Chyba rozhodování

Jiří Vomlel

Laboratoř inteligentních systémů
Vysoká škola ekonomická Praha

Tato prezentace je k dispozici na:
http://www.utia.cas.cz/vomlel/

Chyby rozhodování

Nechť

$$\delta(a,b) = \begin{cases} 1 & \text{když } a = b \\ 0 & \text{jindy.} \end{cases}$$

Chyby rozhodování definujeme:

$$p_d(a_1 \to a_2) = \sum_{\mathbf{r} \in Q : d(\mathbf{r}) = a_2} P(\mathbf{r}|a_1) = \sum_{\mathbf{r} \in Q} P(\mathbf{r}|a_1) \cdot \delta(d(\mathbf{r}), a_2)$$

$$p_d(a_2 \to a_1) = \sum_{\mathbf{r} \in Q : d(\mathbf{r}) = a_1} P(\mathbf{r}|a_2) = \sum_{\mathbf{r} \in Q} P(\mathbf{r}|a_2) \cdot \delta(d(\mathbf{r}), a_1)$$

Rodina všech rozhodovacích funkcí

$\mathcal{D}$ ... rodina všech rozhodovacích funkcí

$$\mathcal{D} = Q_1 \times Q_2 \times \ldots \times Q_N \mapsto \{a_1, a_2\}$$

N ... počet otázek s možností dvou odpovědí

2^N ... počet všech možných kombinací odpovědí

... pro každou odpověd máme dvě možnosti zařazení - a_1 a a_2

2^{2^N} ... rozhodovacích funkcí

Příklad Pro $N = 1$ máme 4 rozhodovací funkce:

$$\begin{aligned} d_1(0) &= a_1 & \qquad d_1(1) &= a_1 \\ d_2(0) &= a_1 & d_2(1) &= a_2 \\ d_3(0) &= a_2 & d_3(1) &= a_1 \\ d_4(0) &= a_2 & d_4(1) &= a_2 \end{aligned}$$

Dominance rozhodovacích funkcí

Definice

d_1 dominuje d_2 jestliže buď

$$\begin{aligned} p_{d_1}(a_1 \longrightarrow a_2) &\leq p_{d_2}(a_1 \longrightarrow a_2) \\ p_{d_1}(a_2 \longrightarrow a_1) &< p_{d_2}(a_2 \longrightarrow a_1) \end{aligned}$$

nebo

$$\begin{aligned} p_{d_1}(a_1 \longrightarrow a_2) &< p_{d_2}(a_1 \longrightarrow a_2) \\ p_{d_1}(a_2 \longrightarrow a_1) &\leq p_{d_2}(a_2 \longrightarrow a_1) \end{aligned}$$

Přípustná rozhodovací funkce

Definice

Rozhodovací funkce $d \in \mathcal{D}$ je **přípustná**, jestliže v $\mathcal{D}$ neexistuje rozhodovací funkce, která by ji dominovala.

Příklad

Pro $N = 1$ máme 4 rozhodovací funkce:

$$\begin{aligned} d_1(0) &= a_1 & \qquad d_1(1) &= a_1 \\ d_2(0) &= a_1 & \qquad d_2(1) &= a_2 \\ d_3(0) &= a_2 & \qquad d_3(1) &= a_1 \\ d_4(0) &= a_2 & \qquad d_4(1) &= a_2 \end{aligned}$$

Z definice chyby rozhodování

$$
\begin{aligned}
p_d(a_1 \to a_2) &= \sum_{\mathbf{r} \in Q} P(\mathbf{r}|a_1) \cdot \delta(d(\mathbf{r}), a_2) \\
p_d(a_2 \to a_1) &= \sum_{\mathbf{r} \in Q} P(\mathbf{r}|a_2) \cdot \delta(d(\mathbf{r}), a_1)
\end{aligned}
$$

dostáváme:

i	$p_d(a_1 \to a_2)$	$p_d(a_2 \to a_1)$
1	0	$P(0 \mid a_2) + P(1 \mid a_2)$
2	$P(1 \mid a_1)$	$P(0 \mid a_2)$
3	$P(0 \mid a_1)$	$P(1 \mid a_2)$
4	$P(0 \mid a_1) + P(1 \mid a_1)$	0

Mějme pravděpodobnostní distribuci $P(Q_1, A)$ definovanou následující tabulkou:

	$Q_1 = 0$	$Q_1 = 1$
$A = a_1$	0.2	0.1
$A = a_2$	0.3	0.4

Z ní spočteme podmíněnou pravděpodobnostní distribuci $P(Q_1 \mid A)$

	$Q_1 = 0$	$Q_1 = 1$
$A = a_1$	$\frac{2}{3}$	$\frac{1}{3}$
$A = a_2$	$\frac{3}{7}$	$\frac{4}{7}$

Pro chyby rozhodování dostáváme:

i	$p_{d_i}(a_1 \to a_2)$	$p_{d_i}(a_2 \to a_1)$
1	0	1
2	$\frac{1}{3}$	$\frac{3}{7}$
3	$\frac{2}{3}$	$\frac{4}{7}$
4	1	0

Rozhodovací funkce d_3 je dominovaná rozhodovací funkcí d_2 a tudíž je **nepřípustná**.

Ostatní rozhodovací funkce nejsou dominované a tudíž jsou **přípustné**.

Příklad

Zjistěte, které rozhodovací funkce z rodiny všech možných rozhodovacích funkcí pro $N = 1$ jsou přípustné pro pravděpodobnostní distribuci $P(Q_1, A)$ definovanou následující tabulkou:

	$Q_1 = 0$	$Q_1 = 1$
$A = a_1$	0.1	0.2
$A = a_2$	0.3	0.4

Řešení příkladu

Z $P(Q_1, A)$ spočteme podmíněnou pravděpodobnostní distribuci $P(Q_1 \mid A)$

	$Q_1 = 0$	$Q_1 = 1$
$A = a_1$	$\frac{1}{3}$	$\frac{2}{3}$
$A = a_2$	$\frac{3}{7}$	$\frac{4}{7}$

a pro chyby rozhodování dostáváme:

i	$p_{d_i}(a_1 \rightarrow a_2)$	$p_{d_i}(a_2 \rightarrow a_1)$
1	0	1
2	$\frac{2}{3}$	$\frac{3}{7}$
3	$\frac{1}{3}$	$\frac{4}{7}$
4	1	0

Všechny rozhodovací funkce jsou **připustné**.

Jsme schopni zkonstruovat přípustné rozhodovací funkce?

Tvrzení

Nechť $w(a_1)$ a $w(a_2)$ jsou nezáporná čísla. Potom rozhodovací funkce

$$d(\mathbf{r}) = \begin{cases} a_1 & \text{když } w(a_1) \cdot P(\mathbf{r} \mid a_1) > w(a_2) \cdot P(\mathbf{r} \mid a_2) \\ a_2 & \text{když } w(a_1) \cdot P(\mathbf{r} \mid a_1) < w(a_2) \cdot P(\mathbf{r} \mid a_2) \end{cases}$$

je přípustná. Navíc tato rozhodovací funkce minimalizuje funkci

$$S(d) = w(a_1) \cdot p_d(a_1 \longrightarrow a_2) + w(a_2) \cdot p_d(a_2 \longrightarrow a_1) \ .$$

Důkaz tvrzení - část 1

$$
\begin{aligned}
S(d) &= w(a_1) \cdot p_d(a_1 \to a_2) + w(a_2) \cdot p_d(a_2 \to a_1) \\
&= \sum_{\mathbf{r} \in Q} w(a_1) \cdot P(\mathbf{r}|a_1) \cdot \delta(d(\mathbf{r}), a_2) + \sum_{\mathbf{r} \in Q} w(a_2) \cdot P(\mathbf{r}|a_2) \cdot \delta(d(\mathbf{r}), a_1) \\
&= \sum_{\mathbf{r} \in Q} \left[w(a_1) \cdot P(\mathbf{r}|a_1) \cdot \delta(d(\mathbf{r}), a_2) \ + \ w(a_2) \cdot P(\mathbf{r}|a_2) \cdot \delta(d(\mathbf{r}), a_1) \right]
\end{aligned}
$$

Povšimněme, si, že

$$
\delta(d(\mathbf{r}), a_2) + \delta(d(\mathbf{r}), a_1) = 1 \ .
$$

Proto S minimalizuje rozhodovací funkce d' taková, že pro každé $\mathbf{r}$ vždy přičte menší z chyb

$$
\begin{aligned}
x_1(\mathbf{r}) &= w(a_1) \cdot P(\mathbf{r}|a_1) \\
x_2(\mathbf{r}) &= w(a_2) \cdot P(\mathbf{r}|a_2)
\end{aligned}
$$

Důkaz tvrzení - část 2

Minimalizujeme

$$S(d) \quad = \quad \sum_{\mathbf{r} \in Q} [x_1(\mathbf{r}) \cdot \delta(d(\mathbf{r}), a_2) \; + \; x_2(\mathbf{r}) \cdot \delta(d(\mathbf{r}), a_1)]$$

Přičtení menší z chyb je realizováno funkcí

$$d(\mathrm{r}) \; = \; d'(\mathrm{r}) \quad = \quad \begin{cases} a_1 & \text{když } x_1(\mathbf{r}) \; > \; x_2(\mathbf{r}) \\ a_2 & \text{když } x_1(\mathbf{r}) \; < \; x_2(\mathbf{r}) \end{cases}$$

což je přesně funkce d z Tvrzení:

$$d(\mathbf{r}) \quad = \quad \begin{cases} a_1 & \text{když } w(a_1) \cdot P(\mathbf{r} \mid a_1) \; > \; w(a_2) \cdot P(\mathbf{r} \mid a_2) \\ a_2 & \text{když } w(a_1) \cdot P(\mathbf{r} \mid a_1) \; < \; w(a_2) \cdot P(\mathbf{r} \mid a_2) \end{cases}$$

Tím jsme ukázali, že d minimalizuje S.

Důkaz tvrzení - část 3

Kdyby existovala rozhodovací funkce $d'' \in \mathcal{D}$ která by dominovala d pak by buď

$$\begin{aligned} p_{d''}(a_1 \to a_2) &\leq p_d(a_1 \to a_2) \\ p_{d''}(a_2 \to a_1) &< p_d(a_2 \to a_1) \end{aligned}$$

nebo

$$\begin{aligned} p_{d''}(a_1 \to a_2) &< p_d(a_1 \to a_2) \\ p_{d''}(a_2 \to a_1) &\leq p_d(a_2 \to a_1) \end{aligned}$$

což by nutně znamenalo, že $S(d'') < S(d)$, což je ve sporu s tím, že d minimalizuje S.

Tudíž neexistuje $d'' \in \mathcal{D}$, která by dominovala d. Tím jsme dokázali, že d je přípustná. $\square$

Rozhodovací funkce optimální vzhledem ke kritériu maximální věrohodnosti

minimalizuje

$$S(d) \quad = \quad w(a_1) \cdot p_d(a_1 \to a_2) + w(a_2) \cdot p_d(a_2 \to a_1) \ .$$

pro

$$\boxed{w(a_1) = w(a_2) = 1.}$$

To znamená, že

$$e'_P(d) \quad = \quad p_d(a_1 \to a_2) + p_d(a_2 \to a_1) \ .$$

Optimální rozhodovací funkce je (dle Tvrzení):

$$d(\mathbf{r}) \quad = \quad \begin{cases} a_1 & \text{když } P(\mathbf{r} \mid a_1) \ > \ P(\mathbf{r} \mid a_2) \\ a_2 & \text{když } P(\mathbf{r} \mid a_1) \ < \ P(\mathbf{r} \mid a_2) \end{cases}$$

Bayesovsky optimální rozhodovací funkce

minimalizuje

$$S(d) \quad = \quad w(a_1) \cdot p_d(a_1 \to a_2) + w(a_2) \cdot p_d(a_2 \to a_1) \quad .$$

pro

$$w(a_1) = P(a_1), \;\; w(a_2) = P(a_2).$$

To znamená, že

$$e_P(d) \quad = \quad P(a_1) \cdot p_d(a_1 \to a_2) + P(a_2) \cdot p_d(a_2 \to a_1) \quad .$$

Optimální rozhodovací funkce je (dle Tvrzení):

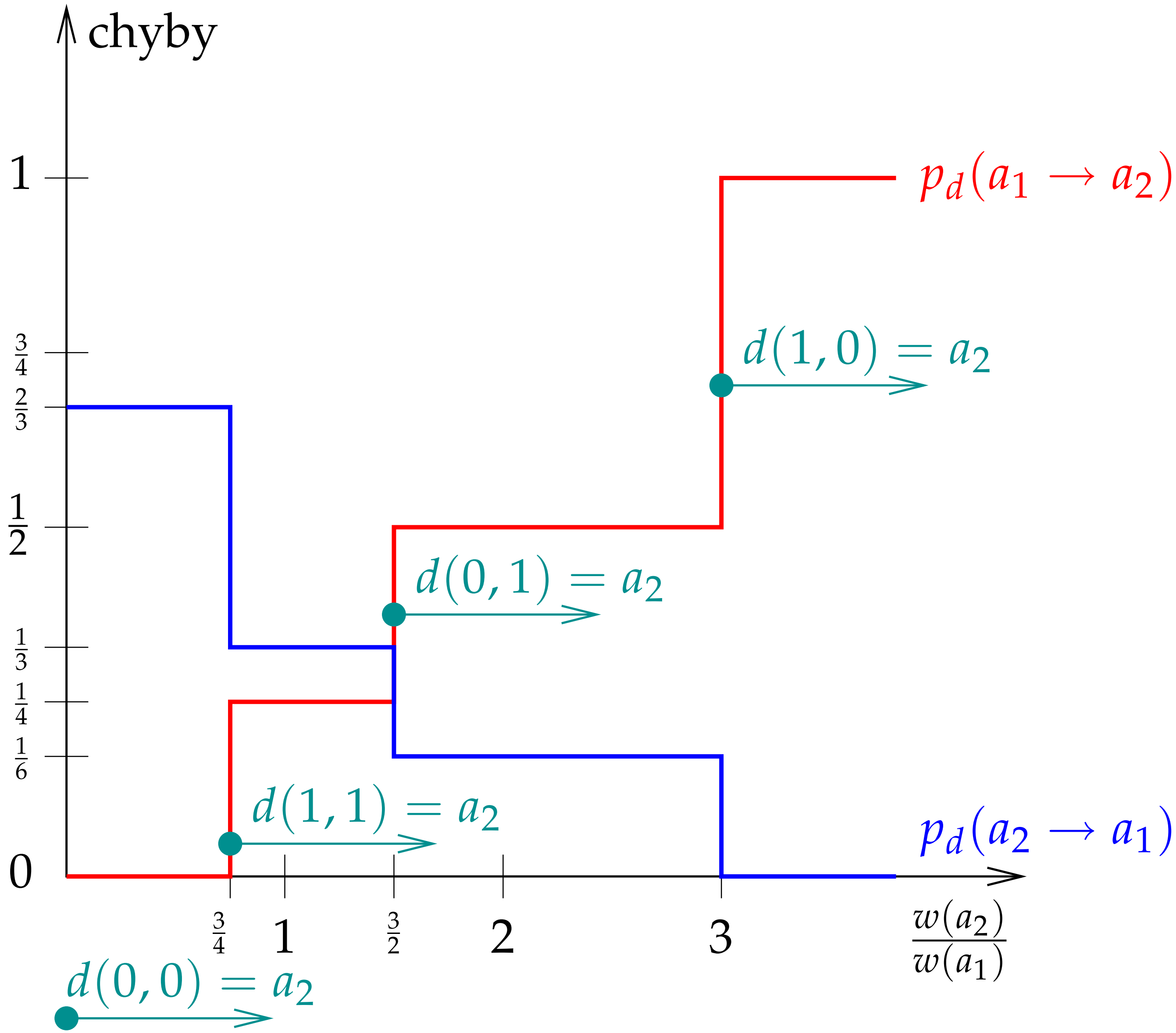
$$d(\mathbf{r}) \quad = \quad \begin{cases} a_1 & \text{když } P(a_1) \cdot P(\mathbf{r} \mid a_1) \; > \; P(a_2) \cdot P(\mathbf{r} \mid a_2) \\ a_2 & \text{když } P(a_1) \cdot P(\mathbf{r} \mid a_1) \; < \; P(a_2) \cdot P(\mathbf{r} \mid a_2) \end{cases}$$

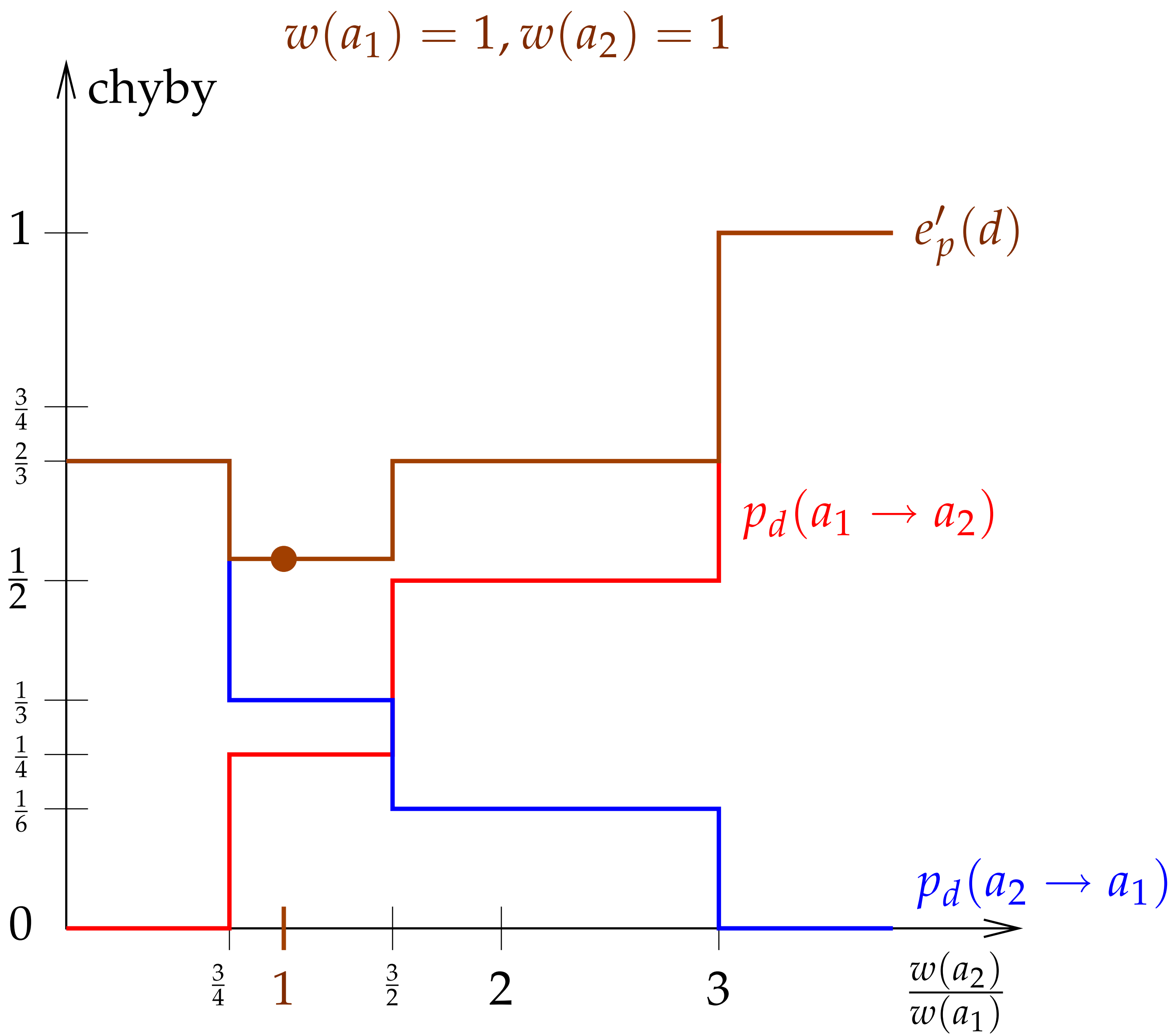
Příklad z minulého týdne

Pravděpodobnostní distribuce $P(Q_1, Q_2, A)$

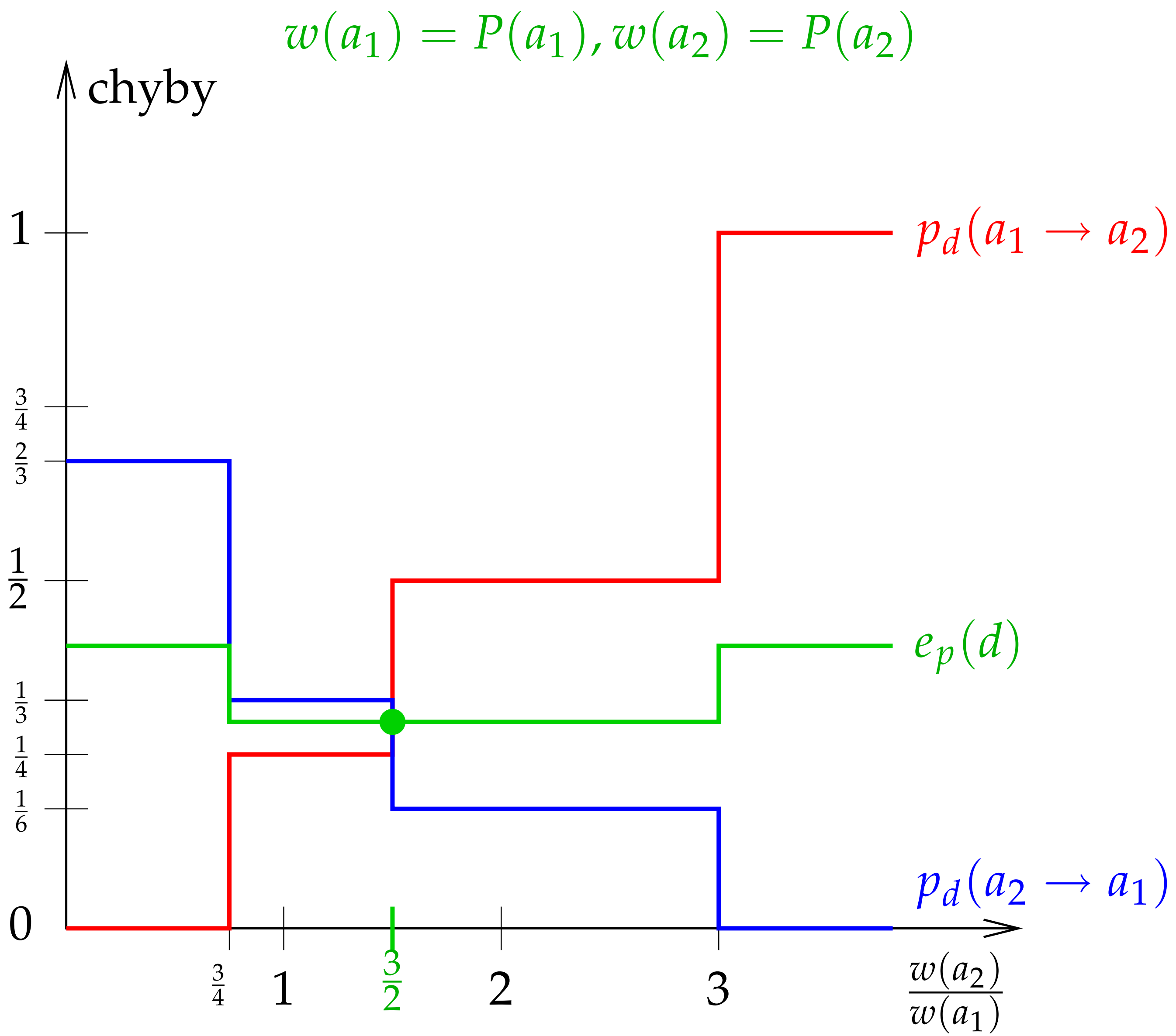
	$Q_1 = 0$		$Q_1 = 1$	
	$Q_2 = 0$	$Q_2 = 1$	$Q_2 = 0$	$Q_2 = 1$
$A = a_1$	0	0.1	0.2	0.1
$A = a_2$	0.2	0.1	0.1	0.2

(Q_1, Q_2)	$P(Q_1, Q_2 \mid A = a_1)$	$P(Q_1, Q_2 \mid A = a_2)$
$(0, 0)$	0	$\frac{1}{3}$
$(0, 1)$	$\frac{1}{4}$	$\frac{1}{6}$
$(1, 0)$	$\frac{1}{2}$	$\frac{1}{6}$
$(1, 1)$	$\frac{1}{4}$	$\frac{1}{3}$


chyby
1
$\frac{3}{4}$
$\frac{2}{3}$
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{1}{4}$
$\frac{1}{6}$
0
$p_d(a_1 \to a_2)$
$d(1,0) = a_2$
$d(0,1) = a_2$
$d(1,1) = a_2$
$p_d(a_2 \to a_1)$
$\frac{3}{4}$
1
$\frac{3}{2}$
2
3
$\frac{w(a_2)}{w(a_1)}$
$d(0,0) = a_2$



$w(a_1) = 1, w(a_2) = 1$
chyby
1
$\frac{3}{4}$
$\frac{2}{3}$
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{1}{4}$
$\frac{1}{6}$
0
$e'_p(d)$
$p_d(a_1 \longrightarrow a_2)$
$p_d(a_2 \longrightarrow a_1)$
$\frac{3}{4}$
1
$\frac{3}{2}$
2
3
$\frac{w(a_2)}{w(a_1)}$


$w(a_1) = P(a_1), w(a_2) = P(a_2)$
chyby
1
$\frac{3}{4}$
$\frac{2}{3}$
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{1}{4}$
$\frac{1}{6}$
0
$p_d(a_1 \longrightarrow a_2)$
$e_p(d)$
$p_d(a_2 \longrightarrow a_1)$
$\frac{3}{4}$
1
$\frac{3}{2}$
2
3
$\frac{w(a_2)}{w(a_1)}$