přednáška 03: další pravidla pro výpočet psti

Jeden princip součtu je třetím axiomem v definici psti, a sice pro neslučitelné jevy:

Axiom 3: $P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \ldots$ pro navzájem neslučitelné jevy

(axiom součtu pravděpodobností konečně mnoha nebo nekonečně mnoha neslučitelných jevů)

- Musíme se ovšem zabývat ještě situací, kdy jevy A_i NEJSOU neslučitelné, tj. nějaké elementární výsledky experimentu leží v jejich průniku

Začněme obecnou situací dvou jevů:

Větička 3: $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$ pro libov. $A_1, A_2 \in \mathcal{A}$.

Důkaz: Pokud existují nějaké prvky v průniku obou množin, lze sjednocení $A_1 \cup A_2$ rozložit na tři disjunktní podmnožiny: $A_1 - A_2$, $A_2 - A_1$ a $A_1 \cap A_2$.

- Pak podle axiomu 3 platí:

$P(A_1 \cup A_2) = P(A_1 - A_2) + P(A_1 \cap A_2) + P(A_2 - A_1)$ (ax. 3 pro tři disjmnožiny)

$P(A_1) = P(A_1 - A_2) + P(A_1 \cap A_2)$ (ax. 3 pro dvě disjunktní množiny)

$P(A_2) = P(A_1 \cap A_2) + P(A_2 - A_1)$ (ax. 3 pro dvě disjunktní množiny

Dosazením těchto tří vztahů do levé a právě strany rovnice ve větičce 3 vidíme:

$L = P(A_1 \cup A_2) = P(A_1 - A_2) + P(A_1 \cap A_2) + P(A_2 - A_1)$

$P = P(A_1) + P(A_2) - P(A_1 \cap A_2) = P(A_1 - A_2) + P(A_1 \cap A_2) + P(A_1 \cap A_2) + \quad + P(A_2 - A_1) - P(A_1 \cap A_2) = P(A_1 - A_2) + P(A_1 \cap A_2) + P(A_2 - A_1)$

Obě strany se rovnají, tvrzení platí.

Lze dozákat větu o psti sjednocení n jevů (otázka č. 11):

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n) -$$
$$- P(A_1 \cap A_2) - P(A_1 \cap A_3) - \dots - P(A_{n-1} \cap A_n) +$$
$$+ P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) + \dots P(A_{n-2} \cap A_{n-1} \cap A_n) -$$
$$\dots$$
$$+ (-1)^{n-1} P(A_1 \cap A_2 \cap A_3 \dots \cap A_n)$$

pro libov. náhodné jevy $A_1, A_2, \dots A_n \in \mathcal{A}$.

(této větě se říká věta o součtu pstí)

Příklad: Elektrikář vytiskl štítky na zvonky domy se 4 byty a zapojuje je náhodně, protože jeho parťák onemocněl

Jaká je pst, že

a) Všechny zvonky zapojí ke správným bytům

b) Aspoň jeden zvonek zapojí správně?

c) Žádný zvonek nezapojí správně?

- A_1 … zvonek k bytu č. 1 bude zapojen správně
  A_2 … zvonek k bytu č. 2 bude zapojen správně
  A_3 … zvonek k bytu č. 3 bude zapojen správně
  A_4 … zvonek k bytu č. 4 bude zapojen správně

Řešení: všech možných připojení zvonků k bytům je …

(řešíme pomocí modelu klasické psti … každá varianta připojení zvonků je stejně pravděpodobná)

- a) $P(A_1 \cap A_2 \cap A_3 \cap A_4) = \frac{1}{} =$

Řešení: všech možných připojení zvonků k bytům je 24

a) $P(A_1 \cap A_2 \cap A_3 \cap A_4) = \frac{1}{24} = 0{,}04166666\ldots$

b) $P(A_1 \cup A_2 \cup A_3 \cup A_4) = P(A_1) + P(A_2) + P(A_3) + P(A_4) -$

- $- P(A_1 \cap A_2) - P(A_1 \cap A_3) - \ldots - P(A_3 \cap A_4) +$
  $+ P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) + P(A_1 \cap A_3 \cap A_4) + \; + P(A_2 \cap A_3 \cap A_4) - P(A_1 \cap A_2 \cap A_3 \cap A_4) =$

$= \frac{1}{24}(\ldots)$

a) $P(A_1 \cap A_2 \cap A_3 \cap A_4) = \frac{1}{24} = 0{,}04166666\ldots$

b) $P(A_1 \cup A_2 \cup A_3 \cup A_4) = P(A_1) + P(A_2) + P(A_3) + P(A_4) -$

$- P(A_1 \cap A_2) - P(A_1 \cap A_3) - \ldots - P(A_3 \cap A_4) +$

$+ P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) + P(A_1 \cap A_3 \cap A_4) + \; + P(A_2 \cap A_3 \cap A_4) - P(A_1 \cap A_2 \cap A_3 \cap A_4) =$

$= \frac{1}{24}(3! + 3! + 3! + 3! - 2! - 2! - 2! - 2! - 2! - 2! + 1 + 1 + 1 + 1 - 1) = 0{,}625$

Jaká je pst, že

a) Všechny zvonky zapojí ke správným bytům

b) Aspoň jeden zvonek zapojí správně?

c) Žádný zvonek nezapojí správně?

$$1 - P(A_1 \cup A_2 \cup A_3 \cup A_4) = 1 - 0{,}625 = 0{,}375$$

Otázka 12: Stochastická nezávislost jevů

stocazomai = tuším, domnívám se, odhaduji

stocastikoj = důvtipný, bystý (o lidech)

stocastikoj = odhadovaný, náhodný

Stochastické metody = odhadové metody, přibližné metody

Definice: (Králová, Budíková, Maroš, str. 59)

jevy $\mathcal{A}_1, \mathcal{A}_2 \in \mathcal{A}$ se naz. stochasticky nezávislé, když

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$$

jevy $A_1, A_2, \dots A_n \in \mathcal{A}$ se naz stochasticky nezávislé, když

(i) $P(A_i \cap A_j) = P(A_i) \cdot P(A_j)$ pro $1 \leq i < j \leq n$

(ii) $P(A_i \cap A_j \cap A_k) = P(A_i) \cdot P(A_j) \cdot P(A_k)$ pro $1 \leq i < j < k \leq n$

(iii) ...

(iv) $P(A_1 \cap A_2 \cap A_3 \dots \cap A_n) = P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_n)$

(tedy ověřujeme celkem $2^n - n - 1$ vztahů)

Příklad: Házíme dva hody mincí, padá líc nebo rub

A … v prvním hodu padne líc

B … ve druhém hodu padne líc

C … v obou hodech padne stejná strana

Ověřte, zda jsou jevy A, B, C stochasticky nezávislé

$\Omega=\{LL, LR, RL, RR\}$ ... všechny element výsl nastávají stejně často

A ... v prvním hodu padne líc

$A=\{LL, LR\}$

B ... ve druhém hodu padne rub

$B=\{LR, RR\}$

C ... v obou hodech padne stejná strana

$C=\{LL, RR\}$

Máme ověřit čtyři vztahy:

$P(A \cap B) = P(A) \cdot P(B)$ $\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$ … OK

$P(A \cap C) = P(A) \cdot P(C)$ $\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$ … OK

$P(B \cap C) = P(B) \cdot P(C)$ $\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$ … OK

$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$???????????????

$P(A \cap B) = P(A) \cdot P(B)$ $\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$ ... OK

$P(A \cap C) = P(A) \cdot P(C)$ $\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$ ... OK

$P(B \cap C) = P(B) \cdot P(C)$ $\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$ ... OK

$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ $\quad 0 \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$ ... not OK

Tj jevy A, B, C nejsou stochasticky nezávislé

(závislost je ukryta až při interakci všech tří jevů)

Další příklad: Házíme čtyřstěnem; jsou náhodné jevy R,G,B stochasticky nezávislé?

stěna A … obarvena červeně

stěna B … obarvena zeleně

stěna C … obarvena modře

stěna D … obarvena část červeně, část zeleně, část modře

- R … padne aspoň část stěny červené ….. R $=\{A,D\}$

G … padne aspoň část stěny zelené …. G $=\{B,D\}$

B … padne aspoň část stěny modré ….. B $=\{C\ D\}$

$$\Omega=\{A,B,C,D\}$$

Máme ověřit čtyři vztahy:

$P(R \cap G) = P(R) \cdot P(G)$ $\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$ ... OK

$P(R \cap B) = P(R) \cdot P(B)$ $\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$ ... OK

$P(G \cap B) = P(G) \cdot P(B)$ $\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$ ... OK

$P(R \cap G \cap B) = P(R) \cdot P(G) \cdot P(B)$???????????????

$P(R \cap G) = P(R) \cdot P(G)$ $\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$ ... OK

$P(R \cap B) = P(R) \cdot P(B)$ $\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$ ... OK

$P(G \cap B) = P(G) \cdot P(B)$ $\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$ ... OK

- $P(R \cap G \cap B) = P(R) \cdot P(G) \cdot P(B)$ $\quad \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$ ... not OK

Tj jevy A, B, C nejsou stochasticky nezávislé

(závislost je ukryta až při interakci všech tří jevů)

Otázka 13: podmíněná pst (označení i vzorec jsou vysvětleny na příkladu)

Příklad: Uvažujme situaci, kdy zásilkový prodejce vyřizuje objednávky zboží telefonicky, emailem nebo formulářem při osobním odběru

T … objednávka bude se děje telefonicky

E … objednávka se děje emailem

F … objednávka se děje formulářem při osobním odběru

M … objednávka je malá

S … objednávka je střední

V … objednávka je velká

H … objednávka je prioritní, HIGH PRIORITY

Jsou známa následující data z letošního roku:

Typ obj.	Malá obj.	Střední obj.	Velká obj.	Prior. obj	celkem
Telefon	1021	216	109	14	1360
Email	86	371	308	49	814
Formulář	1497	230	86	13	1826
celkem	2604	817	503	76	4000

Dvě otázky v tomto příkladu:

a) Právě volá zákazník, který si chce objednat zboží. S jakou pstí bude jeho objednávka prioritní?

b) Referentka řekla, že právě vyřídila prioritní objednávku. S jakou pstí ji vyřizovala se zákazníkem telefonicky?

Charakter otázek:

1. Je nám známá část informace – víme, že nastal jev V (podmínka)
2. Situace ještě stále obsahuje náhodnost – nevíme, zda nastal N

Chceme spočítat tzv. podmíněnou pst $P(N \mid V)$ ... pst výskytu jevu N, pokud je známo, že nastala situace V (obr. ... $N \subseteq V$)

Odpověď na první otázku:

Vzorec: $P(N|V)=\frac{P(N\cap V)}{P(V)}$ ... platí, pokud jev V má nenulovou pst

a) Právě volá zákazník, který si chce objednat zboží. S jakou pstí bude jeho objednávka prioritní?

Máme počítat P(H|T) ... pst, že objednávka je prioritní, pokud víme, že k ní došlo telefonicky: $P(H|T)=\frac{P(H\cap T)}{P(T)}=\frac{14}{1360}\doteq 0{,}0103$

Odpověď na druhou otázku:

Vzorec: $P(N|V)=\frac{P(N\cap V)}{P(V)}$ … platí, pokud jev V má nenulovou pst

- b) Referentka řekla, že právě vyřídila prioritní objednávku. S jakou pstí ji vyřizovala se zákazníkem telefonicky?

Máme počítat P(T|H) … pst, že objednávka je telefonická, pokud víme, že je HIGH priority: $P(T|H)=\frac{P(T\cap H)}{P(H)}=\frac{14}{76}\doteq 0{,}1842$

Otázka č. 14: věta o součinu pstí

Ze vzorce pro podmíněnou pst plyne: P(V ∩ N)=P(V)·P(N | V)

Obecně pro n jevů (věta o součinu pstí):

- Pokud P(A_1 ∩ A_2 ∩ A_3… ∩ A_{n-1})≠0 (tj. jmenovatel posledního člene v následující rovnosti je nenulový), tak platí

P(A_1 ∩ A_2 ∩ A_3… ∩ A_n)= P(A_1)· P(A_2 | A_1)·P(A_3 | A_1 ∩ A_2) · … ·

· P(A_n | A_1 ∩ A_2 ∩ A_3… ∩ A_{n-1})

Příklad: Ze sady 100 výrobků je 10 zmetků. Při kontrole jakosti vybereme náhodně tři výrobky z této sady.

a) Určete pst, že první dva výrobky budou kvalitní a třetí bude zmetek (tj. záleží na pořadí zmetku). Jen tato část a) se týká podmíněné psti, naučte se jen tu.

b) Určete pst, že ze tří vybraných výrobků budou dva kvalitní a jeden zmetek. Nezáleží na pořadí zmetku. Počítejte podle věty o součinu.

c) Totéž jako b), ale počítejte pomocí klasické psti (tak nějak najednou, když nezáleží na pořadí).

Příklad: Ze sady 100 výrobků je 10 zmetků. Při kontrole jakosti vybereme náhodně tři výrobky z této sady.

a) Určete pst, že první dva výrobky budou kvalitní a třetí bude zmetek (tj. záleží na pořadí zmetku).

- A_1 … první vybraný výrobek bude kvalitní

A_2 … druhý vybraný výrobek bude kvalitní

$\overline{A_3}$ … třetí vybraný výrobek bude zmetek

$P(A_1 \cap A_2 \cap \overline{A_3}) = P(A_1) \cdot P(A_2 \mid A_1) \cdot P(\overline{A_3} \mid A_1 \cap A_2) = \ldots$

a) Určete pst, že první dva výrobky budou kvalitní a třetí bude zmetek (tj. záleží na pořadí zmetku).

A_1 … první vybraný výrobek bude kvalitní

A_2 … druhý vybraný výrobek bude kvalitní

$\overline{A_3}$ … třetí vybraný výrobek bude zmetek

$$P(A_1 \cap A_2 \cap \overline{A_3}) = P(A_1) \cdot P(A_2 \mid A_1) \cdot P(\overline{A_3} \mid A_1 \cap A_2) = \frac{90}{100} \cdot \frac{89}{99} \cdot \frac{10}{98} \doteq 0{,}0826$$

b) Určete pst (jevu B), že ze tří vybraných výrobků budou dva kvalitní a jeden zmetek. Nezáleží na pořadí zmetku.

$P(B) = P(A_1 \cap A_2 \cap \overline{A_3}) + P(A_1 \cap \overline{A_2} \cap A_3) + P(\overline{A_1} \cap A_2 \cap A_3) = \ldots$

b) Určete pst (jevu B), že ze tří vybraných výrobků budou dva kvalitní a jeden zmetek. Nezáleží na pořadí zmetku.

$P(B) = P(A_1 \cap A_2 \cap \overline{A_3}) + P(A_1 \cap \overline{A_2} \cap A_3) + P(\overline{A_1} \cap A_2 \cap A_3) =$

$= \frac{90}{100} \cdot \frac{89}{99} \cdot \frac{10}{98} + \frac{90}{100} \cdot \frac{10}{99} \cdot \frac{89}{98} + \frac{10}{100} \cdot \frac{90}{99} \cdot \frac{89}{98} \doteq 02477$

Některé příklady lze řešit i více způsoby:

c) Totéž jako b), ale počítejte pomocí klasické psti (tak nějak najednou, když nezáleží na pořadí).

$P(B) = \frac{?}{??}$

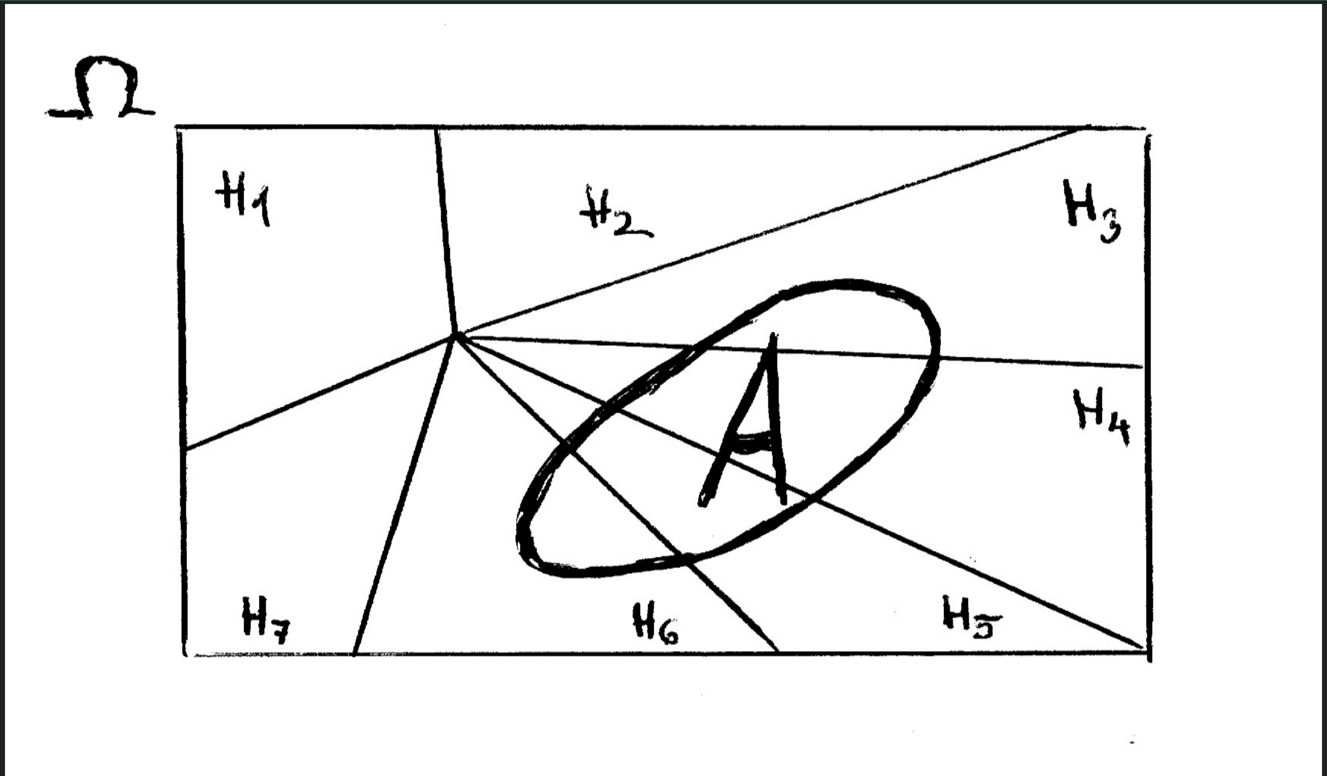
c) Totéž jako b), ale počítejte pomocí klasické psti (tak nějak najednou, když nezáleží na pořadí).

$$P(B)=\frac{|B|}{|\Omega|}=\frac{??}{???}$$

c) Totéž jako b), ale počítejte pomocí klasické psti (tak nějak najednou, když nezáleží na pořadí).

$$P(B)=\frac{|B|}{|\Omega|}=\frac{\binom{10}{1}\cdot\binom{90}{2}}{\binom{100}{3}} \doteq 02477$$ ... stejný výsledek!!

Otázka č. 15: věta o úplné psti - obrazem



Ot. č 15: Věta o úplné psti - slovem

Pokud $\{H_i\,;\, i = 1,2,\dots,n\}$ je rozklad množiny Ω a P(H_i)>0 pro všechna i, tak pst jevu A⊆Ω lze určit ze vztahu

P(A)= P(H_1)·P(A | H_1)+ P(H_2)·P(A | H_2)+ … + P(H_n)·P(A | H_n)

- (větu užijeme tehdy, když P(A) nelze spočítat přímo, ale P(H_i)·P(A | H_i) lze určit relativně snadno)

Důkaz: plyne z axiomu 3 a vyjádření A jako sjednocení neslučitelných jevů A= (H_1∩ A)∪ (H_2∩ A) ∪ … ∪ (H_n∩ A)

(a pro každý průnik 2 jevů užijeme větu o součinu pstí)

Příklad: v čokoládovně se kompletují bonboniéry na třech výrobních linkách

Linka 1 … kompletuje 40 % produkce, pokazí 5 % bonb.

Linka 2 … kompletuje 45 % produkce, pokazí 4 % bonb.

Linka 3 … kompletuje zbytek produkce, pokazí 2 % bonb.

Zkontrolujeme náhodně vybranou bonboniéru … jaká je šance, že nebude v normě (= není správně zabalena)?

Řešení úlohy na úplnou pst je hračkou tehdy, pokud dobře určíme jevy H_1, H_2, H_3, jejichž psti známe

H_1 … ? $P(H_1)=??$

H_2 … ? $P(H_2)=??$

H_3 … ? $P(H_3)=??$

A … ??????

H_1 … čokoláda byla balena na lince 1 $P(H_1)=??$

H_2 … čokoláda byla balena na lince 2 $P(H_2)=??$

H_3 … čokoláda byla balena na lince 3 $P(H_3)=??$

A … náhodně vybraná bonboniéra je zmetkově zabalena

H_1 ... čokoláda byla balena na lince 1 $P(H_1)=0,40$

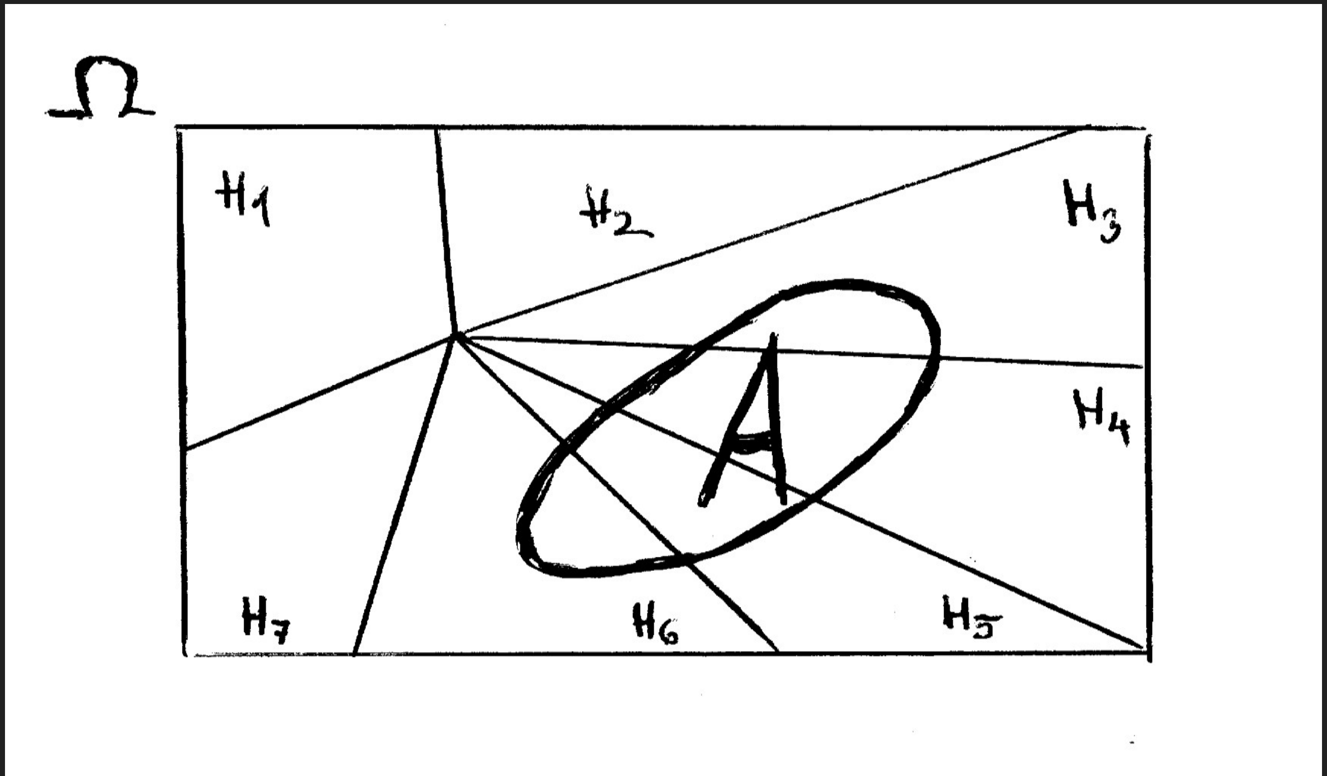
H_2 ... čokoláda byla balena na lince 2 $P(H_2)=0,45$

H_3 ... čokoláda byla balena na lince 3 $P(H_3)=0,15$

A ... náhodně vybraná bonboniéra je zmetkově zabalena

$P(A)=0,40 \cdot 0,05 + 0,45 \cdot 0,04 + 0,15 \cdot 0,02=0,041$

Otázka č. 16: Bayesův vzorec – obrazem (stejný obr, ale jiný vzorec!!!)



Ot. č 16: Bayesův vzorec – slovem
(co bychom tak mohli chtít ještě spočítat v té situaci??)

řekněte

Ot. č 16: Bayesův vzorec – slovem
(co bychom tak mohli chtít ještě spočítat v té situaci??)

$P(A \mid H_1)$ je relativně snadné většinou určit; ovšem

- $P(H_1 \mid A)$ není často jednoduché určit, i když máme k dispozici vzorec $P(H_1 \mid A) = \frac{P(H_1 \cap A)}{P(A)}$ jako základní vzorec pro podmíněnou pst

Veškerá věda Bayesova vzorce spočívá v tom, že

Do čitatele základního vzorce pro podmíněnou pst

$P(H_1 | A) = \frac{P(H_1 \cap A)}{P(A)}$ dosadíme větu o součinu

- A do jmenovatele větu o úplné psti:

$$P(H_1 | A) = \frac{P(H_1 \cap A)}{P(A)} = \frac{P(H_1) \cdot P(A|H_1)}{P(H_1) \cdot P(A | H_1) + P(H_2) \cdot P(A | H_2) + \dots + P(H_n) \cdot P(A | H_n)}$$

(všimněte si, že v čitateli Bayesova vzorce se vyskytuje jeden člen sumy ze jmenovatele ... pomůcka pro zapamatování)

Tímto způsobem nespočteme jen $P(H_1 | A)$, ale také $P(H_i | A)$ pro jakékoli *i*:

- $P(H_i | A) = \frac{P(H_i \cap A)}{P(A)} = \frac{P(H_i) \cdot P(A|H_i)}{P(H_1) \cdot P(A | H_1) + P(H_2) \cdot P(A | H_2) + \ldots + P(H_n) \cdot P(A | H_n)}$

(v čitateli je pevné í, ve jmenovateli se í mění)

Vraťme se k příkladu s čokoládovnou:
v čokoládovně se kompletují bonboniéry na třech výrobních linkách

Linka 1 … kompletuje 40 % produkce, pokazí 5 % bonb.

Linka 2 … kompletuje 45 % produkce, pokazí 4 % bonb.

Linka 3 … kompletuje zbytek produkce, pokazí 2 % bonb.

Zkontrolovali jsme náhodně vybranou bonboniéru a zjistili, že není v normě (tato podmínka nastala) … s jakou pstí byla zkompletována na lince 1 (nebo 2, nebo 3)?

Tohle všechno už jsme spočítali v otázce 15:

H_1 … čokoláda byla balena na lince 1 $P(H_1)=0{,}40$

H_2 … čokoláda byla balena na lince 2 $P(H_2)=0{,}45$

H_3 … čokoláda byla balena na lince 3 $P(H_3)=0{,}15$

A … náhodně vybraná bonboniéra je zmetkově zabalena

$P(A)=0{,}40\cdot 0{,}05+0{,}45\cdot 0{,}04+0{,}15\cdot 0{,}02=0{,}041$

$P(H_1|A)=?$ $P(H_2|A)=??$ $P(H_3|A)=???$

Nyní přidáme $P(H_1|A)=?$ $P(H_2|A)=??$ $P(H_3|A)=???$

H_1 … čokoláda byla balena na lince 1 $P(H_1)=0{,}40$

H_2 … čokoláda byla balena na lince 2 $P(H_2)=0{,}45$

H_3 … čokoláda byla balena na lince 3 $P(H_3)=0{,}15$

- A … náhodně vybraná bonboniéra je zmetkově zabalena

$P(A)=0{,}40\cdot 0{,}05+0{,}45\cdot 0{,}04+0{,}15\cdot 0{,}02=0{,}041$

$$P(H_1|A)=\frac{0{,}40\cdot 0{,}05}{0{,}041},\ P(H_2|A)=\frac{0{,}45\cdot 0{,}04}{0{,}041},\ P(H_1|A)=\frac{0{,}15\cdot 0{,}02}{0{,}041},$$

$P(H_i)$ … apriorní psti (a priori = předem, před měřením)
$P(H_i| A)$ … aposteriorní psti = psti po měření

Apriorní psti:

$P(H_1)$=0,40

$P(H_2)$=0,45

$P(H_3)$=0,15

Aposteriorní psti:

$P(H_1| A)$=0,4878

$P(H_2| A)$=0,439

$P(H_3| A)$=0,0732

Po nalezení chybně balené čokolády narostla pst, že byla balena na lince 1, kdežto klesly psti, že byla balena na lince 2 nebo že byla balena na lince 3

Rekapitulace otázek:

11. Věta o součtu pstí.
12. Stochasticky nezávislé jevy.
13. Podmíněná pst
14. Věta o součinu pstí.
15. Věta o úplné psti
16. Bayesův vzorec