# Bayesovské rozhodování

X – sada pozorování

K – sada tříd

D – sada možných rozhodnutí

P(x,k) – pravděpodobnostní rozdělení jevů; P(x,k) = p(x|k)p(k)

W(K,D) – ztrátová funkce

q – strategie, máme D^X strategií

R(q) – Bayesovský risk

$R\left(q\right)= \sum\_{x\in X}^{}\sum\_{k\in K}^{}p\left(x,k\right)∙W\left(k,q\left(x\right)\right)$

$q^{\*}$ - optimální Bayesovská strategie minimalizuje Bayesovský risk

$q^{\*}=argmin R(q) $

Pro diskrétní přepisujeme R(q) jako

$R\left(q\right)= \sum\_{k\in K}^{}p(k)\sum\_{x\in X}^{}p\left(x|k\right)∙W\left(k,q\left(x\right)\right)$

Pro spojité v R(q) 2. suma přechází na integrál

$R\left(q\right)= \sum\_{k\in K}^{}p(k)\int\_{x\in X}^{}p\left(x|k\right)∙W\left(k,q\left(x\right)\right)$

Spojitý případ s normálním rozdělením můžeme přepsat jako

$q^{\*}=arg\max\_{K} p(x,k) $ (1)

*((1) platí pro každý případ ale pouze s vyváženou ztrátovou funkcí)*

To pak přechází v problém

$p\left(x|A\right)∙p\left(A\right)>p(x|C)∙p(C)$

Výpočtem získáme parabolu a dva threshold body

Např. – líný krátkozraký student, mincovní automat

Bayesovská strategie je jasně daná, není vhodné náhodné rozhodování

Limity Bayesovské metody

* W(k, q(x)) je nutná, ne vždy ale dává smysl a ne vždy se dají hodnoty srovnat (cena za smrt pacienta, věznění nevinného vs běhání vraha na svobodě)
* P(k) – musí existovat a musíme ji znát; někdy ale nedává smysl o ní přemýšlet (nepřátelské letadlo na radaru)
* P(x|k) může být funkcí vedlejší informace -> P(x|k, z)
	+ x – pozorování; k – to na čem záleží; z – náhodný zásah, to jak jsme ovlivnili vztah x a k

#  Non-Bayes rozhodování

## Neyman – Pearson

K ={D, N}; D – dangerous, N – normal

D = K; rozhodnutí

X – set pozorování

Známe p(x|K) ale NENÍ známé p(K)

Hledáme strategii q: X->D

Dostáváme podmínku, že chyba ve třídě D ($ε\_{D}$)musí být menší, než uživatelem stanovaná chybovost $(ε)$.

Chyba ve třídě N $(ε\_{N})$ musí být tak malá jak to jde.

$ε\_{D}=\sum\_{x:q(x)\ne D}^{}p(x|D) overlooked danger$

$ε\_{N}=\sum\_{x:q(x)\ne N}^{}p\left(x|N\right) false alarm$

Formulujeme optimalizační úlohu

$q^{\*}=argmin\sum\_{x:q\left(x\right)\ne N}^{}p\left(x|N\right) $

$s.t. \sum\_{x:q(x)\ne D}^{}p(x|D) \leq ε$

Používáme likelihood ratio (ukazuje v jakém pořadí postupovat, na něm hledáme prahy)

$LR=\frac{p\left(x|N\right)}{p\left(x|D\right)} $

$LR>μ \rightarrow q\left(x\right)=N$

$LR \leq μ \rightarrow q\left(x\right)=D$

NP stejný jako testování hypotéz ve statistice

Problém – na třídy koukám různě

Strategii je možné randomizovat

* např v 1/3 případů pro dané x\_i z X se rozhodnout pro N a ve zbylých 2/3 se rozhodnout pro D
* ? randomizace funguje pouze pro diskrétní případy? (Pro cont. dist není randomizace vůbec nutná, pokud existuje řešení, pak ho najdeme přesně)

## Minimax

K = {1, 2, .., N}; třídy

X – set pozorování

Známe p(x|k) pro všechna k z K

Hledáme strategii q

$q^{\*}=argmin max ε\left(k\right) $

$kde ε\left(k\right)=\sum\_{x:q\left(x\right)\ne k}^{}p\left(x|k\right)$

Řešíme pomocí LR

* Nakreslíme si ji
* Juknem kolik hledáme prahů
	+ Speciální případy
		- Monotónní funkce – víme, že ať práh hodíme kamkoli, vždy ji rozsekneme jen na 2 části, budeme mít jen jedno u
		- Symetrická fce – víme, že ať ji rozdělíme jakkoli, vždy ji rozsekneme na 3 části, budeme mít 2 symetrické u *(např. cosinus je symetrický a existuje nekonečné mnoho řešení, ale u funkcí s jedním extrémem platí)*
* Hledáme nejlepší práh

## Wald task

K ={1, 2} + {?} -> mám možnost říkat nevím

Máme omezení chyby pro obě třídy

q:X->D charakterizuje

$ε\_{1}=\sum\_{X:q\left(x\right)=2}^{}p(x|1)$

$ε\_{2}=\sum\_{X:q\left(x\right)=1}^{}p(x|2)$

$κ\_{1}=\sum\_{X:q\left(x\right)=?}^{}p(x|1)$

$κ\_{2}=\sum\_{X:q\left(x\right)=?}^{}p(x|2)$

Optimalizační formulace hledání strategie:

$q^{\*}= argmin (q:X\rightarrow D) max (i=\{1,2\}) κi $

$s.t. ε\_{1}\leq ε ; ε\_{2}\leq ε $

Opět používáme likelihood ratio

$r\left(x\right)=\frac{p\left(1\right)}{p\left(2\right)} $

Optimální strategie:

$q\left(x\right)=\{2; r\left(x\right)<μ\_{l} 1; r\left(x\right)>μ\_{h} ?; μ\_{l}\leq r(x)\leq μ\_{h} $

# Shrnutí

**Bayes** – znám P(K)

| Mám | K; X; D |
| --- | --- |
|  | P(x, k)=P(x|k)P(K) |
| Rizika/omezení | W(k,d) |
| Strategie | Minimalizujeme risk strategie R(q)$q^{\*}=argmin R(q) $$R\left(q\right)= \sum\_{x\in X}^{}\sum\_{k\in K}^{}p\left(x,k\right)∙W\left(k,q\left(x\right)\right)$ |
| Speciálně | Pro 2 třídy s normálním rozdělením a vyváženou ztrátovou funkcí můžeme použít nerovnici$p\left(x|A\right)∙p\left(A\right)>p(x|C)∙p(C)$ |

**Non-Bayes** – P(K) neznámé nebo neexistující

|  | Neuman-Pearson | Minimax | Wald task |
| --- | --- | --- | --- |
| Mám | K = {D, N}; X; D | K={1, 2, …,N}; X; D | K= {1, 2} + {?}; X; D |
|  | P(x|k) | P(x|k) | P(x|k) |
| Rizika/ omezení | Maximální přípustná chyba třídy D | --- | Maximální přípustná chyba tříd 1 a 2 |
| Strategie | Minimalizujeme chybu třídy N za podmínku, že D je menší než přípustná chyba | Minimalizujeme největší chybu z chyb | Minimalizujeme větší chybu z rozhodnutí ?, za podmínky že chyby tříd 1 a 2 nesmí přesáhnout danou mez |
| Speciálně | Používám likelihood ratioMožnost randomizovat | Používám likelihood ratio | Používám likelihood ratioMožnost říkat nevím {?} |