# Přednáška 2 – Bayesovský odhad jednotlivých modelů

**Obecný tvar**

$f(y_t|\Theta)$

**Neznámé parametry $\Theta$**

- $\{\theta, R\}$ – regresní • $\Theta$ – diskrétní • $\lambda$ – Poissonův

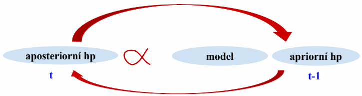
**Bayesovo pravidlo**

$$\underbrace{f(\Theta|y(t))}_{\text{aposteriorní}} \propto \underbrace{f(y_t|\Theta)}_{\text{model}}\, \underbrace{f(\Theta|y(t-1))}_{\text{apriorní}}$$

$y_t$ – data v čase $t$

$$\underbrace{y(t) = \{y_0, y_1, \ldots, y_t\}}_{\text{stará data } + \; y_t}$$

Rekurze – online odhad




- Reprodukující se (konjugovaná) apriorní hp (stejný tvar v čase $t$)

**Update statistik (rekurzivní přepočet):**

Normální model: $V_t = V_{t-1} + \begin{bmatrix} y_t \\ 1 \end{bmatrix} [y_t' \;\; 1], \quad \kappa_t = \kappa_{t-1} + 1$

Kategorický model: $\nu_{i;t} = \nu_{i;t-1} + \delta(i;\, y_t)$

Poissonův model: $S_t = S_{t-1} + y_t, \quad \kappa_t = \kappa_{t-1} + 1$

# Mnohorozměrný statický normální model

$$y_t = \theta + e_t$$

**Algoritmus odhadu mnohorozměrného normálního modelu:**

1. Pro čas $t = 0$ nastavíme počáteční informační matici $V_0$ a počítadlo $\kappa_0$
2. Pro čas $t = 1, 2, \ldots$
   1. Měříme data $y_t$
   2. Update statistik $V_t = V_{t-1} + \begin{bmatrix} y_t \\ 1 \end{bmatrix} [y_t' \;\; 1], \quad \kappa_t = \kappa_{t-1} + 1$
   3. Jdeme na krok 2.1
3. Rozklad informační matice $$V_t = \begin{bmatrix} V_{n\times n} & V_y' \\ V_y & V_n \end{bmatrix} = \left[\begin{array}{cc} \left[\begin{array}{cc} \cdots & \cdots \\ \cdots & \cdots \end{array}\right] & | \\ \underline{\qquad\quad} & . \end{array}\right]$$
4. Výpočet bodových odhadů parametrů $\hat{\theta}_t = V_n^{-1} V_y,$ $$\hat{R}_t = \frac{V_{n\times n} - V_y' V_n^{-1} V_y}{\kappa_t}$$

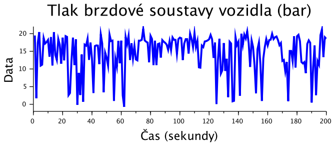
# Příklad – vozidlo na silnici

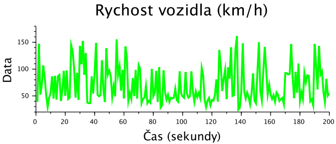
## Data:

$y_{1;t}$ – tlak v hydraulickém systému brzdové soustavy vozidla (bar)

$y_{2;t}$ – rychlost vozidla (km/h)

$t$ – sekundy




| | | | | | |
|---|---|---|---|---|---|
| $y_{1;t}$ | 19.48 | 1.71 | 17.38 | 20.44 | ... |
| $y_{2;t}$ | 38.85 | 147.11 | 52.67 | 70.7 | ... |

**2-rozměrný statický model:**

$$\begin{bmatrix} y_{1;t} \\ y_{2;t} \end{bmatrix} = \underbrace{\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}}_{\theta} + \begin{bmatrix} e_{1;t} \\ e_{2;t} \end{bmatrix}$$

$\theta$, $R$ – neznámé parametry

## Algoritmus:

1. $t = 0$, $V_0$ – dimenze informační matice

   $V_0$ – symetrická matice $m \times m$

   $m = dim(y_t) + 1 = 3$

   $dim(V_0) = 3 \times 3$

   zeros$(3, 3)$, $\quad 0.001 * \text{eye}(3, 3)$

   $\kappa_0 = 0$

# Algoritmus – pokračování, kroky 2.1–2.3

2. $t = 1, 2, \ldots$
   1. Měříme data: $y_{1;t} = 19.48$, $y_{2;t} = 38.85$
   2. Update statistik $V_1 = V_0 + \begin{bmatrix} y_{1;1} \\ y_{2;1} \\ 1 \end{bmatrix} [y_{1;1} \;\; y_{2;1} \;\; 1] =$

      $$= \underbrace{V_0 + \begin{bmatrix} 19.48 \\ 38.85 \\ 1 \end{bmatrix} [19.48 \;\; 38.85 \;\; 1],}_{\text{symetrická matice}(3\times 3)} \quad \kappa_1 = 0 + 1$$

   3. Jdeme na krok 2.1

      $$V_2 = V_1 + \begin{bmatrix} y_{1;2} \\ y_{2;2} \\ 1 \end{bmatrix} [y_{1;2} \;\; y_{2;2} \;\; 1], \quad \kappa_2 = 1 + 1 \ldots\ldots\ldots$$

      $$V_3 = V_2 + \begin{bmatrix} y_{1;3} \\ y_{2;3} \\ 1 \end{bmatrix} [y_{1;3} \;\; y_{2;3} \;\; 1], \quad \kappa_3 = 2 + 1 \ldots\ldots\ldots$$

      $$V_{800} = V_{799} + \begin{bmatrix} y_{1;800} \\ y_{2;800} \\ 1 \end{bmatrix} [y_{1;800} \;\; y_{2;800} \;\; 1], \quad \kappa_{800} = 799 + 1 \quad \ldots$$

# Algoritmus – pokračování, kroky 3–4

3. Rozklad informační matice

$$V_{800} = \begin{bmatrix} 193508.43 & 692639.56 & 11702.588 \\ 692639.56 & 4824915.7 & 56139.732 \\ 11702.588 & 56139.732 & 800.0001 \end{bmatrix}$$

$$= \begin{bmatrix} V_{nn} & V_y' \\ V_y & V_n \end{bmatrix} = \left[ \begin{array}{c|c} \begin{bmatrix} \cdots & \cdots \\ \cdots & \cdots \end{bmatrix} & \\ \hline & . \end{array} \right], \quad \kappa_{800} = 800 \quad \rightarrow f(\Theta|d(t))$$

4. Bodové odhady parametrů

$$\hat{\theta}_{800} = V_n^{-1} V_y = \begin{bmatrix} \hat{\theta}_{1;800} & \hat{\theta}_{2;800} \end{bmatrix} = \begin{bmatrix} 14.63 & 70.18 \end{bmatrix}$$

$$\hat{R}_t = \frac{V_{n\times n} - V_y' V_n^{-1} V_y}{\kappa_t} = \begin{bmatrix} 27.900302 & -160.73194 \\ -160.73194 & 1106.6616 \end{bmatrix}$$

**Poznámka**

- Význam bodových odhadů – střední hodnoty – bodový odhad $\hat{y}_t$
- Online/offline odhad

# Algoritmus Bayesovského odhadu kategorického modelu

| $y_t$ | 1 | 2 | ... | N |
|---|---|---|---|---|
| $\Theta$ | $\Theta_1$ | $\Theta_2$ | ... | $\Theta_N$ |

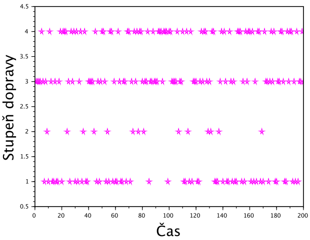
1. Pro čas $t = 0$ nastavíme počáteční statistiku $\nu_0$
2. Pro čas $t = 1, 2, \ldots$
   1. Měříme data $y_t$
   2. Update statistik $\nu_{i;t} = \nu_{i;t-1} + \delta(i;\, y_t)$
   3. Jdeme na krok 2.1
3. Výpočet bodových odhadů parametrů $\underbrace{\hat{\Theta}_{i;t}}_{\text{model}} = \frac{\nu_{i;t}}{\sum_{j=1}^{n} \nu_{j;t}}$

Příklad: úsek silnice

$y_t \in \{1, 2, 3, 4\}$ – stupeň dopravy (plynulá jízda až dopravní kolaps)

Čas $t$ – 5 minut

| $y_t$ | 3 | 4 | 3 | 1 | 4 | 3 | 3 | ... |
|---|---|---|---|---|---|---|---|---|



# $\Theta$ – neznámé parametry

1. $t = 0$, dimenze počáteční statistiky
   $\dim(\nu_0) = \dim(\text{model}) = 1 \times 4$
   zeros$(1,4)$, ones$(1,4)$
2. $t = 1, 2, \ldots$
   1. Měříme data: $y_t = 3$
   2. Update statistik
      $\nu_{i;t} = \nu_{i;t-1} + \delta(i;\, 3)$
      $\nu_{3;t} = \nu_{3;t-1} + 1$
   3. Jdeme na krok 2.1:
      $y_t = 4$
      $\nu_{4;t} = \nu_{4;t-1} + 1$
      ...
3. Bodové odhady parametrů $\hat{\Theta}_{i;t} = \frac{\nu_{i;t}}{\sum_{j=1}^{n} \nu_{j;t}}$

$\hat{\Theta}_{1;t} = \frac{46}{46+96+305+53}$, $\hat{\Theta}_{2;t} = \frac{96}{500}$, $\hat{\Theta}_{3;t} = \frac{305}{500}$, $\hat{\Theta}_{4;t} = \frac{53}{500}$

## Model

| $y_t$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\Theta$ | $\Theta_1$ | $\Theta_2$ | $\Theta_3$ | $\Theta_4$ |

## Vývoj statistiky $\nu_t$ v čase

| | | | | |
|---|---|---|---|---|
| $\nu_0 =$ | 0 | 0 | 0 | 0 |
| $\nu_1 =$ | 0 | 0 | 0+1 | 0 |
| $\nu_2 =$ | 0 | 0 | 1 | 0+1 |
| ... | | | | |
| $\nu_{500} =$ | 46 | 96 | 305 | 53 |

| $y_t$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\Theta$ | 0.092 | 0.192 | 0.61 | 0.106 |

- Význam bodových odhadů
- Online/offline odhad

# Poissonův model pro sčítací data

$f(y_t|\lambda) = \exp\{-\lambda\}\frac{\lambda^{y_t}}{y_t!}$

## Algoritmus odhadu Poissonova modelu:

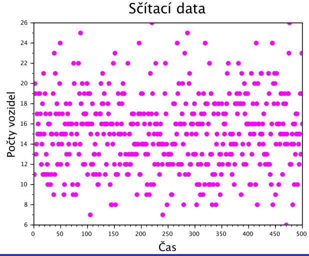
1. Pro čas $t = 0$ nastavíme počáteční statistiku $S_0$ a počítadlo $\kappa_0$
2. Pro čas $t = 1, 2, \ldots$
   1. Měříme data $y_t$
   2. Update statistik $S_t = S_{t-1} + y_t, \quad \kappa_t = \kappa_{t-1} + 1$
   3. Jdeme na krok 2.1
3. Výpočet bodových odhadů parametrů $\hat{\lambda}_t = \frac{S_t}{\kappa_t}$

Příklad: úsek silnice

$y_t$ – počty vozidel za jednotku času

Čas $t$ – minuty

| $y_t$ | 11 | 20 | 19 | 13 | 14 | 15 | ... |
|---|---|---|---|---|---|---|---|



1. $t = 0$, počáteční statistiky
   $S_0 = 0$, $\kappa_0 = 0$
2. $t = 1, 2, \ldots$
   1. Měříme data: $y_t = 11$
   2. Update statistik
      $S_t = 0 + 11$, $\kappa_t = 0 + 1$
   3. Jdeme na krok 2.1:
      $y_t = 20$
      $S_t = 11 + 20$, $\kappa_t = 1 + 1$
      $\ldots$
3. Bodové odhady parametrů
   $\hat{\lambda}_t = \frac{S_t}{\kappa_t}$
   $\hat{\lambda}_{500} = \frac{7519}{500} = 15.038$

**Poznámka**

- Význam bodových odhadů
  – bodový odhad $\hat{y}_t$
- Online/offline odhad