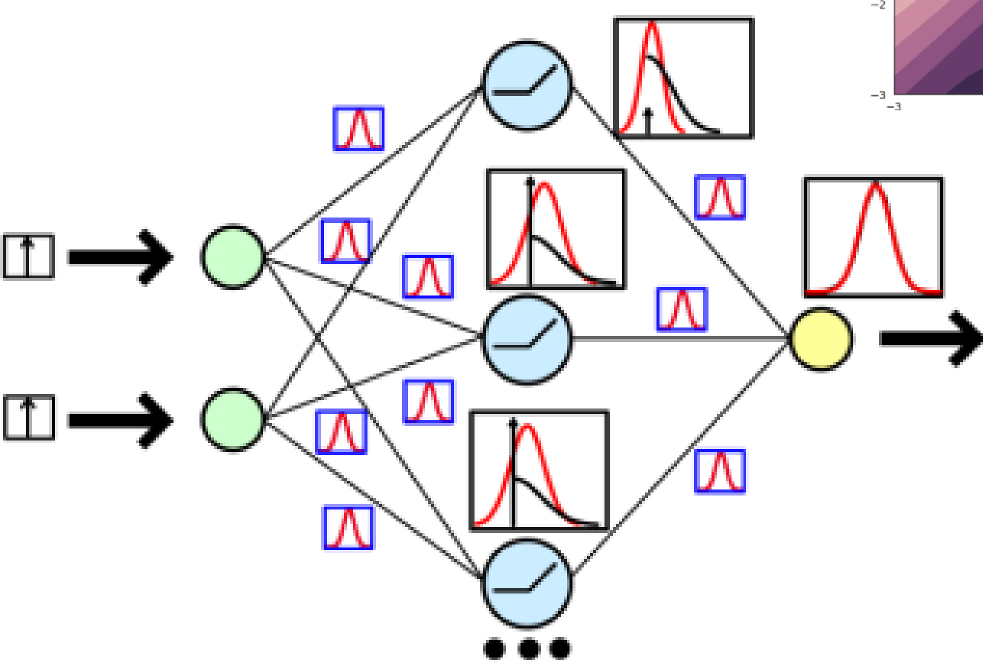
Megjegyzés: az egész számolás két kulcs-feltételen nyugszik:

- Létezik egy priori eloszlás (a kiterjesztett valószínűség-definícióval ez megoldható)
- Kiszámolhatóak a Likelihood értékek, és a teljes Ω-t le tudjuk fedni, nem felejtünk ki semmit.

Orvosi alkalmazási példák

Döntéstámogató rendszerek
Kategorizálás
Kutatás
Mesterséges intelligencia
(neurális hálózatok)





Döntéstámogató rendszer - döntéselmélet

A Bayes elmélet segít betekintést nyerni -> kapunk egy (poszteriori) eloszlást ami a világ számunkra kérdéses részének állapotát mutatja valószínűségi alapon.
(ezzel a bizonytalan helyzetben meglevő **bizonytalanság** a valószínűségekben kódolódik.)

Hogyan döntsünk?

nyereség

Döntés -> Haszon?

Azt a döntést fogjuk választani aminek a várható haszna a legnagyobb.

A **lehetőségekre** vonatkozó preferenciáinkra két axiómát tekintünk érvényesnek:

Teljesség: $1 \preccurlyeq 2 \ \ vagy \ 2 \preccurlyeq 1$

Tranzitivitás: $1 \preccurlyeq 2 \ \ és \ 2 \preccurlyeq 3 \implies 1 \preccurlyeq 3$

Itt $\preccurlyeq$ olyan reláció ami azt jelzi, hogy áldozni is hajlandóak vagyunk a „nagyobb"-ért

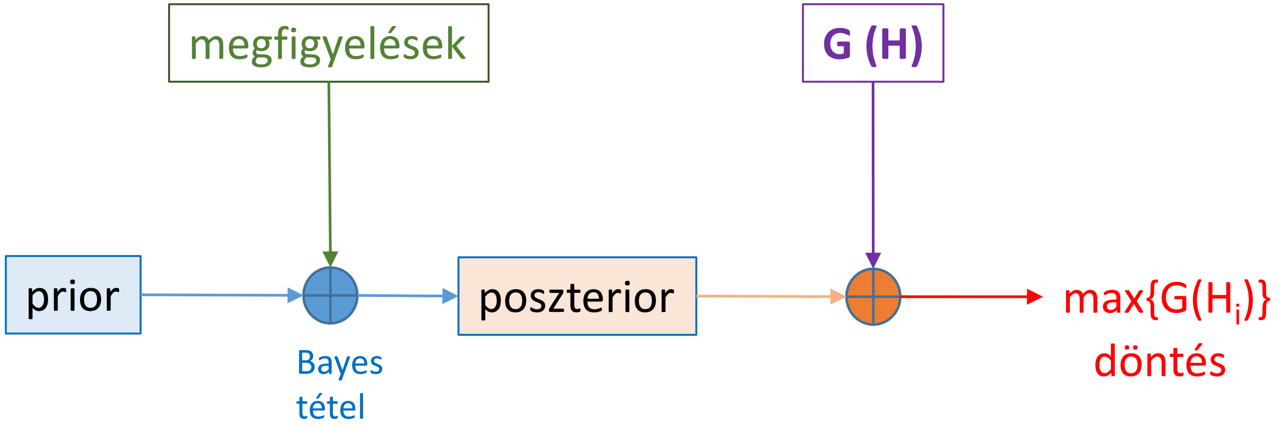
Nem biztos, hogy minden lehetőséget ismerünk, de a preferenciáinkat *racionálisan ki tudjuk terjeszteni* amennyiben új lehetőségek (pl. új terápia) merülnek fel.

A preferenciáinkat kifejezhetjük egy **hasznossági függvénnyel (G)** ami számot rendel hozzá minden lehetőséghez. (pl az adott kezelés várható haszna)

Megjegyzés: Ha a számok közötti különbséggel a preferenciáink közötti „erősség" eltérést **is** meg akarjuk adni, akkor kell egy standard, ez többnyire egy szerencsejáték (pl. lottó, sorsjegy, stb). Ekkor megadhatjuk hogy mennyiért „éri meg" nekünk „megvenni" egy játékot aminek a kimenetét értékelni akarjuk egy bázis-opcióhoz képest. (ez a *vNM várható nyereség/haszon elmélet* alapja)

-> bővebben a választható kurzusban!

A Bayes-i döntési fa



-> bővebben a választható kurzusban!