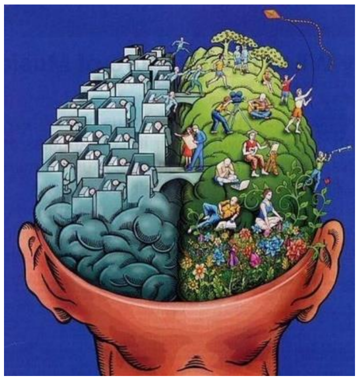
Budapesti Műszaki és Gazdaságtudományi Egyetem
Méréstechnika és Információs rendszerek Tanszék



Bizonytalan tudás kezelése

Előadó:	Hullám Gábor
	Pataki Béla
Előadás anyaga:	Dobrowiecki Tadeusz

Valószínűségi axiómák

MŰEGYETEM 1782

BME I.E. 414, 463-26-79
pataki@mit.bme.hu,
http://www.mit.bme.hu/general/staff/pataki

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:

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

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

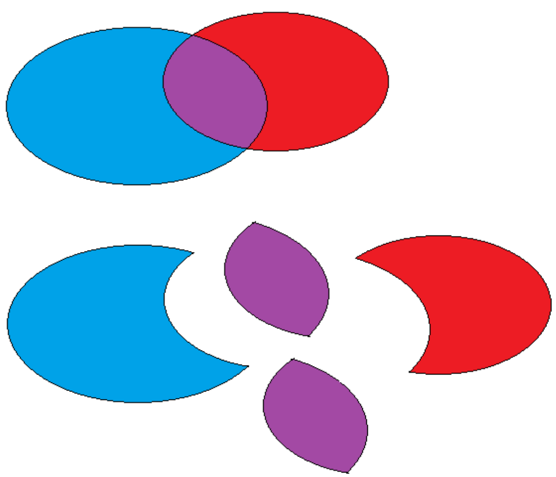
$$\forall p.\,Van(p, Fogszuvasodás) \rightarrow Tünet(p, Fogfájás)$$

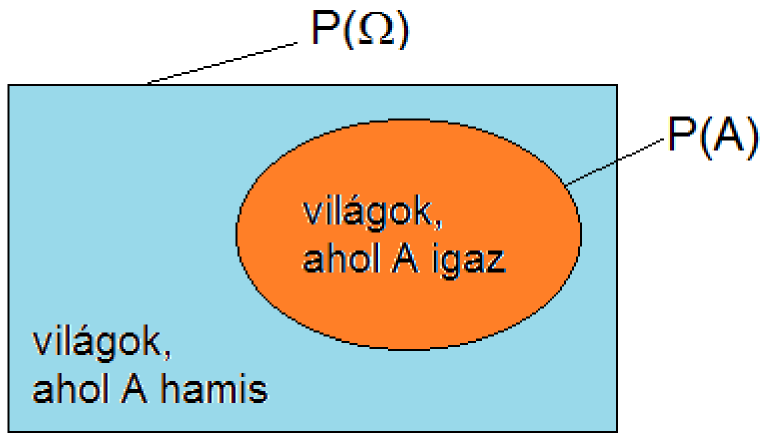
- Fogfájás és fogszuvasodás közötti kapcsolat egyik irányban sem feltétlen logikai következmény.

- 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. $P(Igaz) = 1$ $P(Hamis) = 0$
3. Diszjunkció valószínűsége: $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$





Valószínűségi axiómákból bizonyítottuk pl., hogy:

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

$P(A) = P(A \wedge B) + P(A \wedge \neg B)$ stb.

Volt már

Valószínűségi állítások

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

Feltételes valószínűség: P(A | B) = P(A B) / P(B)

Volt már

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) = …
P(A | B C D E) P(B | C D E) P(C | D E) P(D | E) P(E)

$$P(X_1 X_2 X_3 \ldots X_N) = P(X_1 X_2 X_3 \ldots X_{N-1})\, P(X_N \mid X_1 X_2 \ldots X_{N-1})$$

$$= P(X_1 X_2 X_3 \ldots X_{N-2})\, P(X_{N-1} \mid X_1 X_2 \ldots X_{N-2})\, P(X_N \mid X_1 X_2 \ldots X_{N-1})$$

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

Volt már

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:

Ok Okozat

Betegség Tünet $P(tünet \mid betegség)$

Tűz Riasztás $P(riasztás \mid tűz)$

Bayes-tétel

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

$P(betegség \mid tünet)$ Betegség Tünet

$P(tűz \mid riasztás)$ Tűz Riasztás

Bayes-tétel jelentőssége

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?

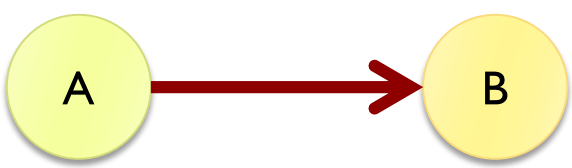
$$P(Betegség \mid Tünet) = \frac{P(Tünet \mid Betegség)P(Betegség)}{P(Tünet)}$$

Mi a Betegség meglétének valószínűsége a Tünet ismeretében?

Mi a Tünet létrejöttének a valószínűsége?

Ok

Okozat



Betegség

Tünet

Bayes-tétel jelentősége

Mi a Tünet létrejöttének valószínűsége a Betegség ismeretében?

Mi a Betegség kialakulásának a valószínűsége?

$$P(Betegség \mid Tünet) = \frac{P(Tünet \mid Betegség)P(Betegség)}{P(Tünet)}$$

Mi a Betegség meglétének valószínűsége a Tünet ismeretében?

Mi a Tünet létrejöttének a valószínűsége?

a posteriori valószínűség (poszterior)

Likelihood

a priori valószínűség (prior)

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

Normalizációs konstans

Bayes-tétel - műveletek

Összetett feltétel

A értéke: $\{v_1, v_2, \ldots v_N\}$

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

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

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

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

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

Bayes-tétel - műveletek

Bővítés:

A értéke: $\{v_1, v_2, \ldots v_N\}$

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

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

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

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

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

Pl. ha A egy bináris valószínűségi változó:

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

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)

$$P(A) \rightarrow P(A) \rightarrow \ldots \rightarrow P(A) \rightarrow \text{Bayes-tétel} \rightarrow P(A \mid B)$$

B

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(„a lap pikk ász (lesz)”) = 1/52. Miután megnézte: P(...) = 0, vagy 1.

Orvos páciens megvizsgálása után P1 valószínűségűnek lát egy konkrét betegséget. Labor lelet ezt vagy erősíti (> P1), vagy gyengíti (< P1).

Események – véletlen kísérlet

- **Elemi esemény:** minden lehetséges kimenetel ($e_1, e_2, \ldots, e_n$), amiről egy kísérlet elvégzése után eldönthető, hogy bekövetkezett vagy sem
  - $e_1 \wedge e_2 \wedge \ldots \wedge e_n = \emptyset$
- **Eseménytér:** egy kísérlet összes kimenetele, az összes elemi esemény halmaza (Ω).
  - $e_1 \, U \, e_2 \, U \ldots U \, e_n = \Omega$
- **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

Volt már

Együttes valószínűség-eloszlás

- Az együttes valószínűség-eloszlás P(X_1, ... ,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(*időjárás*, *hőmérséklet*):

Volt már

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(*időjárás*, *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.**

Volt már

Marginális és más eloszlások

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

P(*hőmérséklet*) =

Meleg	Közepes	Hideg
.15	.55	.3

P(*időjárás*) =

Napos	Felhős
.4	.6

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

Volt már

Marginális és más eloszlások

P(*hőmérséklet*| *időjárás* = Napos) =

Meleg	Közepes	Hideg
.25	.50	.25

P(*időjárás*| *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

$P(x_1, x_2, \ldots, x_N)$ 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:

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

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

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

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

A bayesi frissítés -példa

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

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

$$P(Fogsz \mid Ffáj \wedge Lyuk)$$

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

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.

$$P(Lyuk \mid Ffáj \wedge Fogsz) = P(Lyuk \mid Fogsz)$$
$$P(Ffáj \mid Lyuk \wedge Fogsz) = P(Ffáj \mid Fogsz)$$

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

Feltételes függetlenség - alkalmazás

$$P(Fogsz \mid Ffáj \wedge Lyuk)$$

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

$$P(Lyuk \mid Ffáj \wedge Fogsz) = P(Lyuk \mid Fogsz)$$

$$P(Ffáj \mid Lyuk \wedge Fogsz) = P(Ffáj \mid Fogsz)$$

$$P(Fogsz \mid Ffáj \wedge Lyuk)$$

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

Normalizálás

Még mindig kérdéses a: $P(Lyuk|Fogfájás)$?!

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

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

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

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

$$P(Fogsz \mid Lyuk) \approx P(Lyuk \mid Fogsz)P(Fogsz)$$

Normalizálás

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

$$P(Fogsz \mid Lyuk) \approx P(Lyuk \mid Fogsz)P(Fogsz)$$

$$P(Fogsz \mid Lyuk) = \alpha P(Lyuk \mid Fogsz)P(Fogsz) = C_1\alpha$$

$$P(\neg Fogsz \mid Lyuk) = \alpha P(Lyuk \mid \neg Fogsz)P(\neg Fogsz) = C_2\alpha$$

$$P(Fogsz \mid Lyuk) + P(\neg Fogsz \mid Lyuk) = (C_1 + C_2)\alpha = 1$$

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

$$\boxed{\begin{aligned} &P(Ok \mid Hatás_1, Hatás_2, \ldots Hatás_N) = \\ &\quad \alpha P(Ok)P(Ok \mid Hatás_1)P(Ok \mid Hatás_2)\ldots P(Ok \mid Hatás_N) \end{aligned}}$$

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