# Bayesovská rozhodovací úloha, Neyman-Pearson, Minimax

June 8, 2012

## 1 Bayesovská rozhodovací úloha

Situace objektu je charakterizována dvěma parametry.

- $x$, což je pozorování objektu
- $k$, což je nepozorovatelný skutečný stav objektu

Pojmy:

- X je konečná množina pozorování, $x \in X$
- K je konečná množine skrytých stavů, $k \in K$
- D je konečná množina rozhodnutí
- $p_{XK} : X \times K \to \mathbb{R}$ je pravděpodobnost, že objekt je ve stavu $k$ a je pozorováno $x$
$$p_{XK} = p_{Xk}(x|k) \cdot p_K(k) \tag{1}$$
- $W : K \times D \to \mathbb{R}$ ztrátová funkce, $W(k, d)$ je ztráta pro situaci kdy je objekt ve stavu $k$ a bylo učiněno rozhodnutí $d$
- $q : X \to D$ je rozhodovací funkce která přiřazuje každému $x \in X$ $q(x) \in D$
- Bayesovský risk:
$$R(q) = \sum_{x \in X} \sum_{k \in K} p_{XK}(x, k) W(k, q(x)) \tag{2}$$

Bayesovský rozhodovací problém se dá formulovat jako nalezení strategie $q : X \to D$ minimalizující bayesovský risk. Ze vztahů ve Figure 1 plyne, že optimální strategie $q^*(x)$ může být nalezena pomocí určení optimálního rozhodnutí pro každé pozorování. Příklad aplikace na zjednodušenou situaci ve Figurách 2 a 3. Příklad s reject option v Figurách 4, 5 a 6 (prakticky se najde stav $k$, který má největší aposteriorní pravděpodobnost pro $x$, pokud je tato pravděpodobnost větší než $1 - \epsilon$ pak se rozhodujeme pro k pokud ne pak říkáme nevím).

$$R(q^*) = \min_{q \in X \to D} \sum_{x \in X} \sum_{k \in K} p_{XK}(x,k) W(k, q(x))$$

$$R(q^*) = \sum_{x \in X} \min_{q(x) \in D} \sum_{k \in K} p_{XK}(x,k) W(k, q(x))$$

$$R(q^*) = \sum_{x \in X} \min_{q(x) \in D} p(x) \sum_{k \in K} p_{K|X}(k|x) W(k, q(x))$$

$$R(q^*) = \sum_{x \in X} p(x) \min_{q(x) \in D} \sum_{k \in K} p_{K|X}(k|x) W(k, q(x))$$

$$R(q^*) = \sum_{x \in X} p(x) R(x, d^*)$$

where

$$R(x, d^*) = \sum_{k \in K} p_{K|X}(k \mid x)\, W(k, d^*) = \min_{q(x) \in D} \sum_{k \in K} p_{K|X}(k|x) W(k, q(x))\,.$$

Figure 1: Bayesovský risk

Consider the following problem:

- The object is in an unknown state $k$.
- The set of possible decisions $D$ and of hidden states $K$ coincide, $D = K$.
- The cost function assigns a unit penalty when $q(x) \neq k$ occurs and no penalty otherwise, i.e.

$$W\big(k, q(x)\big) = \begin{cases} 0 & \text{if} \quad q(x) = k \\ 1 & \text{if} \quad q(x) \neq k \end{cases}$$

The Bayesian risk

$$\begin{aligned} R(q) &= \sum_{x \in X} \sum_{k \in K} p_{XK}(x,k)\, W\big(k, q(x)\big) = \sum_{x \in X} p_X(x) \sum_{k \neq q(x)} p_{Kx}(k|x) \\ &= \sum_{x \in X} p_X(x)(1 - p_{xk}(q(x)|x) \end{aligned}$$

is then equal the probability of the situation $q(x) \neq k$ (probability of classification error) or 1 – probability of correct decision.

Figure 2: Bayesovský risk příklad 1

We have to determine the strategy $q\colon X \to K$ which minimises the risk, i.e.,

$$\begin{aligned} q(x) &= \operatorname*{argmin}_{k\in K} \sum_{k^*\in K} p_{XK}(x,k^*)\, W(k^*,k) \\ &= \operatorname*{argmin}_{k\in K} p_X(x) \sum_{k^*\in K} p_{K|X}(k^*\,|\,x)\, W(k^*,k) = \operatorname*{argmin}_{k\in K} \sum_{k^*\in K} p_{K|X}(k^*\,|\,x)\, W(k^*,k) \\ &= \operatorname*{argmin}_{k\in K} \sum_{k^*\in K\setminus\{k\}} p_{K|X}(k^*\,|\,x) \\ &= \operatorname*{argmin}_{k\in K} \left( \sum_{k^*\in K} p_{K|X}(k^*\,|\,x) - p_{K|X}(k\,|\,x) \right) \\ &= \operatorname*{argmin}_{k\in K} \left(1 - p_{K|X}(k\,|\,x)\right) = \operatorname*{argmax}_{k\in K} p_{K|X}(k\,|\,x)\,. \end{aligned}$$

The result shows that the *a posteriori* probability of each state $k$ is to be calculated for the observation $x$ and the optimal decision is in favour of the most probable state. The the maximum *a posteriori strategy* is the Bayesian strategy for the 0-1 loss function.

Dichotomy. In the situation with two possible decisions (and classes), the optimal decision can be expressed as a sign of discriminative function $g(x) = p_{k|x}(1\,|\,x) - p_{k|x}(0\,|\,x)$.

Figure 3: Bayesovský risk příklad 2

Consider an examination where for each question there are three possible answers: yes, no, `not known`. If your answer is correct, 1 point is added to your score. If your answer is wrong, 3 points are subtracted. If your answer is `not known`, your score is unchanged. What is the optimal Bayesian strategy if for each question you know the probabilities that $p(\text{yes})$ is the right answer?

Note that adding a fixed amount to all penalties and multiplying all penalties by a fixed amount does not change the optimal strategy. Adding 3 and multiplying by $1/4$ leads to 1 point for correct answer, $3/4$ for `not known` and 0 points of a wrong answer.

Any problem of this type can be transformed to an equivalent problem with penalty 0 for the correct answer, 1 for the wrong answer, and $\epsilon$ for `not known`. In realistic problems, $\epsilon \in (0,1)$, since $\epsilon \geq 1$ means it is always better to guess than to say `not known`; $\epsilon \leq 0$ states that saying `not known` is preferred to giving the correct answer.

Let us solve the problem formally.

Figure 4: Bayesovský risk s nevím příklad 1

Let $X$ and $K$ be sets of observations and states, $p_{XK}\colon X \times K \to \mathbb{R}$ be a probability distribution and $D = K \cup \{\texttt{not known}\}$ be a set of decisions.

Let us define $W(k,d)$, $k \in K$, $d \in D$:

$$W(k,d) = \begin{cases} 0, & \text{if } d = k\,, \\ 1, & \text{if } d \neq k \text{ and } d \neq \texttt{not known}\,, \\ \varepsilon, & \text{if } d = \texttt{not known}\,. \end{cases}$$

Find the Bayesian strategy $q\colon X \to D$. The decision $q(x)$ corresponding to the observation $x$ has to minimises the partial risk,

$$q(x) = \operatorname*{argmin}_{d \in D} \sum_{k^* \in K} p_{K|X}(k^* \,|\, x)\, W(k^*, d)\,.$$

Figure 5: Bayesovský risk s nevím příklad 2

There holds for $R(x, \texttt{not known})$

$$\begin{aligned} R(x, \texttt{not known}) &= \sum_{k^* \in K} p_{K|X}(k^* \,|\, x)\, W(k^*, \texttt{not known}) \\ &= \sum_{k^* \in K} p_{K|X}(k^* \,|\, x)\, \varepsilon = \varepsilon\,. \end{aligned}$$

The decision rule becomes

$$q(x) = \begin{cases} \operatorname*{argmax}_{k \in K} p_{K|X}(k \,|\, x)\,, & \text{if } 1 - \max_{k \in K} p_{K|X}(k \,|\, x) < \varepsilon\,, \\ \texttt{not known}\,, & \text{if } 1 - \max_{k \in K} p_{K|X}(k \,|\, x) \geq \varepsilon\,. \end{cases}$$

Figure 6: Bayesovský risk s nevím příklad 3

# 2 Nebayesovské rozhodování

V situacích kdy neznáme apriorní pravděpodobnosti (např protože k není náhodný jev) nebo nemůžeme porovnávat ztrátu pro různé situace.

## 2.1 Neyman-Pearson

Předpokládejme dva stavy $k = 1$ bezpečný a $k = 2$ nepezpečný. Náš úkol je rozdělit množinu pozorování $X$ na množinu $X_1$ bezpečných stavů a množinu $X_2$ nebezpečných stavů takových že $X_1 \cup X_2 = X$ a $X_1 \cap X_2 = \emptyset$. $p_{X|K}(x|k)$ jsou známé. Strategie je charakterizována dvěmi čísly. Přehlédnuté nebezpečí(overlooked danger) $\sum\limits_{x \in X_2} p_{X|K}(x|1)$ a planý poplach(false alarm) $\sum\limits_{x \in X_1} p_{X|K}(x|2)$. Hlavní myšlenka tohoto rozhodování je minimalizace false alarm, za předpokladu že overlooked danger je menší než $\epsilon$.

## 2.2 Minimax

Minimax minimalizuje nejhorší možnou chybu vzhledem k neznámým apriorním pravděpodobnostem, což pro rozhodování mezi 2mi stavy odpovídá minimalizaci $argmin_{d(x)} \max\{\int_{X_2} p(x|1)dx, \int_{X_1} p(x|2)dx\}$. Risk minimaxního rozhodování není nikdy menší než risk bayesovského rozhodování pro nejnepříznivější apriorní pravděpodobnosti.