

Přednáška 6 – Odhad diskrétního modelu

Konjugovaná hp – **Dirichletovo** rozdělení (Dir)

$$f(\Theta|d(t-1)) \propto \prod_{\forall y|\psi} (\Theta_{y|\psi})^{\nu_{y|\psi;t-1}}$$

ν_{t-1} — tabulka, $\nu_{y|\psi;t-1}$ — její prvky

Bayesovo pravidlo

$$\underbrace{f(\Theta|d(t))}_{\text{aposteriorní Dir}} \propto \underbrace{f(y_t|\psi_t, \Theta)}_{\text{alternativní/kategorický}} \underbrace{f(\Theta|d(t-1))}_{\text{apriorní Dir}}$$

Úprava diskrétního modelu

$$f(y_t|\psi_t, \Theta) = \prod_{\forall y|\psi} (\Theta_{y|\psi})^{\delta(y|\psi; y_t|\psi_t)}$$

Kroneckerovo delta

$$\delta = 1 \begin{cases} \text{číslo sloupce } y = y_t \\ \text{číslo řádku } \psi = \psi_t \end{cases}$$

$\delta = 0$ jinak

Update statistik: $\nu_{y|\psi;t} = \nu_{y|\psi;t-1} + \delta(y|\psi; y_t|\psi_t)$

Bodový odhad: $\hat{\Theta}_{y|\psi;t} = \frac{\nu_{y|\psi;t}}{\sum_{i=1}^n \nu_{i|\psi;t}}$

Algoritmus Bayesovského odhadování diskrétního modelu

- 1 Pro čas $t = 0$ nastavíme počáteční statistiku ν_0
- 2 Pro čas $t = 1, 2, \dots$
 - 1 Měříme data $d_t = \{y_t, u_t\}$
 - 2 Update statistik $\nu_{y|\psi;t} = \nu_{y|\psi;t-1} + \delta(y|\psi; y_t|\psi_t)$
 - 3 Jdeme na krok [2.1](#)

- 3 Výpočet bodových odhadů parametrů

$$\underbrace{\hat{\Theta}_{y|\psi;t}}_{\text{model}} = \frac{\nu_{y|\psi;t}}{\sum_{i=1}^n \nu_{i|\psi;t}}$$

Příklad:

System: úsek silnice

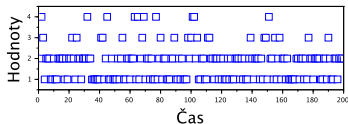
$y_t \in \{1, 2, 3, 4\}$ – pohon automobilu
(benzínový, naftový, hybridní, elektrický)

$v_t \in \{1, 2, 3\}$ – zkušenosti řidiče
(začátečník, pokročilý, profesionál)

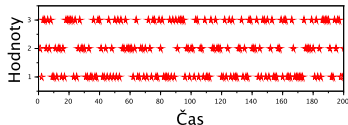
Čas t – minuty (cca 8h)

| | | | | | | | | | |
|-------|---|---|---|---|---|---|---|---|-----|
| y_t | 2 | 4 | 3 | 1 | 2 | 2 | 2 | 1 | ... |
| v_t | 2 | 1 | 3 | 3 | 2 | 3 | 2 | 3 | ... |

Typ pohonu automobilu



Zkušenosti řidiče



Algoritmus, kroky 1–2.3

| | | | | |
|-------------------------|----------------|----------------|----------------|----------------|
| $\psi_t \backslash y_t$ | 1 | 2 | 3 | 4 |
| $v_t = 1$ | $\Theta_{1 1}$ | $\Theta_{2 1}$ | $\Theta_{3 1}$ | $\Theta_{4 1}$ |
| $v_t = 2$ | $\Theta_{1 2}$ | $\Theta_{2 2}$ | $\Theta_{3 2}$ | $\Theta_{4 2}$ |
| $v_t = 3$ | $\Theta_{1 3}$ | $\Theta_{2 3}$ | $\Theta_{3 3}$ | $\Theta_{4 3}$ |

Diskrétní model $f(y_t | v_t, \Theta)$

$\psi_t = v_t$ – data

Θ – neznámé parametry

Algoritmus:

- $t = 0$, dimenze počáteční statistiky ν_0
 $\dim(\nu_0) = \dim(\text{model}) = 3 \times 4$
`zeros(3, 4)`, `ones(3, 4)`
- $t = 1, 2, \dots$
 - Měříme data: $y_1 = 2, v_1 = 2$
 - Update statistik
 $\nu_{y|\psi;1} = \nu_{y|\psi;0} + \delta(y|\psi; 2|2)$
 $\nu_{2|2;1} = \nu_{2|2;0} + 1$
 - Jdeme na krok [2.1](#) ...
 $\nu_{4|1;2} = \nu_{4|1;1} + 1$...
 $\nu_{3|3;3} = \nu_{3|3;2} + 1$...

Vývoj statistiky ν_t v čase

| | | | | |
|---------|---|-----|---|---|
| ν_0 | 0 | 0 | 0 | 0 |
| | 0 | 0 | 0 | 0 |
| | 0 | 0 | 0 | 0 |
| ν_1 | 0 | 0 | 0 | 0 |
| | 0 | 0+1 | 0 | 0 |
| | 0 | 0 | 0 | 0 |

| | | | | |
|---------|---|---|-----|-----|
| ν_2 | 0 | 0 | 0 | 0+1 |
| | 0 | 1 | 0 | 0 |
| | 0 | 0 | 0 | 0 |
| ν_3 | 0 | 0 | 0 | 1 |
| | 0 | 1 | 0 | 0 |
| | 0 | 0 | 0+1 | 0 |

Algoritmus, pokračování, kroky 2.3 – 3

$$t = 200, \quad \nu_{2|2;200} = \nu_{2|2;199} + 1$$

 ν_{200}

| | | | |
|----|----|---|---|
| 33 | 29 | 7 | 8 |
| 17 | 33 | 8 | 2 |
| 31 | 26 | 6 | 0 |

3 Bodové odhady parametrů $\hat{\Theta}_{y|\psi;t} = \frac{\nu_{y|\psi;t}}{\sum_{i=1}^n \nu_{i|\psi;t}}$

$$\hat{\Theta}_{1|1;t} = \frac{33}{33+29+7+8}, \quad \hat{\Theta}_{2|1;t} = \frac{29}{77}, \quad \hat{\Theta}_{3|1;t} = \frac{7}{77}, \quad \hat{\Theta}_{4|1;t} = \frac{8}{77}$$

$$\hat{\Theta}_{1|2;t} = \frac{17}{17+33+8+2}, \quad \hat{\Theta}_{2|2;t} = \frac{33}{60}, \quad \hat{\Theta}_{3|2;t} = \frac{8}{60}, \quad \hat{\Theta}_{4|2;t} = \frac{2}{60}$$

$$\hat{\Theta}_{1|3;t} = \frac{31}{31+26+6+0}, \quad \hat{\Theta}_{2|3;t} = \frac{26}{63}, \quad \hat{\Theta}_{3|3;t} = \frac{6}{63}, \quad \hat{\Theta}_{4|3;t} = \frac{0}{63}$$

Výsledek – model – podmíněná pf $f(y_t | v_t, \Theta)$

| $y_t \backslash \psi_t$ | 1 | 2 | 3 | 4 |
|-------------------------|-----------|-----------|-----------|-----------|
| $v_t = 1$ | 0.4285714 | 0.3766234 | 0.0909091 | 0.1038961 |
| $v_t = 2$ | 0.2833333 | 0.55 | 0.1333333 | 0.0333333 |
| $v_t = 3$ | 0.4920635 | 0.4126984 | 0.0952381 | 0 |

Bodový odhad výstupu (predikce) – maximální $\hat{\Theta}_{y|\psi;t}$ podle v_t

```
% nahrajeme data  
% počet dat  
nd=200;  
% počáteční statistiky  
ni=zeros(max(v),max(y));  
for t=1:nd  
    ni(v(t),y(t))=ni(v(t),y(t))+1; % update statistiky  
end  
% bodový odhad pravděpodobností  
theta_odhad=ni./ sum(ni,2);
```

Odhad logistické regrese s výstupem $y_t \in \{0, 1\}$

Logistický model

$$f(y_t | \psi_t, \theta) = \frac{\exp\{y_t z_t\}}{1 + \exp\{z_t\}}$$

Nemá ~~konjugované rozdělení~~
Bayesovo pravidlo není ~~rekurze~~

Bayesovo pravidlo

$$\underbrace{f(\theta | d(t))}_{\text{aposteriori}} \propto L_t(\theta) \underbrace{f(\theta | d(t-1))}_{\text{apriorní}}$$

Metoda maximální věrohodnosti

$$\ln L_t(\theta) = \ln \prod_{\tau=1}^t \frac{e^{y_\tau z_\tau}}{1 + e^{z_\tau}}$$

$$\hat{\theta} = \arg \max_{\theta} \ln L_t(\theta)$$

$$z_t = \psi_t' \theta + e_t$$

$$\psi_t = \underbrace{[v_{1;t} \ v_{2;t} \ v_{3;t} \ \dots \ 1]'}_{\text{data}}$$

$$\theta = \underbrace{[c_1 \ c_2 \ c_3 \ \dots \ k]'}_{\text{neznámé parametry}}$$

Věrohodnostní funkce

$$L_t(\theta) = \prod_{\tau=1}^t f(y_\tau | \psi_\tau, \theta)$$

Odhad – jednorázový výpočet (offline)

Výhody – známé derivace $\ln L_t(\theta)$
max – Newtonova metoda hledání extrémů

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$$

[Odkaz na web](#)

Příklad – klasifikace s logistickou regresí

System: úsek dálnice

$y_t \in \{0, 1\}$ – nehoda (NE, ANO)

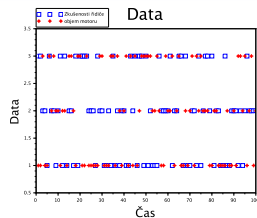
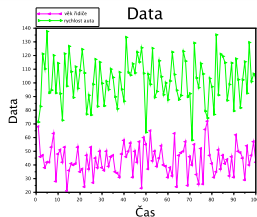
$v_{1;t}$ – věk řidiče

$v_{2;t} \in \{1, 2, 3\}$ – zkušenosti řidiče (začátečník, pokročilý, profesionál)

$v_{3;t} \in \{1, 2, 3\}$ – objem motoru auta (do 2 l; 2-3 l; 3 l a více)

$v_{4;t}$ – rychlost auta [km/h]

| $v_{1;t}$ | $v_{2;t}$ | $v_{3;t}$ | $v_{4;t}$ | y_t |
|-----------|-----------|-----------|-----------|-------|
| 42 | 3 | 3 | 92.676063 | 0 |
| 53 | 2 | 2 | 94.71603 | 1 |
| 63 | 2 | 3 | 120.06312 | 0 |
| 28 | 1 | 2 | 92.568545 | 0 |
| 48 | 2 | 2 | 114.32204 | 0 |
| ... | ... | ... | ... | ... |



Program

```
% nahrajeme data
% pro nominální data y
y=nominal(y);

% odhadování (trénovací data)
[B,dev,stats] = mnrfit(vL,yL)

% klasifikace (testovací data)
py = mnrvl(B,vT)

for i=1:size(py,1)
    [mx ix]=max(py(i,:)); % index max.
    yp(i)=ix-1; % otočíme na 0 a 1
end

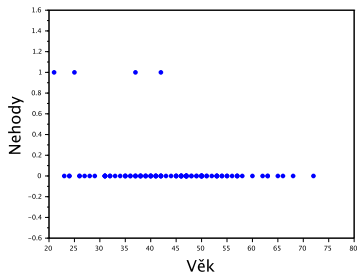
% Chyba predikce
PE=sum(categorical(yT)~=categorical(yp));
Acc=(PE*100) / size(yp,2);
```

Výsledky

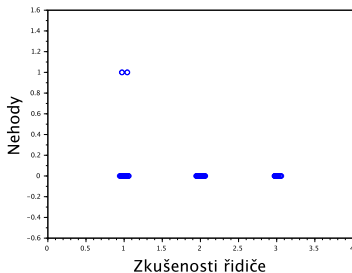
```
68. 3. 1. 71.374685 0.
46. 3. 1. 82.791157 0.
47. 2. 3. 121.15277 0.
38. 2. 1. 110.24375 0.
42. 1. 1. 137.60589 1.
42. 3. 3. 92.676063 0.
53. 2. 2. 94.71603 0.
63. 2. 3. 120.06312 0.
28. 1. 2. 92.568545 0.
48. 2. 2. 114.32204 0.
51. 3. 1. 92.129863 0.
42. 1. 2. 72.665487 0.
53. 3. 2. 121.30124 0.
21. 1. 1. 97.61611 1.
36. 2. 3. 126.78148 0.
```


Klasifikace s logistickou regresí

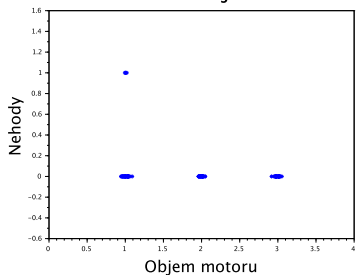
Klasifikace – věk



Klasifikace – zkušenosti řidiče



Klasifikace – objem motoru



Klasifikace – rychlost

