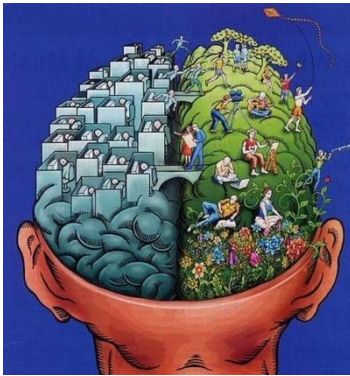




Budapesti Műszaki és Gazdaságtudományi Egyetem

Mérésstechnika és Információs rendszerek Tanszék



Bizonytalan tudás kezelése

Előadó:

Dr. Hullám Gábor

Előadás anyaga:

Dr. Dobrowiecki Tadeusz

Dr. Gézsi András

Valószínűségi axiómák

Témakörök



Bizonytalan tudás

Lehetséges okok

- „Lazaság/lustaság” - a részletes kapcsolatok megfogalmazása túl nehéz, a használatuk szintén nehézkes (véges erőforrások)
- **Elméleti ismeret hiánya** - adott problématerületnek az elméleti feltárása még nem zárult le, vagy lezárni soha nem lehet
- **Gyakorlati ismeret hiánya** - nem minden, a szabályokban hivatkozott feltétel ismert a szabályok alkalmazásakor



Bizonytalan tudás

•Példa:

$$\forall p. Tünet(p, Fogfájás) \rightarrow Van(p, Fogszuvasodás)$$

$$\forall p. Tünet(p, Fogfájás) \rightarrow Van(p, Fogszuvasodás) \vee \\ Van(p, Ínysorvadás) \vee Van(p, Bölcsességfognő) \vee \dots$$

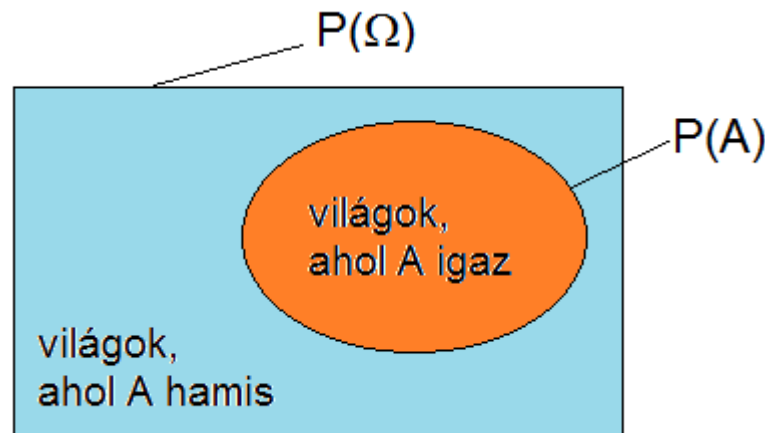
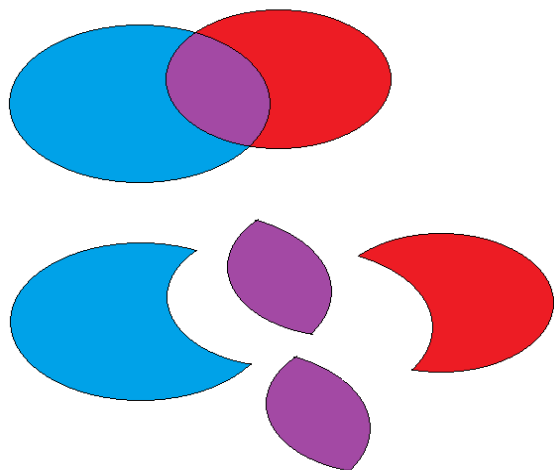
$$\forall p. Van(p, Fogszuvasodás) \rightarrow Tünet(p, Fogfájás)$$

- Fogfájás és fogszuvasodás közötti kapcsolat **egyik irányban sem feltétlen logikai következmény.**
- Az ágens tudása legjobb esetben is csak egy bizonyos **mértékű hiedelmet** jelent az adott állítással kapcsolatban.



Valószínűségi axiómák

1. Minden valószínűség 0 és 1 közé esik $0 \leq P(A) \leq 1$.
2. A biztosan igaz állítás valószínűsége 1, a biztosan hamis állításé 0. $P(\text{Igaz}) = 1$ $P(\text{Hamis}) = 0$
3. Diszjunkció valószínűsége: $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$



Valószínűségi axiómákból bizonyítottuk pl., hogy:

$$P(\neg A) = 1 - P(A)$$

$$P(A) = P(A \wedge B) + P(A \wedge \neg B) \quad \text{stb.}$$



Valószínűségi állítások

bináris	Lyuk = Igaz	$P(\text{Lyuk} = \text{Igaz})$	$P(\text{Lyuk})$
többértékű (kategorikus)	Időjárás 1 értéket vesz fel az alábbi 4 lehetségesből {Napos, Esős, Felhős, Havazás}	$P(\text{Időjárás} = \text{Esős})$	
folytonos változó	Hőmérséklet = 22.1 °C, Hőmérséklet < 22 °C	$P(\text{Hőmérséklet} < 22 \text{ °C})$	

Feltételes valószínűség: $P(A | B) = P(A \cap B) / P(B)$



Valószínűségi állítások átalakítása

Láncszabály: $P(A B C D E) = P(A | B C D E) P(B C D E) =$
 $P(A | B C D E) P(B | C D E) P(C D E) = \dots$
 $P(A | B C D E) P(B | C D E) P(C | D E) P(D | E) P(E)$

$$P(X_1 X_2 X_3 \dots X_N) = P(X_1 X_2 X_3 \dots X_{N-1}) P(X_N | X_1 X_2 \dots X_{N-1})$$

$$= P(X_1 X_2 X_3 \dots X_{N-2}) P(X_{N-1} | X_1 X_2 \dots X_{N-2}) P(X_N | X_1 X_2 \dots X_{N-1})$$

$$= \dots = \prod_{i=1}^N P(X_i | X_1 X_2 \dots X_{i-1})$$



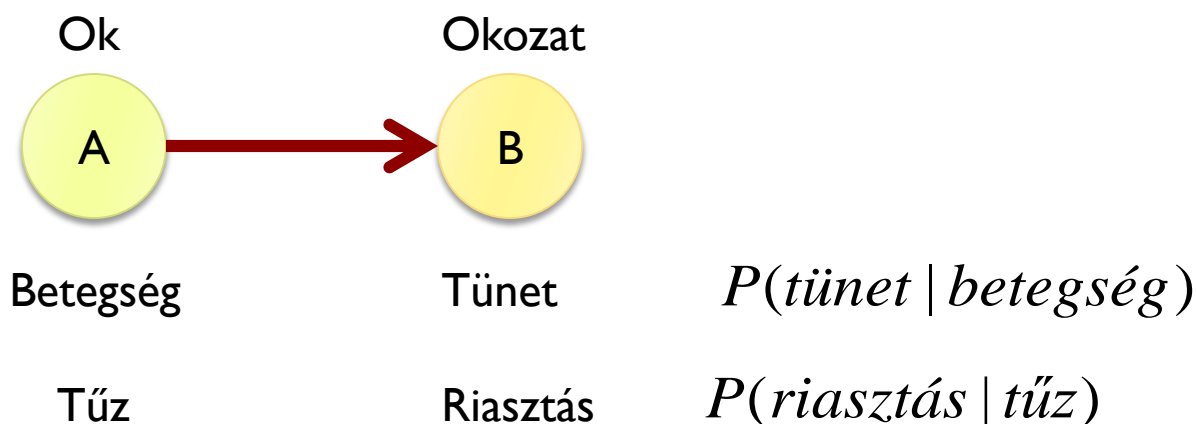
Bayes-tétel

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

$$P(A_k | B) = \frac{P(B | A_k)P(A_k)}{\sum_{i=1}^N P(B | A_i)P(A_i)}, \quad \bigcup_{i=1}^N A_i = \Omega, \quad A_i \cap A_k = \emptyset$$

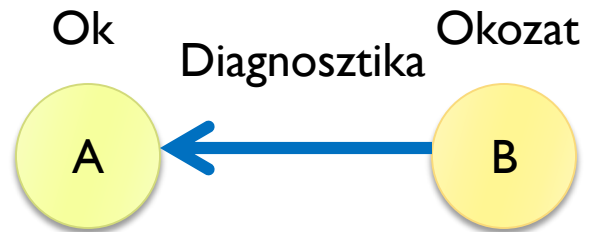
Miért fontos a Bayes-tétel?

Sokszor rendelkezünk kauzális (ok-okozati) tudással:



Bayes-tétel

- Viszont sokszor „ellentétes irányban” szeretnék következtetni, tehát evidenciák (következmények) alapján az ok(ok)ra



$P(\text{betegség} \mid \text{tünet})$

Betegség

Tünet

$P(\text{tűz} \mid \text{riasztás})$

Tűz

Riasztás



Bayes-tétel jelentősége

Lehetővé teszi a valószínűségi állítások átalakítását

- Így kiszámíthatóvá válnak nehezen becsülhető mennyiségek
- **Oksági irány:** általában könnyebb becsülni
- **Diagnosztikai irány:** általában nehezebb

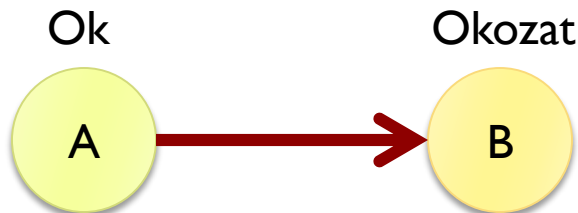
Mi a Tünet létrejöttének valószínűsége a Betegség ismeretében?

Mi a Betegség kialakulásának a valószínűsége?

$$P(\text{Betegség} | \text{Tünet}) = \frac{P(\text{Tünet} | \text{Betegség})P(\text{Betegség})}{P(\text{Tünet})}$$

Mi a Betegség meglétének valószínűsége a Tünet ismeretében?

Mi a Tünet létrejöttének a valószínűsége?



► Betegség

Tünet

Bayes-tétel jelentősége

Mi a Tünet létrejöttének valószínűsége a Betegség ismeretében?

Mi a Betegség kialakulásának a valószínűsége?

$$P(\text{Betegség} | \text{Tünet}) = \frac{P(\text{Tünet} | \text{Betegség})P(\text{Betegség})}{P(\text{Tünet})}$$

Mi a Betegség meglétének valószínűsége a Tünet ismeretében?

Mi a Tünet létrejöttének a valószínűsége?

a posteriori valószínűség (poszterior)

Likelihood

a priori valószínűség (prior)

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

Normalizációs konstans



Bayes-tétel - műveletek

Összetett feltétel

A értéke: $\{v_1, v_2, \dots, v_N\}$

$$P(A = v_k) = P_k$$

$$P(A = v_k) \wedge P(A = v_j) = 0, \quad k \neq j$$

$$P(A = v_1) \vee P(A = v_2) \dots \vee P(A = v_N) = 1$$

$$P(A = v_1) \vee P(A = v_2) \dots \vee P(A = v_k) = \sum_{i=1}^k P_i$$

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)} \leftrightarrow P(A | B \wedge X) = \frac{P(B | A \wedge X)P(A | X)}{P(B | X)}$$



Bayes-tétel - műveletek

Bővítés:

A értéke: $\{v_1, v_2, \dots, v_N\}$

$$P(A = v_k) = P_k$$

$$P(A = v_k) \wedge P(A = v_j) = 0, \quad k \neq j$$

$$P(A = v_1) \vee P(A = v_2) \dots \vee P(A = v_N) = 1$$

$$P(A = v_1) \vee P(A = v_2) \dots \vee P(A = v_k) = \sum_{i=1}^k P_i$$

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)} = \frac{P(B | A)P(A)}{\sum_{i=1}^N P(B | A_i)P(A_i)}$$

Pl. ha A egy bináris valószínűségi változó:

$$P(A = 1 | B = 1) = \frac{P(B = 1 | A = 1)P(A = 1)}{P(B = 1 | A = 0)P(A = 0) + P(B = 1 | A = 1)P(A = 1)}$$



Valószínűség értelmezése

Frekventista vs. bayesi megközelítés

frekventista nézőpont: a valószínűség objektív, események gyakoriságából számítható

Valószínűség (frekventista definíciója): Egy adott **A** esemény valószínűsége az a számérték, amely körül az esemény relatív gyakorisága (f_A) ingadozik, ha egyre több kísérletet végzünk.

$$P(A) = \lim_{n \rightarrow \infty} f_A$$

$P(A)$ tehát azt mutatja meg, hogy az **A** esemény az összes kísérlet mekkora hányadában (%) következik be.



Valószínűség értelmezése

Frekventista vs. bayesi megközelítés

bayesi nézőpont: a valószínűség egy eseménybe vetett hiedelem mértéke.

Valószínűség (bayesi definíciója): Egy adott **A** esemény valószínűsége az adott esemény bekövetkezési esélye.

Ez a gyakorlatban azt a hiedelmet fejezi ki, hogy a pillanatnyi tudásunk birtokában, a jelenlegi helyzettől megkülönböztethetetlen esetek mekkora hányadában fog bekövetkezni az adott esemény.

- Prior valószínűségekből (hiedelmekből) indulunk ki és az új evidencia érkezésekor azokat frissítjük (posterior valószínűségekké)

$$P(A) \rightarrow P(A) \rightarrow \dots \rightarrow P(A) \xrightarrow{\substack{\uparrow \\ B}} \text{Bayes-tétel} \rightarrow P(A | B)$$



A bayesi következtetés, mint a megismerés általános modellje - gondolat kísérlet

Chile országa 15 régióra van felosztva. Az ország teljes mérete 756 096 km². **Mekkora lehet az Atacama régió?** Legyen az ...

- ▶ A_1 esemény az, hogy Atacama régió kisebb, mint 10 000 km²,
 - ▶ A_2 esemény az, hogy 10 000 és 50 000 km² között van a mérete,
 - ▶ A_3 esemény az, hogy 50 000 és 100 000 km² között van,
 - ▶ A_4 esemény az, hogy nagyobb, mint 100 000 km².
-
- ▶ Rendeljen valószínűségeket az A_1, \dots, A_4 eseményekhez!



Chile...

- ▶ az A_1 esemény az, hogy Atacama régió kisebb, mint 10 000 km²,
- ▶ az A_2 esemény az, hogy 10 000 és 50 000 km² között van a mérete,
- ▶ az A_3 esemény az, hogy 50 000 és 100 000 km² között van,
- ▶ az A_4 esemény az, hogy nagyobb, mint 100 000 km².

- ▶ Atacama a negyedik legnagyobb régió.
Hogyan módosítaná a valószínűségeket?



Chile...

- ▶ az A_1 esemény az, hogy Atacama régió kisebb, mint $10\,000\text{ km}^2$,
 - ▶ az A_2 esemény az, hogy $10\,000$ és $50\,000\text{ km}^2$ között van a mérete,
 - ▶ az A_3 esemény az, hogy $50\,000$ és $100\,000\text{ km}^2$ között van,
 - ▶ az A_4 esemény az, hogy nagyobb, mint $100\,000\text{ km}^2$.
 - ▶ Atacama a negyedik legnagyobb régió.
-
- ▶ A legkisebb régió a központi régió, *Santiago Metropolitan*, amelynek mérete $15\,403\text{ km}^2$.
Hogyan módosítanád a valószínűségeket?



Chile...

- ▶ az A_1 esemény az, hogy Atacama régió kisebb, mint 10 000 km²,
- ▶ az A_2 esemény az, hogy 10 000 és 50 000 km² között van a mérete,
- ▶ az A_3 esemény az, hogy 50 000 és 100 000 km² között van,
- ▶ az A_4 esemény az, hogy nagyobb, mint 100 000 km².
- ▶ Atacama a negyedik legnagyobb régió.
- ▶ A legkisebb régió a központi régió, Santiago Metropolitan, amelynek mérete 15 403 km².

- ▶ A harmadik legnagyobb régió *Aysén del General Carlos Ibanez del Campo*, amely 108 494 km².
- ▶ Hogyan módosítaná a valószínűségeket?



Chile...

- ▶ az A_1 esemény az, hogy Atacama régió kisebb, mint $10\,000\text{ km}^2$,
 - ▶ az A_2 esemény az, hogy $10\,000$ és $50\,000\text{ km}^2$ között van a mérete,
 - ▶ az A_3 esemény az, hogy $50\,000$ és $100\,000\text{ km}^2$ között van,
 - ▶ az A_4 esemény az, hogy nagyobb, mint $100\,000\text{ km}^2$.
 - ▶ Atacama a negyedik legnagyobb régió.
 - ▶ A legkisebb régió a központi régió, Santiago Metropolitan, amelynek mérete $15\,403\text{ km}^2$.
- A harmadik legnagyobb régió Aysén del General Carlos Ibanez del Campo, amely $108\,494\text{ km}^2$.



Atacama mérete: $75\,176\text{ km}^2$



Hiedelmek és értelmezésük

Egy adott kijelentéshez rendelt (szubjektív) **0 / 1 valószínűség**:
határozott hiedelem, hogy az állítás **hamis / igaz**.

A **0 és 1 közötti** valószínűség a mondat **igazságtartalmában való hiedelem**
mértékeinek felel meg. **Az állítás valójában persze vagy igaz vagy hamis.**

A 0.8 (szubjektív) valószínűség nem jelenti, hogy az állítás a „80 %-ban igaz”, hanem egy 80 %-os mértékű hiedelmet, azaz igen erős elvárást az állítás igazságával szemben.

Valószínűségi kijelentés szemantikája: egy **kijelentéshez rendelt valószínűség az addigi észlelésektől** (tény, evidencia, tényállás) **függ**.

Ágens kihúz egy lapot egy megkevert kártyapakliból. Mielőtt ránézne a lapra: $P(\text{„a lap pikk ász (lesz)”}) = 1/52$. Miután megnézte: $P(\dots) = 0$, vagy 1.

Orvos páciens megvizsgálása után P_1 valószínűségűnek lát egy konkrét betegséget. Labor lelet ezt vagy erősíti ($> P_1$), vagy gyengíti ($< P_1$).



Események – véletlen kísérlet

- ▶ **Elemi esemény:** minden lehetséges kimenetel (e_1, e_2, \dots, e_n) , amiről egy kísérlet elvégzése után eldönthető, hogy bekövetkezett vagy sem
 - ▶ $e_1 \wedge e_2 \wedge \dots \wedge e_n = \emptyset$
- ▶ **Eseménytér:** egy kísérlet összes kimenetele, az összes elemi esemény halmaza (Ω) .
 - ▶ $e_1 \cup e_2 \cup \dots \cup e_n = \Omega$
- ▶ **Véletlen esemény:** az eseménytér egy részhalmaza
- ▶ **Biztos esemény:** a kísérlet során biztosan (minden kimenetelnél) bekövetkezik.
- ▶ **Ellentett esemény:** akkor és csak akkor következik be, ha az eredeti esemény nem következik be

Együttes valószínűség-eloszlás

- Az együttes valószínűség-eloszlás $P(X_1, \dots, X_n)$ minden egyes elemi eseményhez valószínűséget rendel.
- Ha minden vizsgált valószínűségi változó diszkrét, akkor az együttes valószínűség-eloszlás leírható egy n -dimenziós táblázattal
- Egy cella = az adott állapot valószínűsége.

Írjuk le a budapesti időjárást két (egy bináris és egy ternáris) változóval:

időjárás: {napos, felhős},

hőmérséklet: {meleg, közepes, hideg}.

Akkor az együttes eloszlás $P(\textit{időjárás}, \textit{hőmérséklet})$:



Együttes valószínűség-eloszlás

Írjuk le a budapesti időjárást két (egy bináris és egy ternáris) változóval:

időjárás: {napos, felhős},

hőmérséklet: {meleg, közepes, hideg}.

Akkor az együttes eloszlás $P(\textit{időjárás}, \textit{hőmérséklet})$:

	Meleg	Közepes	Hideg
Napos	.1	.2	.1
Felhős	.05	.35	.2

- Mivel az elemi események egymást kizáróak, ezek együttes bekövetkezése szükségszerűen hamis tény.
- Az axiómákból következően: **a táblázat elemeinek összege 1.**



Marginális és más eloszlások

$$P(X, Y) = \sum_{z \in \text{dom}(Z)} P(X, Y, Z = z)$$

$P(\text{hőmérséklet}) =$	Meleg	Közepes	Hideg
	.15	.55	.3

$P(\text{időjárás}) =$	Napos	Felhős
	.4	.6

	Meleg	Közepes	Hideg
Napos	.1	.2	.1
Felhős	.05	.35	.2



Marginális és más eloszlások

$P(\text{hőmérséklet} | \text{időjárás} = \text{Napos}) =$

	Meleg	Közepes	Hideg
	.25	.50	.25

$P(\text{időjárás} | \text{hőmérséklet}) =$

	Meleg	Közepes	Hideg
Napos	.67	.36	.33
Felhős	.33	.64	.67

	Meleg	Közepes	Hideg
Napos	.1	.2	.1
Felhős	.05	.35	.2



Együttes valószínűség-eloszlás

Jó hír: együttes eloszlás birtokában minden kérdésre kapunk választ, ami a benne szereplő véletlen változók viszonyára és tulajdonságaira vonatkozik.

Rossz hír: nemigen megy 10-nél több változót tartalmazó eloszlások megadása

$P(x_1, x_2, \dots, x_N)$ esetén kell $2^N - 1$ független valószínűségérték.

A diagnózishoz **exponenciális számú valószínűség ismerete szükséges.**

Rémálom: ha valamelyik valószínűség értéke megváltozik?



A bayesi frissítés

- Egyesével gyűjtjük a tényeket, majd módosítjuk az ismeretlen változóval kapcsolatos korábbi hiedelmi mértéket.

1. Fogfájás:

$$P(\text{Fogsz} | \text{Ffáj}) = P(\text{Fogsz}) \frac{P(\text{Ffáj} | \text{Fogsz})}{P(\text{Ffáj})}$$

2. **Lyuk:** a Bayes-tételt úgy alkalmazzuk, hogy a továbbiakban a *Fogfájás*-t állandó feltételnek tekintjük:

$$P(\text{Fogsz} | \text{Lyuk}) = P(\text{Fogsz}) \frac{P(\text{Lyuk} | \text{Fogsz})}{P(\text{Lyuk})}$$

$$P(\text{Fogsz} | \text{Ffáj} \wedge \text{Lyuk}) = P(\text{Fogsz} | \text{Ffáj}) \frac{P(\text{Lyuk} | \text{Ffáj} \wedge \text{Fogsz})}{P(\text{Lyuk} | \text{Ffáj})}$$



A bayesi frissítés -példa

$$P(\text{Fogsz} | \text{Ffáj} \wedge \text{Lyuk}) = P(\text{Fogsz} | \text{Ffáj}) \frac{P(\text{Lyuk} | \text{Ffáj} \wedge \text{Fogsz})}{P(\text{Lyuk} | \text{Ffáj})}$$

$$P(\text{Fogsz} | \text{Ffáj}) = \overbrace{P(\text{Fogsz}) \frac{P(\text{Ffáj} | \text{Fogsz})}{P(\text{Ffáj})}}$$

$$\begin{aligned} &P(\text{Fogsz} | \text{Ffáj} \wedge \text{Lyuk}) \\ &= P(\text{Fogsz}) \frac{P(\text{Ffáj} | \text{Fogsz})}{P(\text{Ffáj})} \frac{P(\text{Lyuk} | \text{Ffáj} \wedge \text{Fogsz})}{P(\text{Lyuk} | \text{Ffáj})} \end{aligned}$$



Feltételes függetlenség

Mind a fogfájásnak, mind a szonda lyukba akadásának közvetlen oka a fogszuvasodás.

Amint **tudjuk**, hogy fogszuvasodás, nem hisszük, hogy a szonda lyukba akadásának valószínűsége a fogfájástól fog függeni.

Hasonlóképpen, a szonda találata nem befolyásolja annak valószínűségét, hogy a szuvasodás fogfájást okoz.

$$P(\text{Lyuk} \mid \text{Ffáj} \wedge \text{Fogsz}) = P(\text{Lyuk} \mid \text{Fogsz})$$

$$P(\text{Ffáj} \mid \text{Lyuk} \wedge \text{Fogsz}) = P(\text{Ffáj} \mid \text{Fogsz})$$

a *Fogszuvasodás* ténye esetén a *Fogfájás* és *Lyuk* között fennáll a **feltételes függetlenség**.



Feltételes függetlenség - alkalmazás

$$P(\text{Fogsz} \mid \text{Ffáj} \wedge \text{Lyuk})$$

$$= P(\text{Fogsz}) \frac{P(\text{Ffáj} \mid \text{Fogsz})}{P(\text{Ffáj})} \frac{P(\text{Lyuk} \mid \text{Ffáj} \wedge \text{Fogsz})}{P(\text{Lyuk} \mid \text{Ffáj})}$$

$$P(\text{Lyuk} \mid \text{Ffáj} \wedge \text{Fogsz}) = P(\text{Lyuk} \mid \text{Fogsz})$$

$$P(\text{Ffáj} \mid \text{Lyuk} \wedge \text{Fogsz}) = P(\text{Ffáj} \mid \text{Fogsz})$$

$$P(\text{Fogsz} \mid \text{Ffáj} \wedge \text{Lyuk})$$

$$= P(\text{Fogsz}) \frac{P(\text{Ffáj} \mid \text{Fogsz})}{P(\text{Ffáj})} \frac{P(\text{Lyuk} \mid \text{Fogsz})}{P(\text{Lyuk} \mid \text{Ffáj})}$$



Normalizálás

Még mindig kérdéses a: $P(\text{Lyuk}|\text{Fogfájás})?!$

- várhatóan figyelembe kell venni a tünetek összes lehetséges párosítását (hármait, stb.), valóságban ez a kifejezés kiesik:

a nevezők szorzata: $P(\text{Lyuk}|\text{Fogfájás}) P(\text{Fogfájás}) = P(\text{Lyuk})$

Ezt az un. **normalizálással** kiküszöbölhetjük, feltéve, hogy pl. a $P(\text{Lyuk}|\neg\text{Fogszuvasodás})$ -t megbecsüljük.

$$P(\text{Fogsz} | \text{Lyuk}) = \frac{P(\text{Lyuk} | \text{Fogsz})P(\text{Fogsz})}{P(\text{Lyuk})}$$

$$P(\text{Fogsz} | \text{Lyuk}) \approx P(\text{Lyuk} | \text{Fogsz})P(\text{Fogsz})$$



Normalizálás

$$P(\text{Fogsz} | \text{Lyuk}) = \frac{P(\text{Lyuk} | \text{Fogsz})P(\text{Fogsz})}{P(\text{Lyuk})}$$

$$P(\text{Fogsz} | \text{Lyuk}) \approx P(\text{Lyuk} | \text{Fogsz})P(\text{Fogsz})$$

$$P(\text{Fogsz} | \text{Lyuk}) = \alpha P(\text{Lyuk} | \text{Fogsz})P(\text{Fogsz}) = C_1 \alpha$$

$$P(\neg \text{Fogsz} | \text{Lyuk}) = \alpha P(\text{Lyuk} | \neg \text{Fogsz})P(\neg \text{Fogsz}) = C_2 \alpha$$

$$P(\text{Fogsz} | \text{Lyuk}) + P(\neg \text{Fogsz} | \text{Lyuk}) = (C_1 + C_2) \alpha = 1$$

$$\alpha = (C_1 + C_2)^{-1}$$



Lehetséges problémák

- ▶ Az a priori feltételes és együttes valószínűségek begyűjtése **nehéz** és **költséges**
- ▶ Az emberek rossz valószínűségbecslők (szubjektív megközelítés esetén)
- ▶ A Bayes-szabály **sok számítást** igényel

