

Analýza a klasifikace dat – přednáška 2



RNDr. Eva Janoušová

Podzim 2014

Typy klasifikátorů – podle reprezentace vstupních dat

1. Podle reprezentace vstupních dat:

- **příznakové klasifikátory:** paralelní (s pevným počtem proměnných) x sekvenční
- **strukturální (syntaktické) klasifikátory**
- **kombinované klasifikátory**

2. Podle jednoznačnosti zařazení do skupin:

- **deterministické klasifikátory**
- **pravděpodobnostní klasifikátory**

3. Podle typů klasifikačních a učících algoritmů:

- **parametrické klasifikátory**
- **neparametrické klasifikátory**

4. Podle způsobu učení:

- **učení s učitelem:** dokonalým x nedokonalým
- **učení bez učitele**

5. Podle principu klasifikace:

- **klasifikace pomocí diskriminačních funkcí**
- **klasifikace pomocí vzdálenosti od etalonů klasifikačních tříd**
- **klasifikace pomocí hranic v obrazovém prostoru**

Typy klasifikátorů – podle reprezentace vstupních dat

1. Podle reprezentace vstupních dat:

- **příznakové klasifikátory: paralelní** (s pevným počtem proměnných) x sekvenční
- strukturální (syntaktické) klasifikátory
- kombinované klasifikátory

2. Podle jednoznačnosti zařazení do skupin:

- **deterministické klasifikátory**
- pravděpodobnostní klasifikátory

→ budeme se zabývat paralelním příznakovým deterministickým klasifikátorům

3. Podle typů klasifikačních a učících algoritmů:

- parametrické klasifikátory
- neparametrické klasifikátory

4. Podle způsobu učení:

- učení s učitelem: dokonalým x nedokonalým
- učení bez učitele

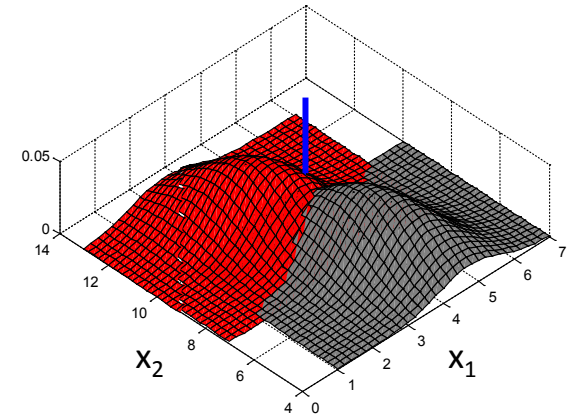
5. Podle principu klasifikace:

- klasifikace pomocí diskriminačních funkcí
- klasifikace pomocí vzdálenosti od etalonů klasifikačních tříd
- klasifikace pomocí hranic v obrazovém prostoru

Typy klasifikátorů – podle principu klasifikace

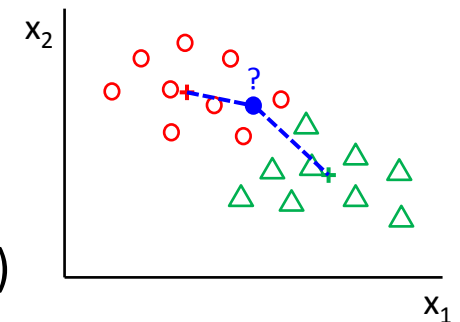
- **klasifikace pomocí diskriminačních funkcí:**

- diskriminační funkce určují míru příslušnosti k dané klasifikační třídě
- pro danou třídu má daná diskriminační funkce nejvyšší hodnotu



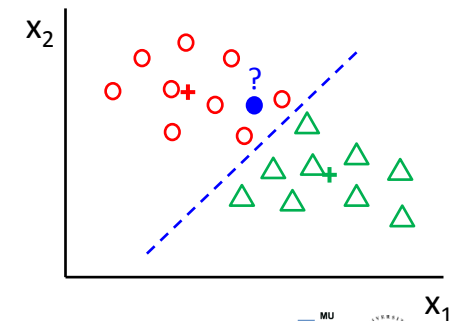
- **klasifikace pomocí vzdálenosti od etalonů klasif. tříd:**

- etalon = reprezentativní objekt(y) klasifikační třídy
- počet etalonů klasif. třídy různý – od jednoho vzorku (např. centroidu) po úplný výčet všech objektů dané třídy (např. u klasif. pomocí metody průměrné vazby)



- **klasifikace pomocí hranic v obrazovém prostoru:**

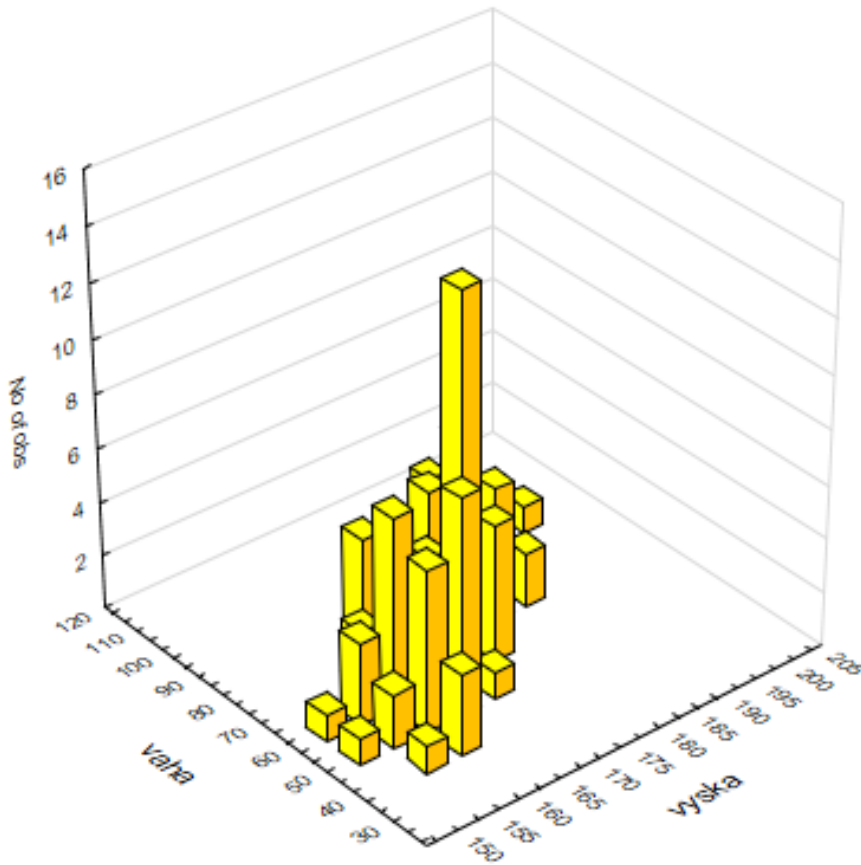
- stanovení hranic (hraničních ploch) oddělujících klasifikační třídy



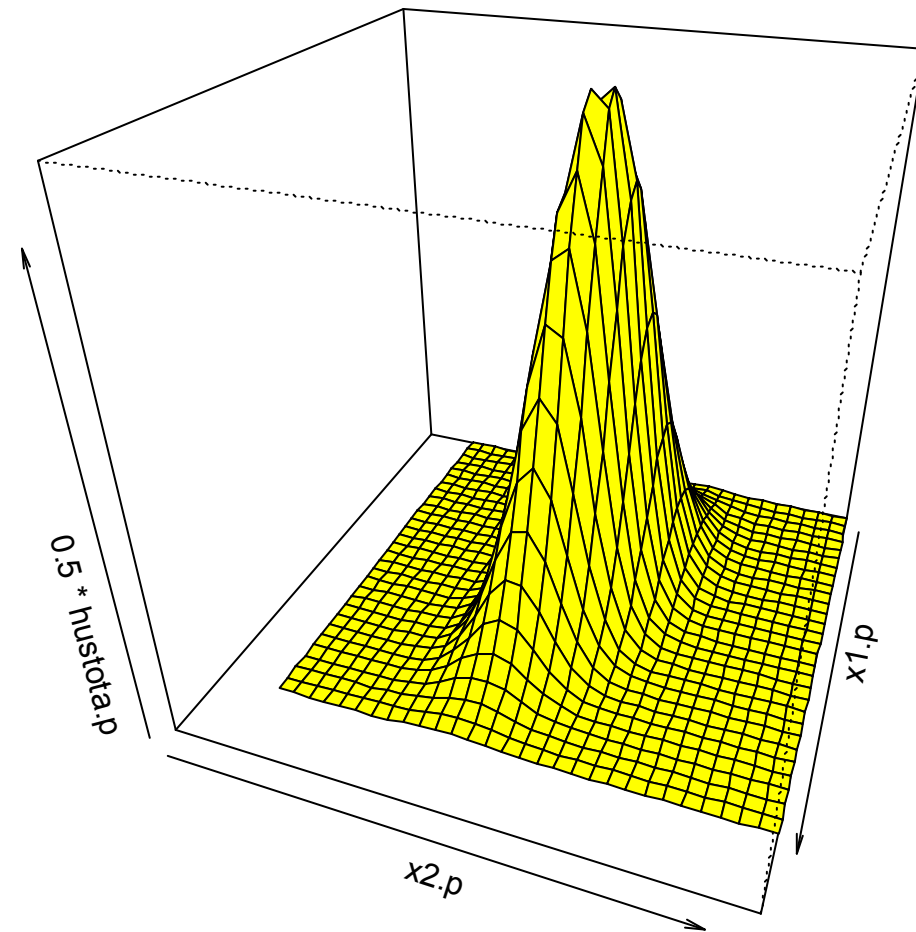
Vícerozměrné normální rozdělení

Motivace

Dvourozměrný histogram



Hustota dvourozměrného normálního rozdělení



Vícerozměrné normální rozdělení

Hustota vícerozměrného normálního rozdělení:

$$f_{\mathbf{x}}(x_1, \dots, x_k) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \cdot \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu)\right)$$

μ - střední hodnota

Σ - kovarianční matice

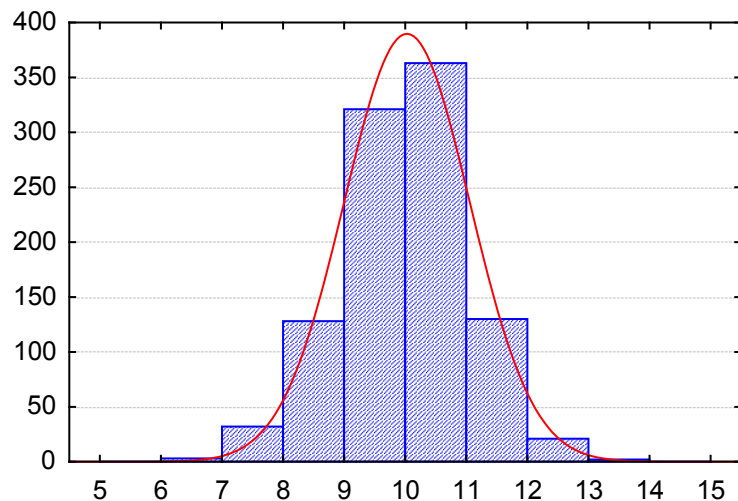
Dvourozměrné normálního rozdělení:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} \right]\right),$$

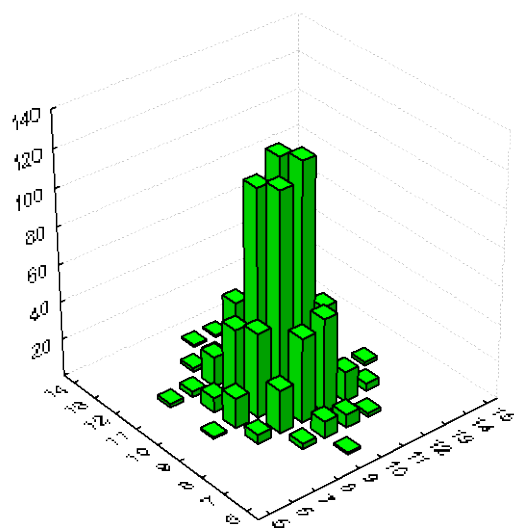
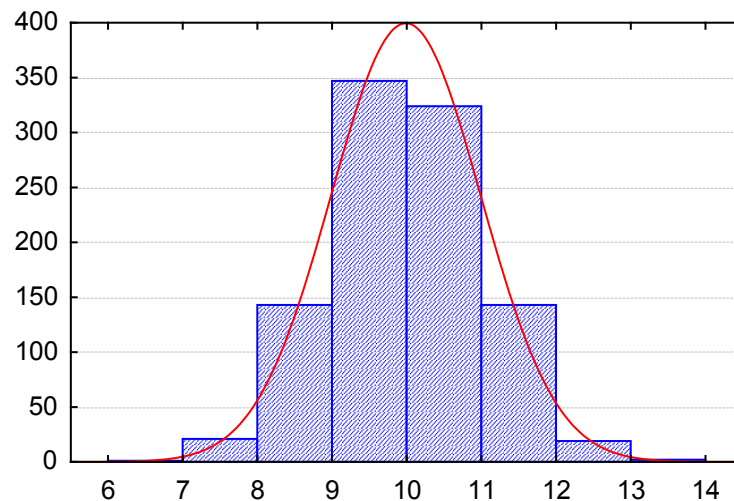
$$\mu = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$$

ρ - korelace mezi X a Y; σ – směrodatná odchylka

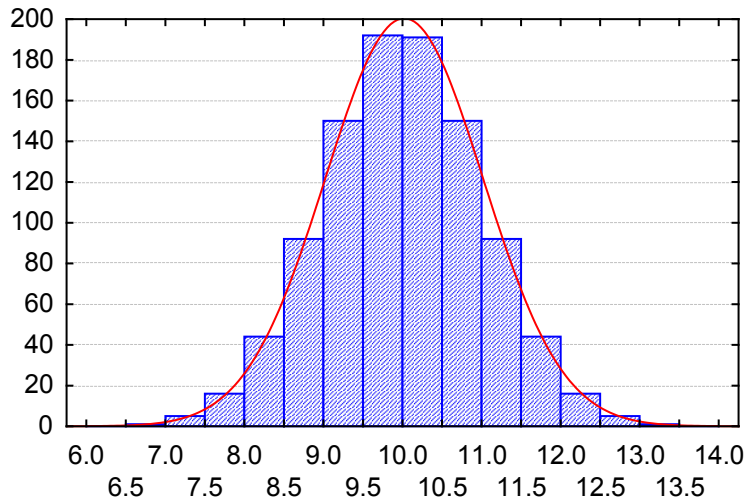
Je normalita v jednorozměrném prostoru jedinou podmínkou vícerozměrné normality?



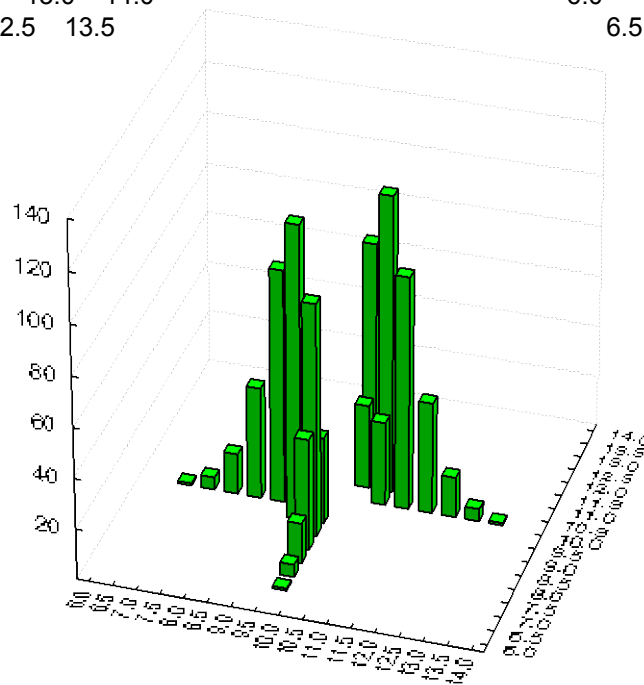
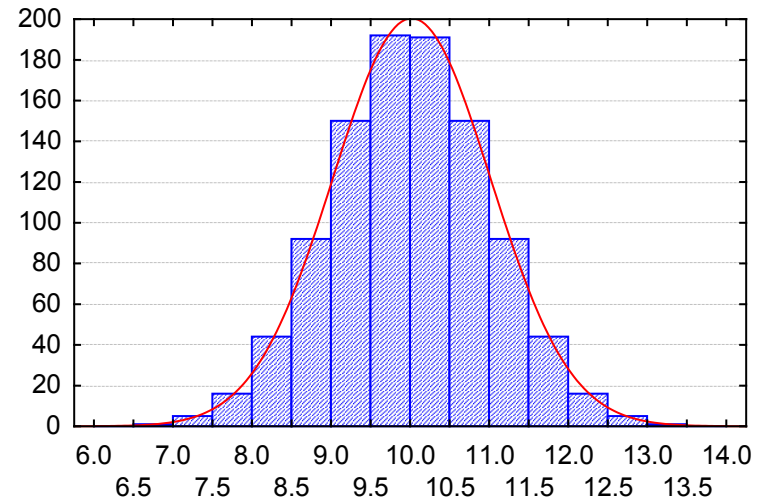
+



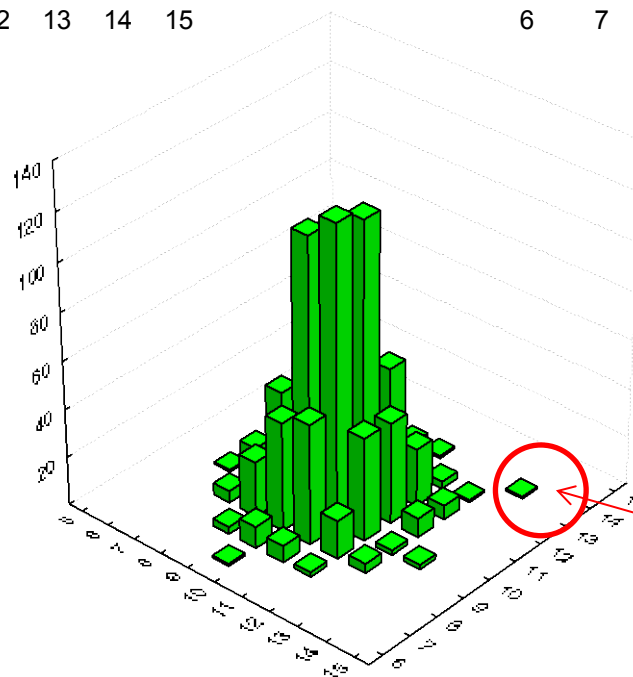
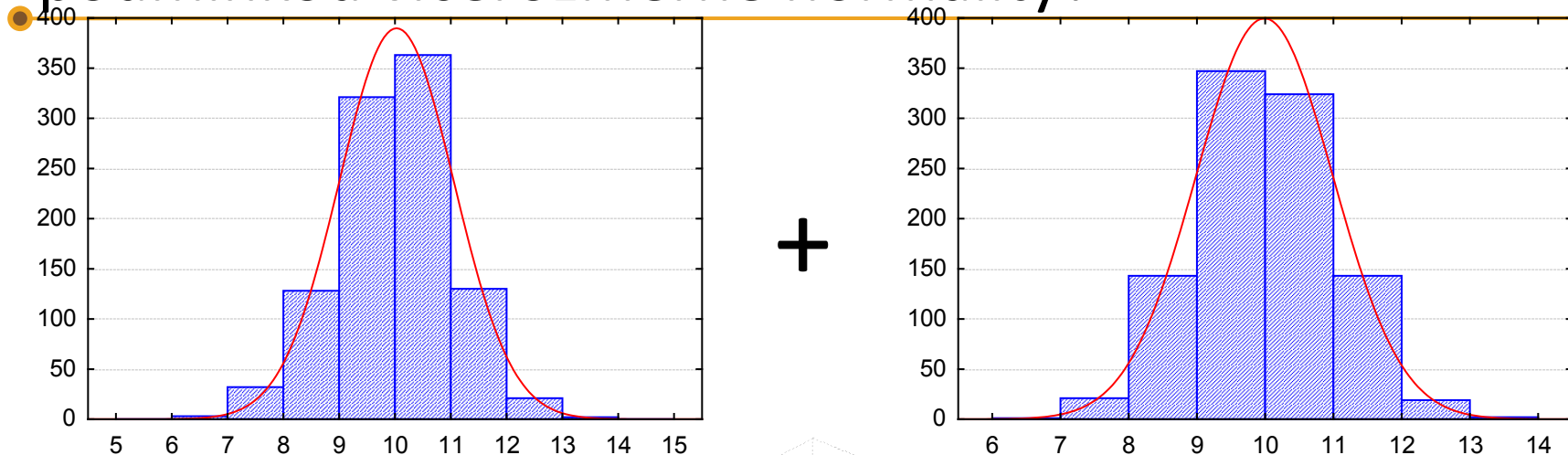
Je normalita v jednorozměrném prostoru jedinou podmínkou vícerozměrné normality?



+



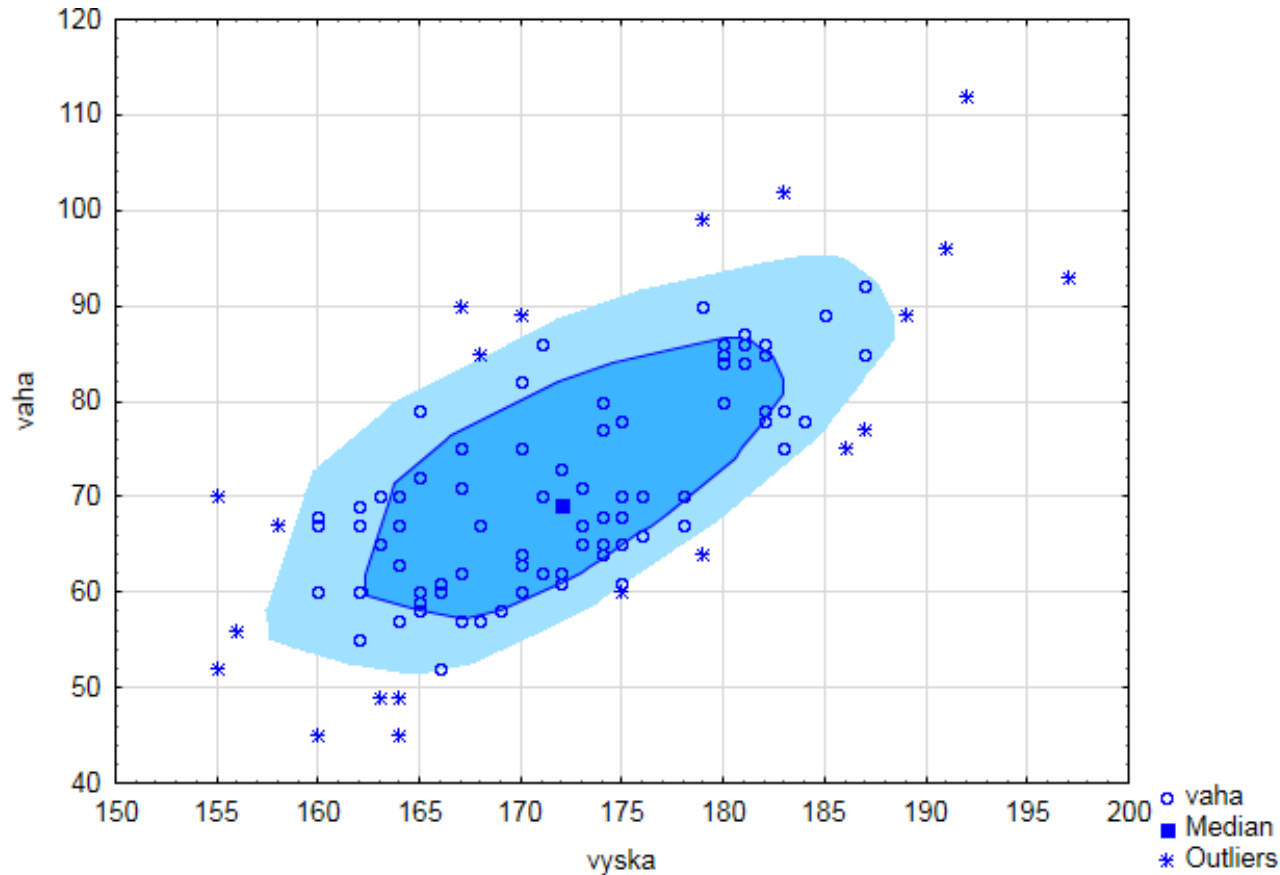
Je normalita v jednorozměrném prostoru jedinou podmínkou vícerozměrné normality?



Vícerozměrný
outlier

Ověření dvourozměrné normality

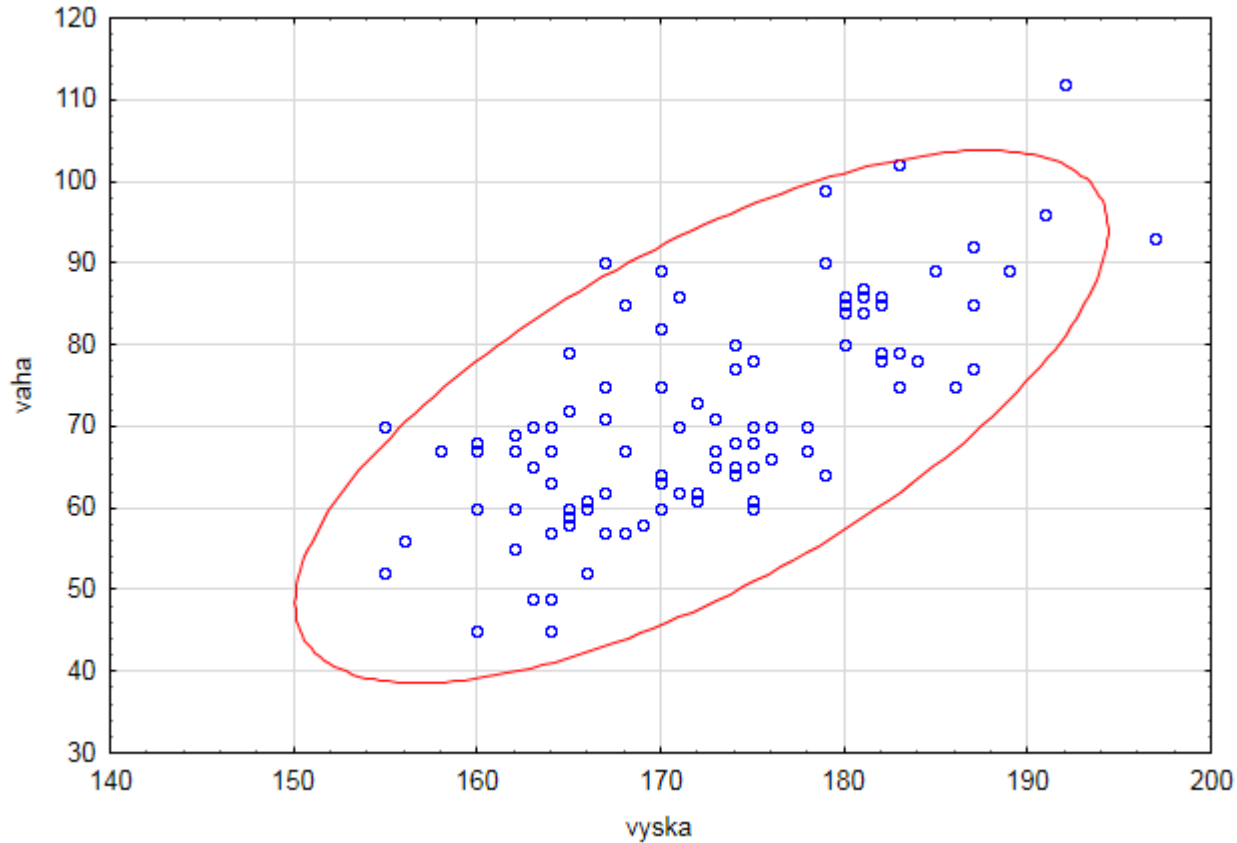
Bagplot = „bivariate boxplot“ (tzn. „dvourozměrný krabicový graf“)



v softwaru Statistica: Graphs – 2D Graphs – Bag Plots

Ověření dvourozměrné normality

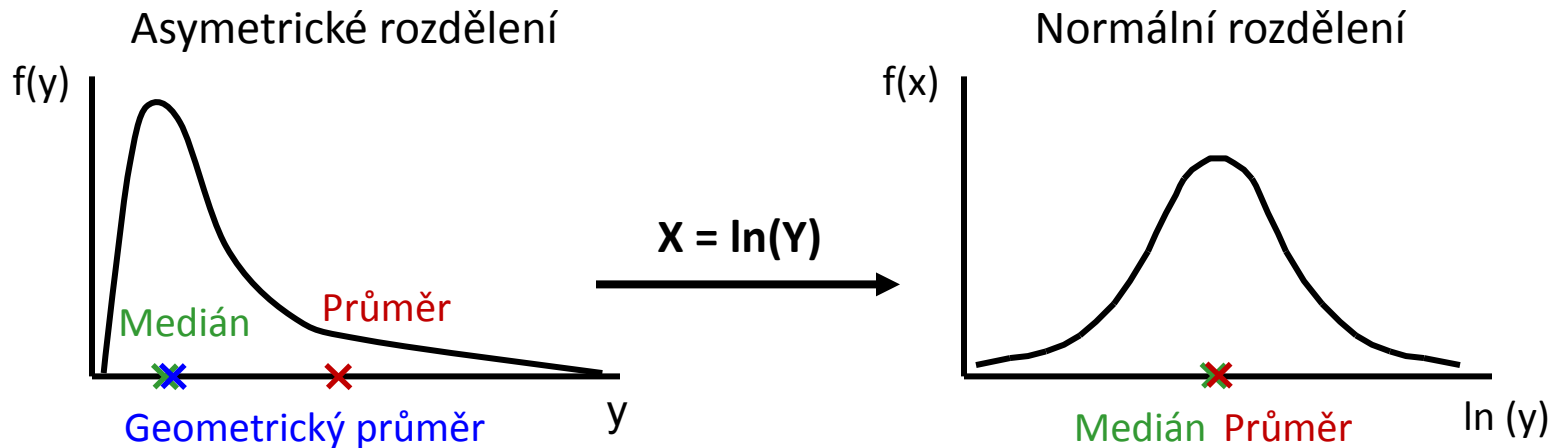
Vykreslení regulační elipsy („control“ ellipse):



v softwaru Statistica: Graphs – Scatterplots – na záložce Advanced zvolit Elipse Normal

Normalizace dat

- Převod na normální rozdělení (normalita je předpokladem řady statistických testů).
- Např. **logaritmická transformace**: $X = \ln(Y)$ nebo $X = \ln(Y+1)$, pokud data obsahují hodnotu 0



- Další příklady:
 - **odmocninová transf.** (pro proměnné s Poissonovým rozložením nebo obecně data typu počet jedinců, buněk apod.: $X = \sqrt{Y}$ nebo $X = \sqrt{Y + 1}$)
 - **arcsin transformace** (pro proměnné s binomickým rozložením)
 - **Box-Coxova transformace**

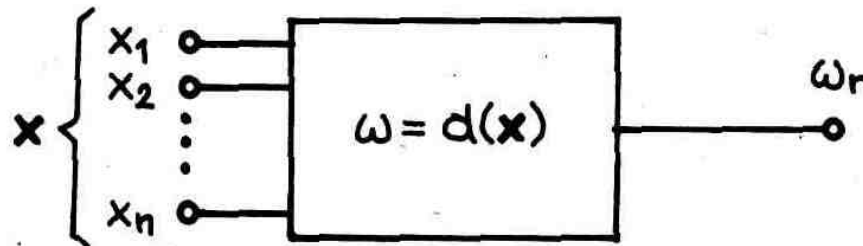
Klasifikace pomocí diskriminačních funkcí

Příznakový klasifikátor

- klasifikátor je algoritmus (stroj, *machine*) s tolika vstupy, kolik je proměnných (příznaků) a s jedním diskretním výstupem, který udává třídu, do které klasifikátor zařadil rozpoznávaný objekt:

$$\omega_r = d(\mathbf{x})$$

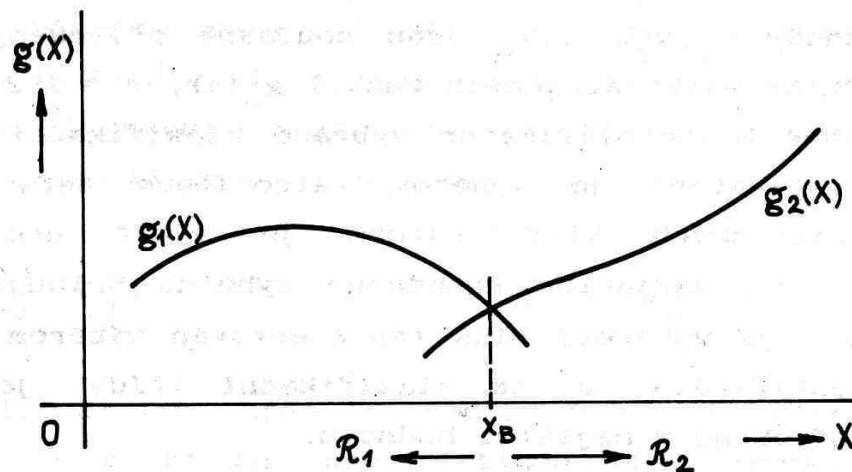
- \mathbf{x} je vektor hodnot jednotlivých proměnných pro daný subjekt:
 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$
- $d(\mathbf{x})$ je skalární funkce vektorového argumentu \mathbf{x} , kterou nazýváme **rozhodovací pravidlo klasifikátoru**
- ω_r je **identifikátor klasifikační třídy**



Klasifikace pomocí diskriminačních funkcí

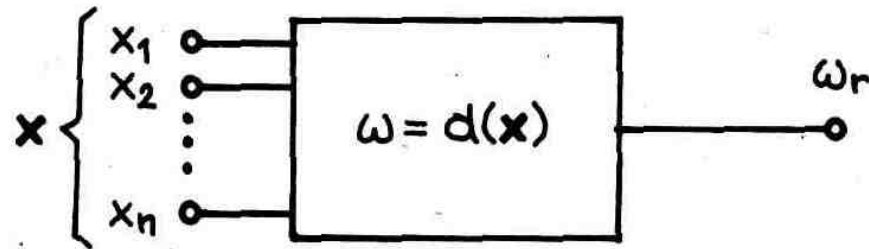
- rozhodovací pravidlo klasifikátoru $d(\mathbf{x})$ založeno na výpočtu diskriminačních funkcí
- **diskriminační funkce** – vyjadřují míru příslušnosti \mathbf{x} do jednotlivých klasifikačních tříd
- zařadíme \mathbf{x} do takové třídy ω_i , pro kterou je $g_i(\mathbf{x})$ maximální
- matematicky: pro obraz \mathbf{x} z podmnožiny \mathcal{R}_r pro všechna r platí

$$g_r(\mathbf{x}) > g_s(\mathbf{x}), \text{ pro } s = 1, 2, \dots, R \text{ a } r \neq s$$

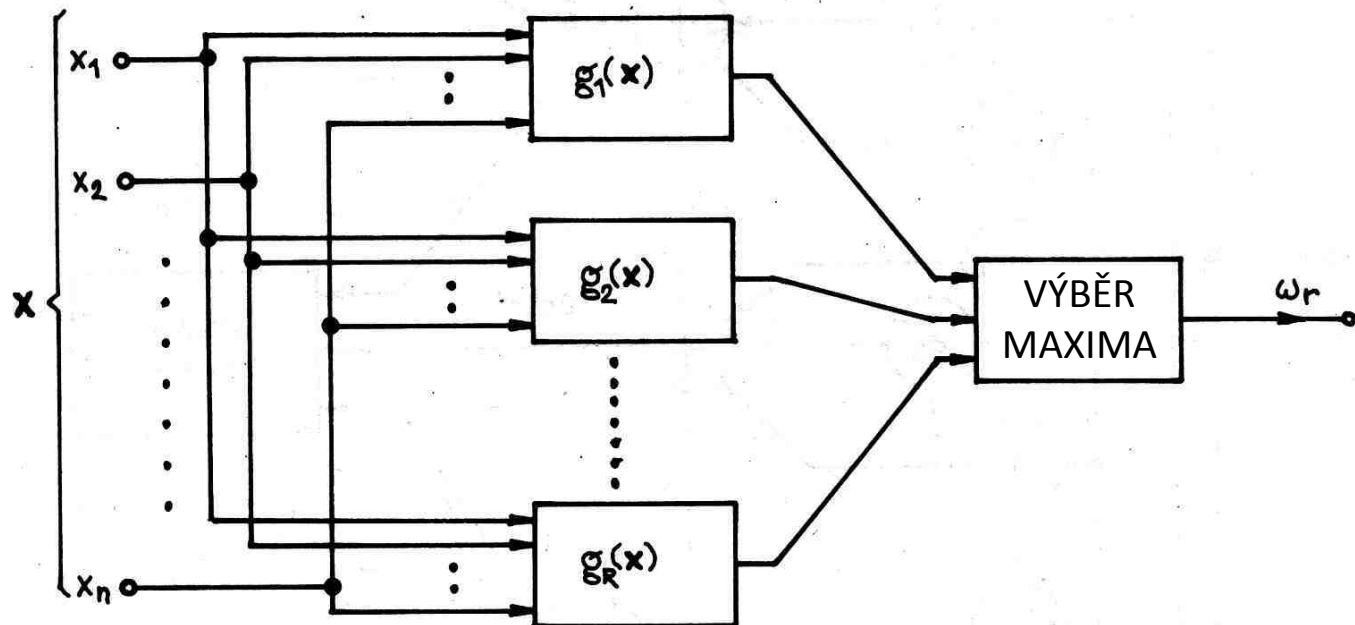


Blokové schéma klasifikátoru

- obecné blokové schéma klasifikátoru:



- blokové schéma klasifikátoru pomocí diskriminačních funkcí:

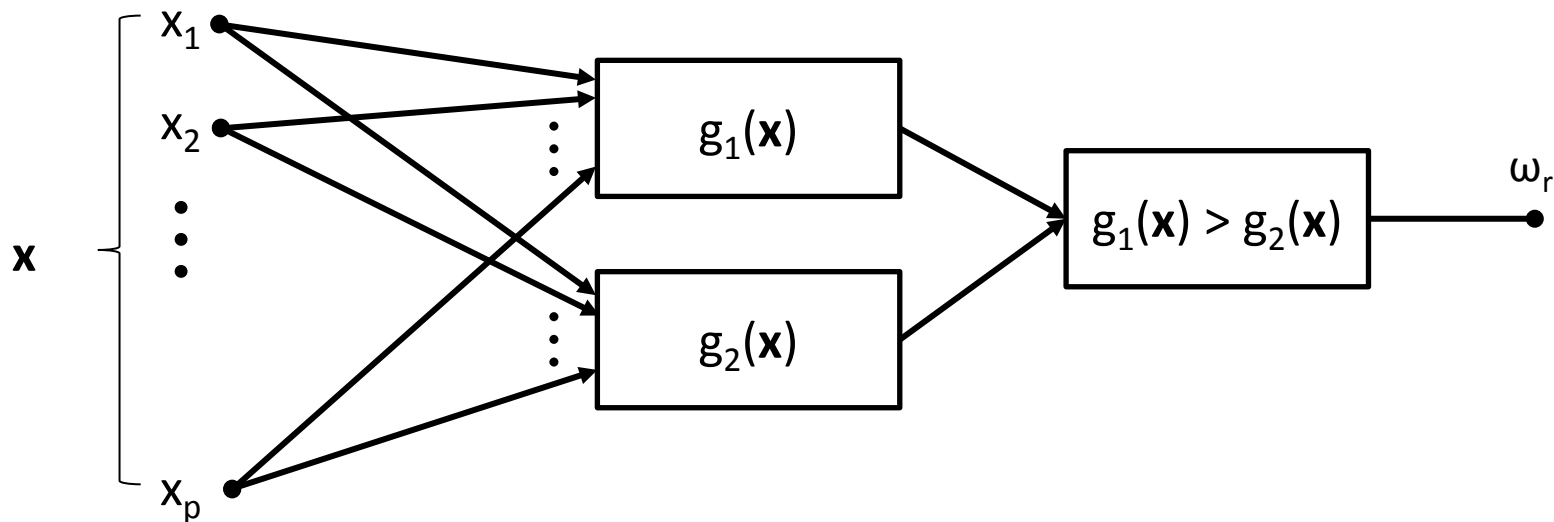


Dichotomický klasifikátor

- pro klasifikaci do dvou tříd
- rozhodovací pravidlo klasifikátoru lze zapsat jako:

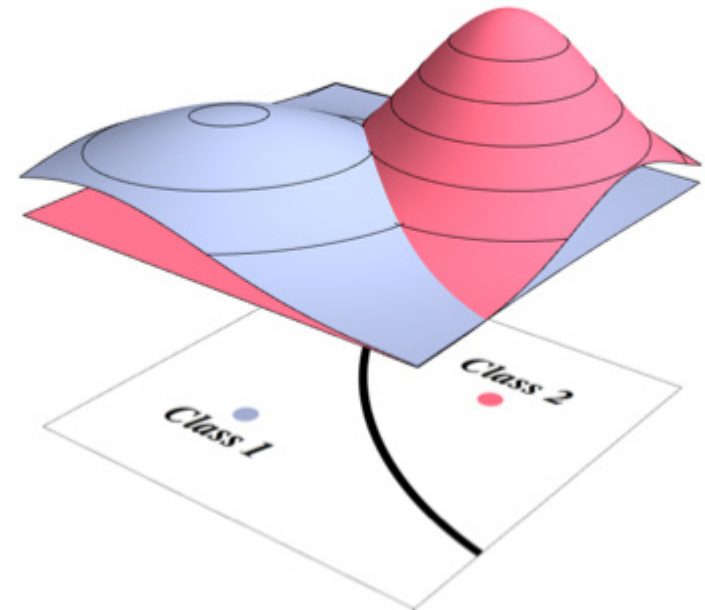
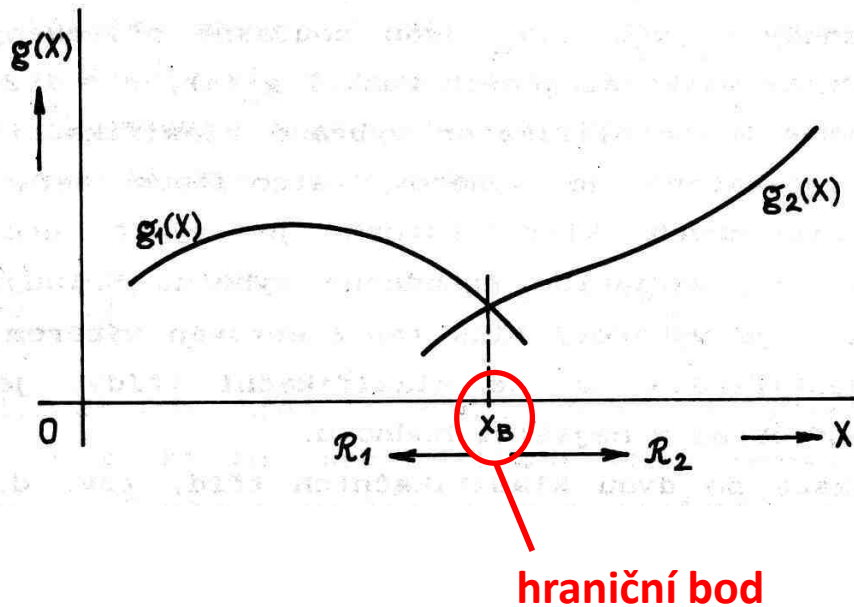
$$\omega_r = d(\mathbf{x}) = \text{sign}(g_1(\mathbf{x}) - g_2(\mathbf{x}))$$

- pokud $d(\mathbf{x}) \geq 0 \rightarrow$ zařazení \mathbf{x} do třídy ω_1
- pokud $d(\mathbf{x}) < 0 \rightarrow$ zařazení \mathbf{x} do třídy ω_2



Souvislost klasifikace pomocí diskriminačních funkcí s klasifikací pomocí hranic

- hranice mezi dvěma sousedními podmnožinami \mathcal{R}_r a \mathcal{R}_s je určena průmětem průsečíku funkcí $g_r(\mathbf{x})$ a $g_s(\mathbf{x})$, definovaného rovnicí $g_r(\mathbf{x}) = g_s(\mathbf{x})$, do obrazového prostoru



Příklady diskriminačních funkcí

- nejjednodušším tvarem diskriminační funkce je lineární funkce:

$$g_r(\mathbf{x}) = a_{r0} + a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rp}x_p$$

- diskriminační funkce na základě statistických vlastností množiny objektů:

$$g_r(\mathbf{x}) = P(\omega_r|\mathbf{x})$$

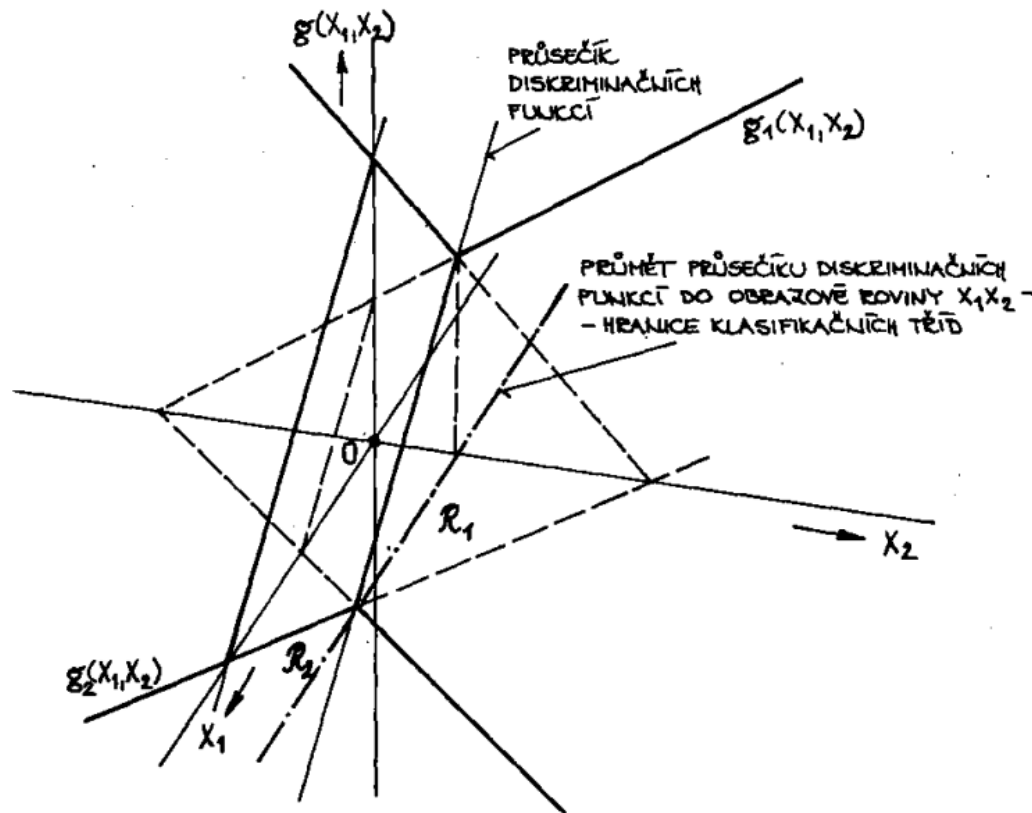
kde $P(\omega_r|\mathbf{x})$ je pravděpodobnost zatřídění \mathbf{x} do třídy ω_r

→ **Bayesův klasifikátor**

Lineární diskriminační funkce

$$g_r(\mathbf{x}) = a_{r0} + a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rp}x_p$$

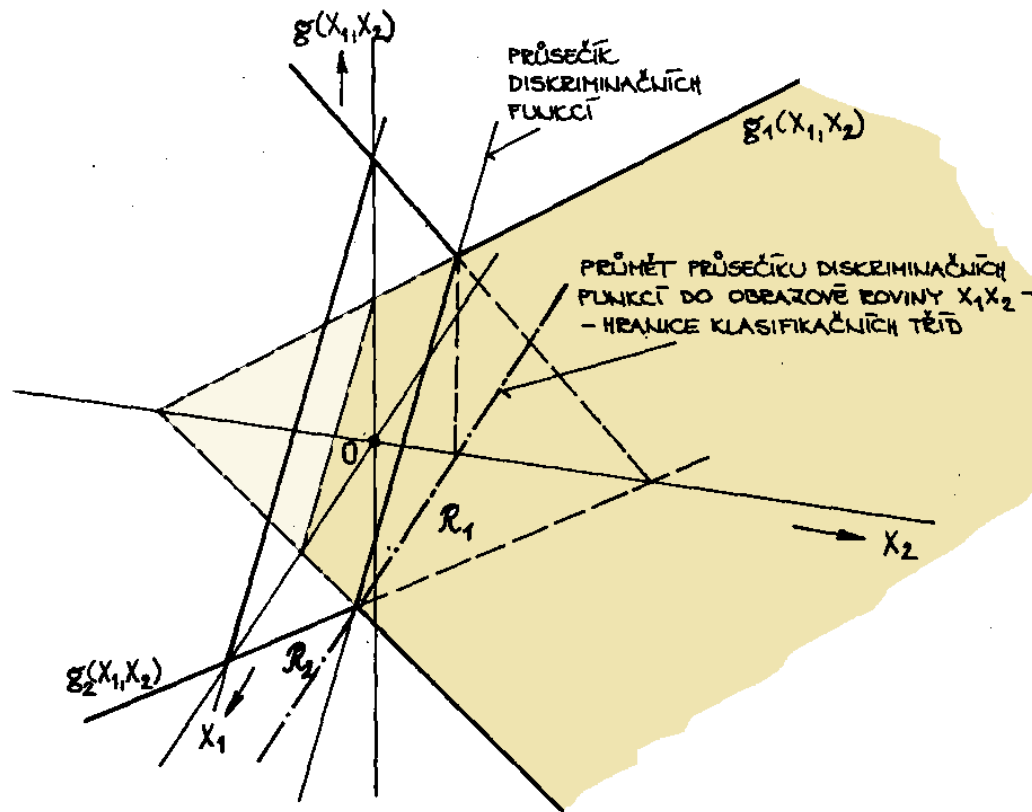
kde a_{r0} je práh diskriminační funkce posouvající počátek souřadného systému a a_{ri} jsou váhové koeficienty i -tého příznaku x_i



Lineární diskriminační funkce

$$g_r(\mathbf{x}) = a_{r0} + a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rp}x_p$$

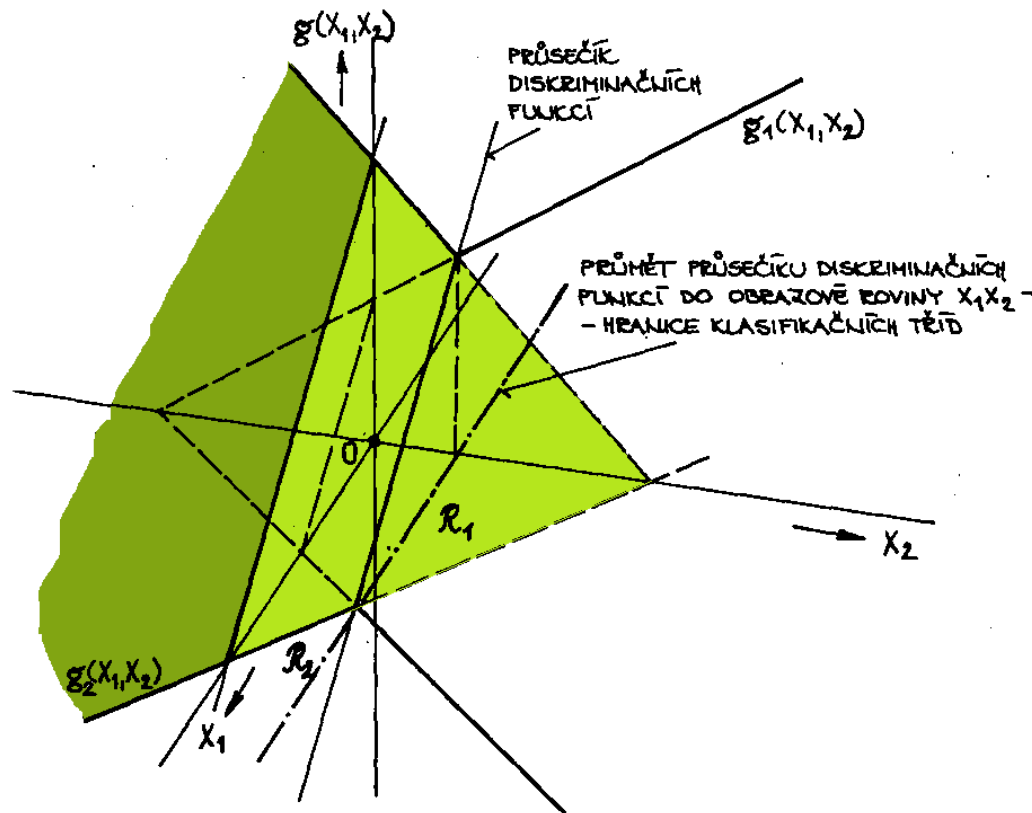
kde a_{r0} je práh diskriminační funkce posouvající počátek souřadného systému a a_{ri} jsou váhové koeficienty i -tého příznaku x_i



Lineární diskriminační funkce

$$g_r(\mathbf{x}) = a_{r0} + a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rp}x_p$$

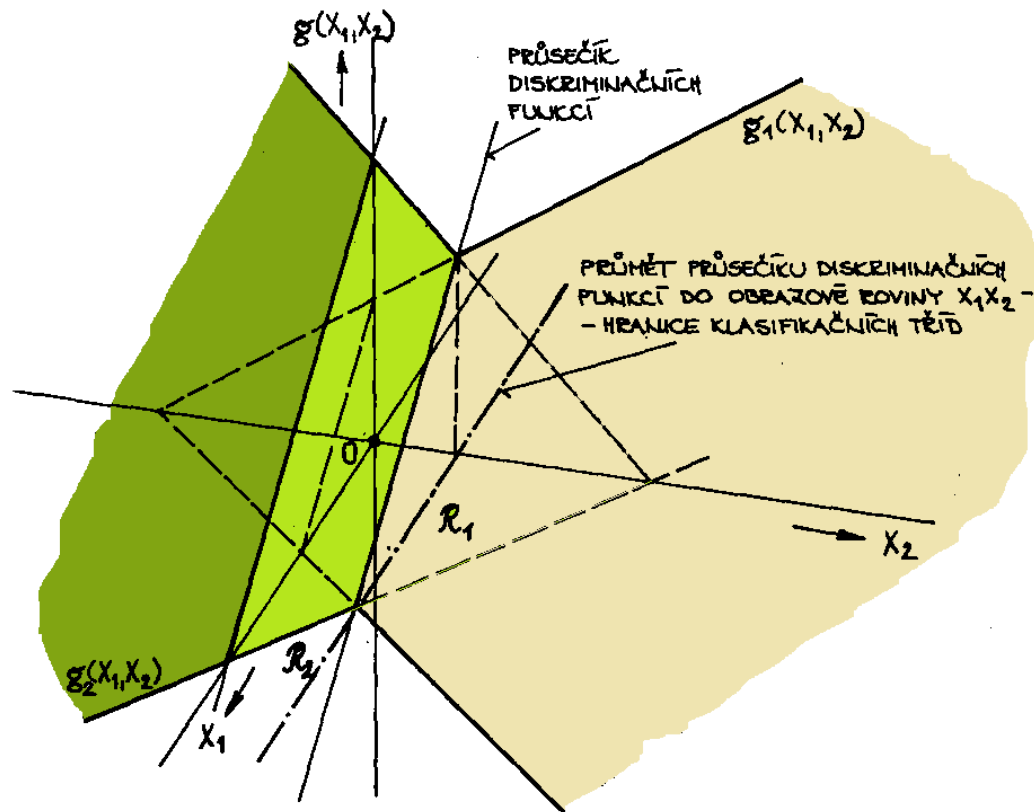
kde a_{r0} je práh diskriminační funkce posouvající počátek souřadného systému a a_{ri} jsou váhové koeficienty i -tého příznaku x_i



Lineární diskriminační funkce

$$g_r(\mathbf{x}) = a_{r0} + a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rp}x_p$$

kde a_{r0} je práh diskriminační funkce posouvající počátek souřadného systému a a_{ri} jsou váhové koeficienty i -tého příznaku x_i



Bayesův klasifikátor

- diskriminační funkce určeny na základě statistických vlastností množiny obrazů
- vyjdeme z **Bayesova vzorce**: $P(\omega_k|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_k) \cdot P(\omega_k)}{p(\mathbf{x})}$, kde
 - $P(\omega_k|\mathbf{x})$ je aposteriorní podmíněná pravděpodobnost zatřídění obrazu \mathbf{x} do třídy ω_r
 - $p(\mathbf{x}|\omega_k)$ je podmíněná hustota pravděpodobnosti výskytu obrazu \mathbf{x} ve třídě ω_k , $k = 1, 2$
 - $P(\omega_k)$ je apriorní pravděpodobnost třídy ω_k
 - $p(\mathbf{x})$ je celková hustota pravděpodobnosti rozložení obrazu \mathbf{x} v celém obrazovém prostoru

Bayesův klasifikátor – kritéria

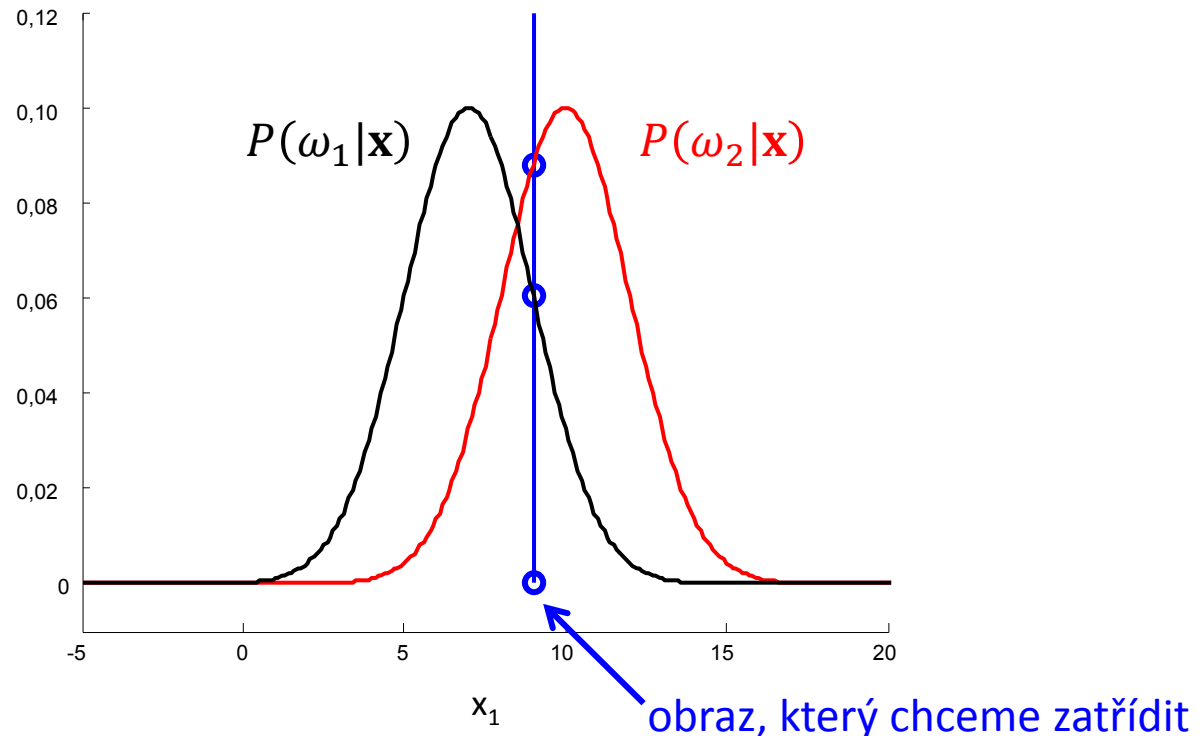
- Kritérium maximální a posteriorní pravděpodobnosti
- Kritérium minimální pravděpodobnosti chybného rozhodnutí
- Kritérium minimální střední ztráty
- Kritérium maximální pravděpodobnosti

Bayesův klasifikátor – kritéria

- Kritérium maximální a posteriorní pravděpodobnosti
- Kritérium minimální pravděpodobnosti chybného rozhodnutí
- Kritérium minimální střední ztráty
- Kritérium maximální pravděpodobnosti

Bayesův kl. – kritérium maximální aposteriori psti

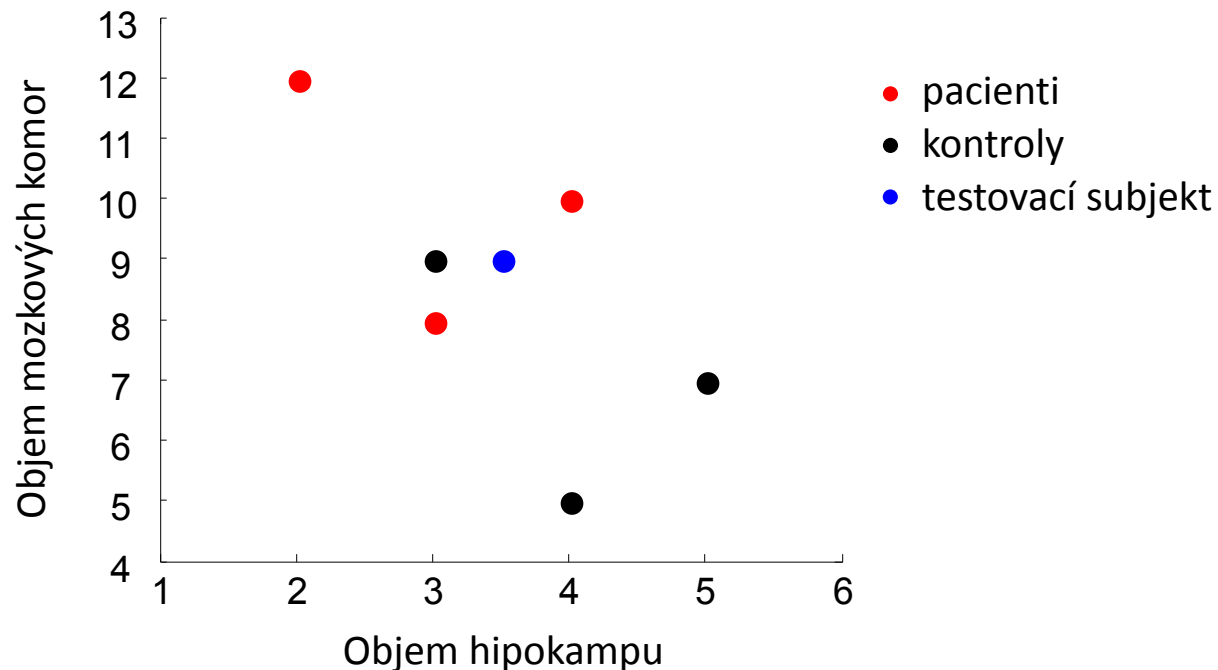
- zatřídění obrazu \mathbf{x} do třídy s větší aposteriori pravděpodobností, tedy:
když $P(\omega_1|\mathbf{x}) \geq P(\omega_2|\mathbf{x}) \rightarrow$ zařazení \mathbf{x} do třídy ω_1
když $P(\omega_1|\mathbf{x}) < P(\omega_2|\mathbf{x}) \rightarrow$ zařazení \mathbf{x} do třídy ω_2



Bayesův kl. – kritérium maximální aposteriorní psti

Příklad: Bylo provedeno měření objemu hipokampu a mozkových komor (v cm^3) u 3 pacientů se schizofrenií a 3 kontrol: $\mathbf{X}_D = \begin{bmatrix} 2 & 12 \\ 4 & 10 \\ 3 & 8 \end{bmatrix}$, $\mathbf{X}_H = \begin{bmatrix} 5 & 7 \\ 3 & 9 \\ 4 & 5 \end{bmatrix}$.

Určete, zda testovací subjekt $\mathbf{x} = [3,5 \ 9]$ patří do skupiny pacientů či kontrolních subjektů pomocí Bayesova klasifikátoru.



Bayesův kl. – kritérium maximální aposteriori psti

Příklad: Bylo provedeno měření objemu hipokampu a mozkových komor (v cm^3) u 3 pacientů se schizofrenií a 3 kontrol: $\mathbf{X}_D = \begin{bmatrix} 2 & 12 \\ 4 & 10 \\ 3 & 8 \end{bmatrix}$, $\mathbf{X}_H = \begin{bmatrix} 5 & 7 \\ 3 & 9 \\ 4 & 5 \end{bmatrix}$.

Určete, zda testovací subjekt $\mathbf{x} = [3,5 \quad 9]$ patří do skupiny pacientů či kontrolních subjektů pomocí Bayesova klasifikátoru.

$$P(\omega_k|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_k) \cdot P(\omega_k)}{p(\mathbf{x})}$$

Označení a pomocné výpočty:

$$n_D = 3; \quad n_H = 3; \quad n = 6$$

Apriorní psti:

$$P(\omega_D) = \frac{n_D}{n} = \frac{3}{6} = 0,5$$

$$P(\omega_H) = \frac{n_H}{n} = \frac{3}{6} = 0,5$$

Podmíněné hustoty psti:

$$p(\mathbf{x}|\omega_k) = \frac{1}{\sqrt{(2\pi)^2 |\mathbf{S}_k|}} \cdot \exp\left(-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{S}_k^{-1}(\mathbf{x} - \bar{\mathbf{x}})\right)$$

Potřebujeme vypočítat: vícerozměrné průměry $\bar{\mathbf{x}}_D$ a $\bar{\mathbf{x}}_H$; kovarianční matice \mathbf{S}_D a \mathbf{S}_H .

Bayesův kl. – kritérium maximální aposteriori psti

Příklad: Bylo provedeno měření objemu hipokampu a mozkových komor (v cm^3) u 3 pacientů se schizofrenií a 3 kontrol: $\mathbf{X}_D = \begin{bmatrix} 2 & 12 \\ 4 & 10 \\ 3 & 8 \end{bmatrix}$, $\mathbf{X}_H = \begin{bmatrix} 5 & 7 \\ 3 & 9 \\ 4 & 5 \end{bmatrix}$.

Určete, zda testovací subjekt $\mathbf{x} = [3,5 \quad 9]$ patří do skupiny pacientů či kontrolních subjektů pomocí Bayesova klasifikátoru.

$$P(\omega_k|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_k) \cdot P(\omega_k)}{p(\mathbf{x})}$$

Označení a pomocné výpočty:

Vícerozměrné průměry:

$$\bar{\mathbf{x}}_D = \left[\frac{1}{n_D} \sum_{i=1}^{n_D} x_{i1} \quad \frac{1}{n_D} \sum_{i=1}^{n_D} x_{i2} \right] = [3 \quad 10]$$

$$\bar{\mathbf{x}}_H = \left[\frac{1}{n_H} \sum_{i=1}^{n_H} x_{i1} \quad \frac{1}{n_H} \sum_{i=1}^{n_H} x_{i2} \right] = [4 \quad 7]$$

Výběrové kovarianční matice:

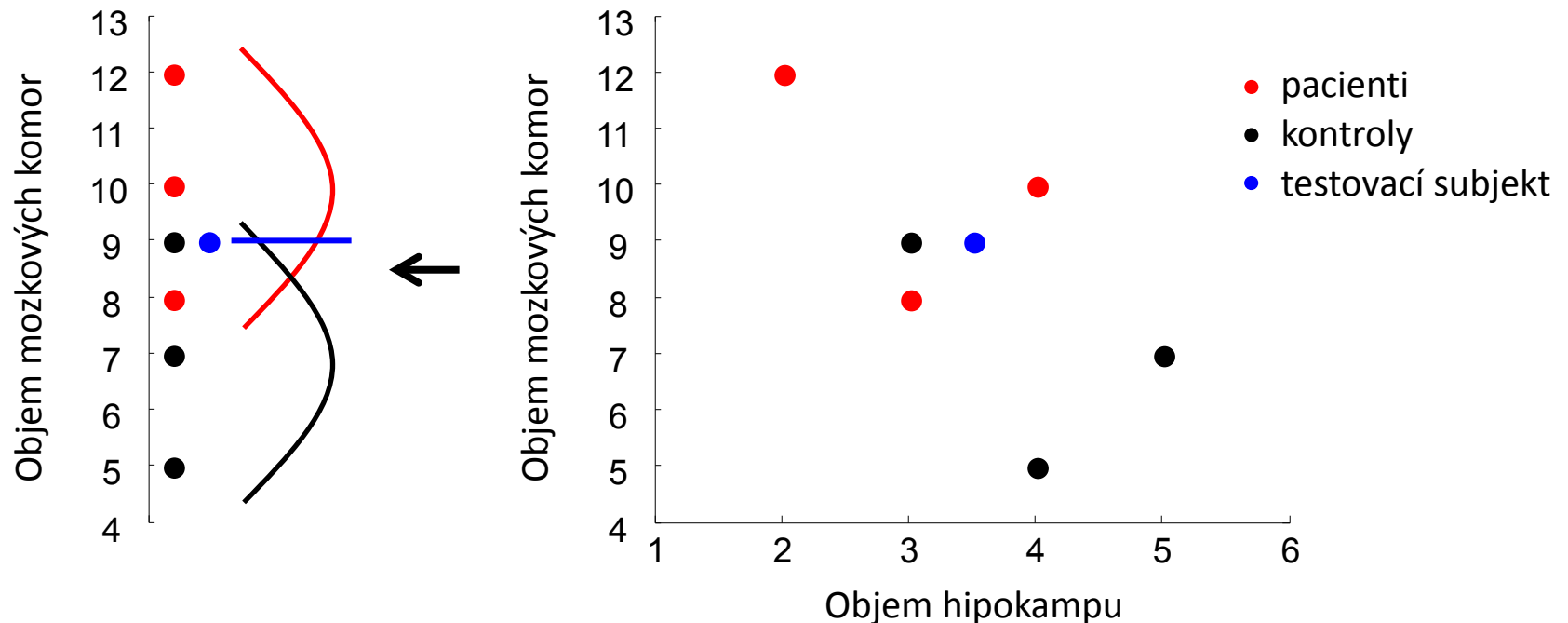
$$\mathbf{S}_D = \begin{bmatrix} s_{11}^D & s_{12}^D \\ s_{21}^D & s_{22}^D \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix}$$

$$\mathbf{S}_H = \begin{bmatrix} s_{11}^H & s_{12}^H \\ s_{21}^H & s_{22}^H \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix}$$

Bayesův kl. – kritérium maximální aposteriorní psti

Příklad:

1. Klasifikace podle objemu mozkových komor:



$$P(\omega_D|x_2) = \frac{0,176 \cdot 0,5}{0,1485} = 0,593$$

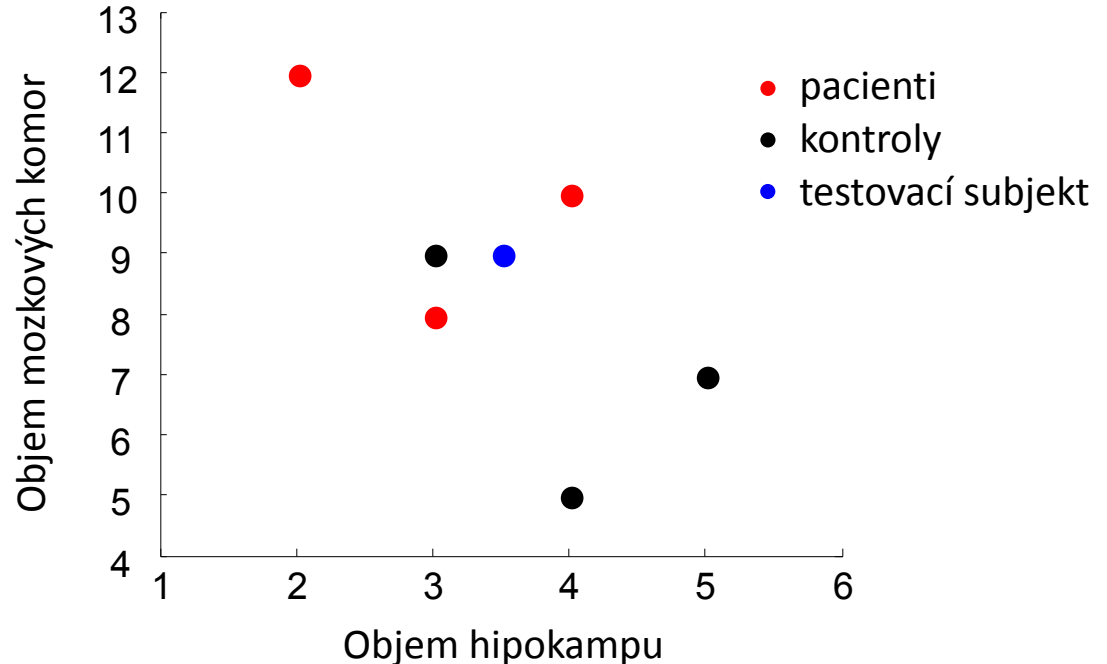
$$P(\omega_H|x_2) = \frac{0,121 \cdot 0,5}{0,1485} = 0,407$$

→ subjekt zařazen do třídy pacientů

Bayesův kl. – kritérium maximální aposteriori psti

Příklad:

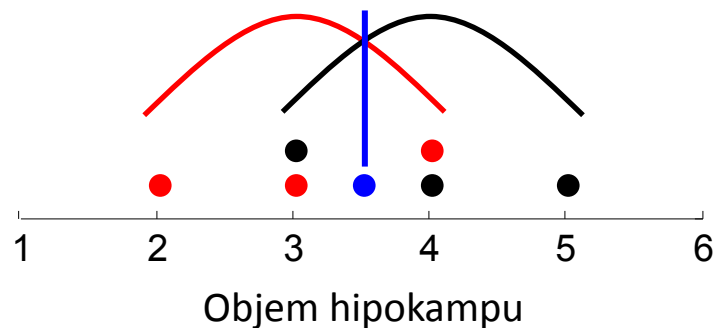
2. Klasifikace podle objemu hipokampu:



$$P(\omega_D|x_1) = \frac{0,352 \cdot 0,5}{0,352} = 0,5$$

$$P(\omega_H|x_1) = \frac{0,352 \cdot 0,5}{0,352} = 0,5$$

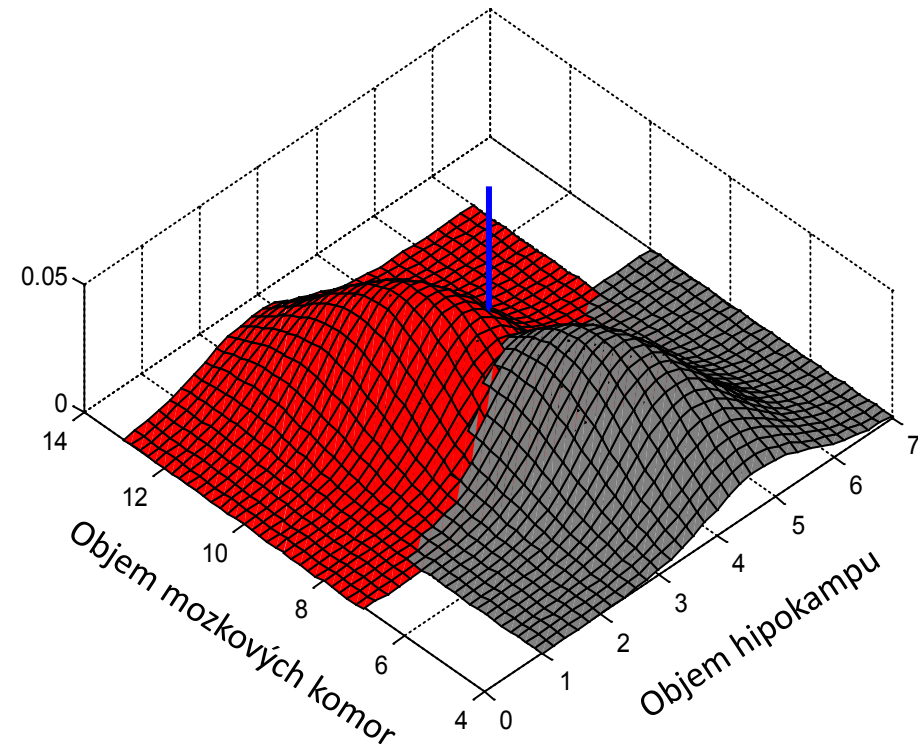
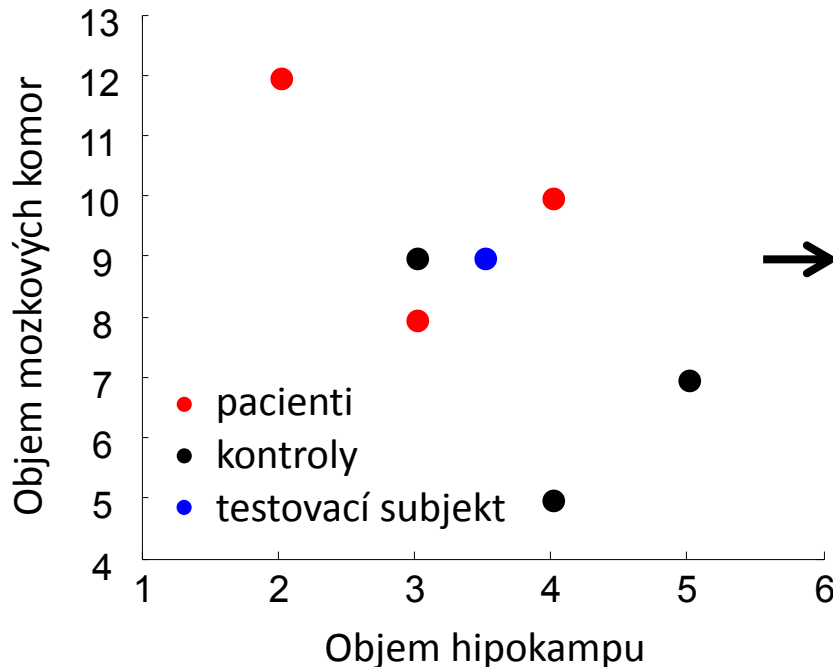
→ nelze jednoznačně určit, kam subjekt zařadíme



Bayesův kl. – kritérium maximální aposteriorní psti

Příklad:

3. Klasifikace podle obou proměnných:



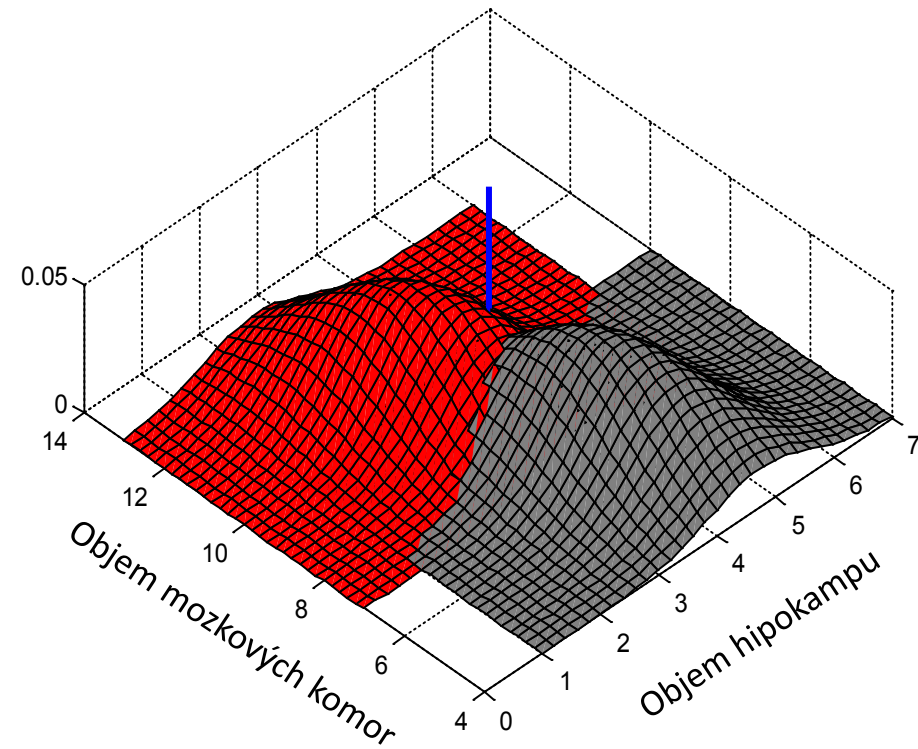
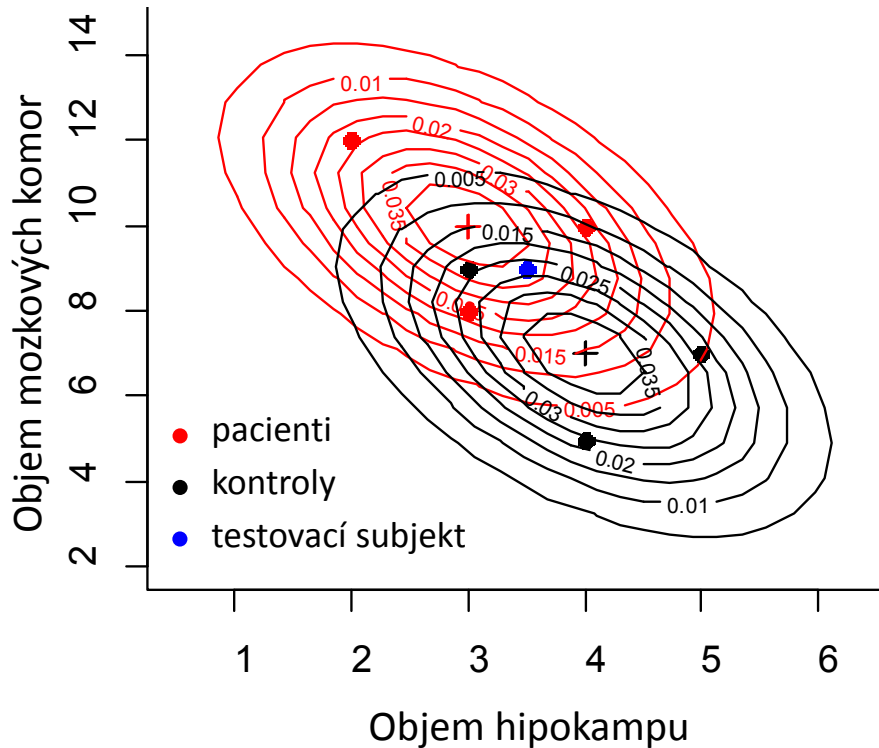
$$P(\omega_D|\mathbf{x}) = \frac{0,078 \cdot 0,5}{0,067} = 0,582$$

$$P(\omega_H|\mathbf{x}) = \frac{0,056 \cdot 0,5}{0,067} = 0,418$$

→ subjekt zařazen do třídy pacientů

Bayesův kl. – kritérium maximální a posteriorní psti

Příklad - klasifikace podle obou proměnných:



Bayesův kl. – kritérium maximální aposteriori psti

Příklad - Výpočet hranice pomocí diskriminačních funkcí:

Pro hranici je rozdíl diskriminačních funkcí roven 0:

$$P(\omega_D|\mathbf{x}) - P(\omega_H|\mathbf{x}) = 0$$

$$P(\omega_D|\mathbf{x}) = P(\omega_H|\mathbf{x})$$

$$\frac{P(\omega_D|\mathbf{x})}{P(\omega_H|\mathbf{x})} = 1 \rightarrow \text{kritérium maximální aposteriori pravděpodobnosti}$$

Levá strana je rovna: $\frac{P(\omega_D|\mathbf{x})}{P(\omega_H|\mathbf{x})} = \frac{0,582}{0,418} =$

1,4 Pravá strana je rovna: 1

→ protože věrohodnostní poměr (na levé straně) je větší než výraz na pravé straně, subjekt zařadíme do třídy pacientů

Bayesův klasifikátor – kritéria

- Kritérium maximální a posteriori pravděpodobnosti
- Kritérium minimální pravděpodobnosti chybného rozhodnutí
- Kritérium minimální střední ztráty
- Kritérium maximální pravděpodobnosti

Bayesův kl. – kritérium min. psti chybného rozhodnutí

- Vyjdeme z výpočtu hranice pomocí diskriminačních funkcí:

$$P(\omega_D|\mathbf{x}) - P(\omega_H|\mathbf{x}) = 0$$

$$\frac{p(\mathbf{x}|\omega_D) \cdot P(\omega_D)}{p(\mathbf{x})} - \frac{p(\mathbf{x}|\omega_H) \cdot P(\omega_H)}{p(\mathbf{x})} = 0$$

- Můžeme vykrátit $p(\mathbf{x})$, protože celková hustota pravděpodobnosti je stejná pro obě diskriminační funkce:

$$p(\mathbf{x}|\omega_D) \cdot P(\omega_D) - p(\mathbf{x}|\omega_H) \cdot P(\omega_H) = 0$$

- $\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{P(\omega_H)}{P(\omega_D)} \rightarrow$ kritérium minimální psti chybného rozhodnutí

Bayesův kl. – kritérium min. psti chybného rozhodnutí

$$\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{P(\omega_H)}{P(\omega_D)} \rightarrow \text{kritérium minimální psti chybného rozhodnutí}$$

Pro náš příklad:

$$\text{Levá strana je rovna: } \frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{0,078}{0,056} = 1,4$$

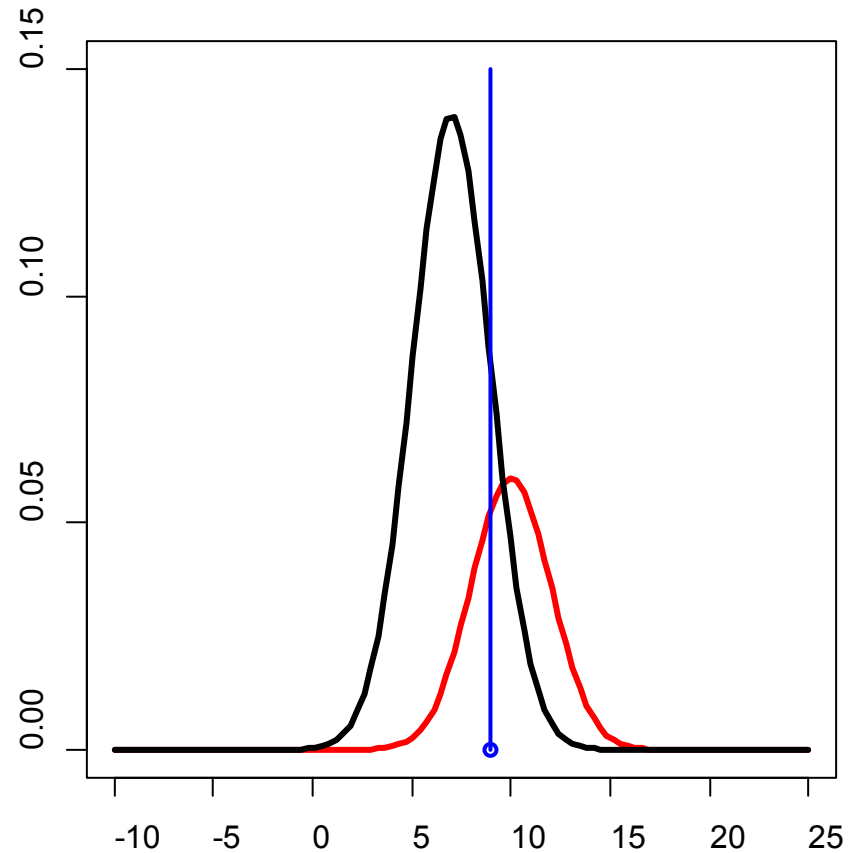
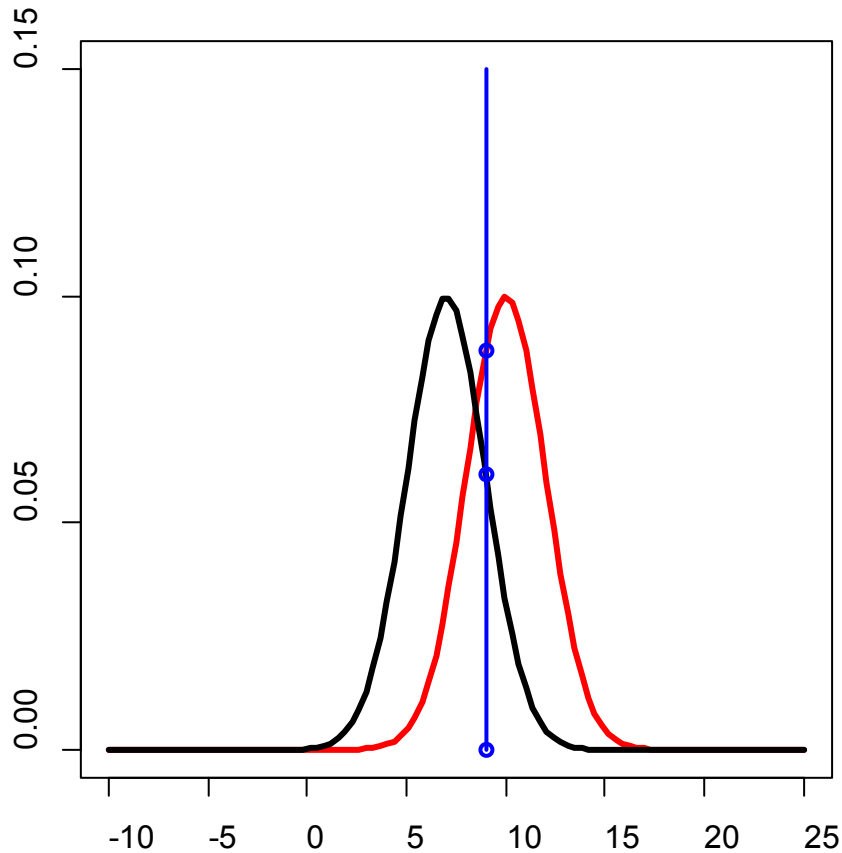
$$\text{Pravá strana rovna } \frac{P(\omega_H)}{P(\omega_D)} = \frac{0,5}{0,5} = 1$$

→ protože věrohodnostní poměr (na levé straně) je větší než výraz na pravé straně, subjekt zařadíme do třídy pacientů

Poznámka: Kdyby byly apriorní pravděpodobnosti jiné, např. $\frac{P(\omega_H)}{P(\omega_D)} = \frac{6/9}{3/9} = 2$, byl by testovací subjekt zařazen do třídy kontrolních subjektů.

Bayesův kl. – kritérium min. psti chybného rozhodnutí

Poznámka: Kdyby byly apriorní pravděpodobnosti jiné, např. $\frac{P(\omega_H)}{P(\omega_D)} = \frac{6/9}{3/9} = 2$, byl by testovací subjekt zařazen do třídy kontrolních subjektů.



Bayesův klasifikátor – kritéria

- Kritérium maximální a posteriorní pravděpodobnosti
- Kritérium minimální pravděpodobnosti chybného rozhodnutí
- **Kritérium minimální střední ztráty**
- Kritérium maximální pravděpodobnosti

Bayesův kl. – kritérium minimální střední ztráty

- Chceme do výpočtů zahrnout ztrátu při chybné klasifikaci obrazu ze třídy ω_s do třídy ω_r - ztráta definována pomocí **ztrátové funkce** $\lambda(\omega_r|\omega_s)$
- Ztrátové funkce zapíšeme do **matice ztrátových funkcí**:

$$\mathbf{\lambda} = \begin{bmatrix} \lambda(\omega_1|\omega_1) & \lambda(\omega_1|\omega_2) & \cdots & \lambda(\omega_1|\omega_R) \\ \lambda(\omega_2|\omega_1) & \lambda(\omega_2|\omega_2) & \cdots & \lambda(\omega_2|\omega_R) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda(\omega_R|\omega_1) & \lambda(\omega_R|\omega_2) & \cdots & \lambda(\omega_R|\omega_R) \end{bmatrix}$$

- např. $\lambda = \begin{bmatrix} \lambda(\omega_D|\omega_D) & \lambda(\omega_D|\omega_H) \\ \lambda(\omega_H|\omega_D) & \lambda(\omega_H|\omega_H) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix} \rightarrow$ víc penalizují, když je pacient nesprávně zařazen do třídy kontrolních subjektů, než když je kontrolní subjekt nesprávně zařazen do třídy pacientů

Bayesův kl. – kritérium minimální střední ztráty

Odvození kritéria minimální střední ztráty:

- Celková střední ztráta $J(\mathbf{a})$ je průměrná hodnota z dílčích středních ztrát:

$$J(\mathbf{a}) = \int_{\mathcal{X}} \sum_{s=1}^R \lambda(\omega_r | \omega_s) \cdot p(\mathbf{x} | \omega_s) \cdot P(\omega_s) d\mathbf{x}$$

- Chceme minimalizovat střední ztrátu:

$$J(\mathbf{a}^*) = \min_r \int_{\mathcal{X}} \sum_{s=1}^R \lambda(\omega_r | \omega_s) \cdot p(\mathbf{x} | \omega_s) \cdot P(\omega_s) d\mathbf{x}$$

- Hledáme minimální střední ztrátu, pokud ale chceme využít principu diskriminačních funkcí, budeme hledat maximum z výrazu se záporným znaménkem \rightarrow diskriminační funkce potom bude tvaru:

$$\mathbf{g}_r(\mathbf{x}) = - \sum_{s=1}^R \lambda(\omega_r | \omega_s) \cdot p(\mathbf{x} | \omega_s) \cdot P(\omega_s)$$

Bayesův kl. – kritérium minimální střední ztráty

Diskriminační funkce obecně: $g_r(\mathbf{x}) = - \sum_{s=1}^R \lambda(\omega_r|\omega_s) \cdot p(\mathbf{x}|\omega_s) \cdot P(\omega_s)$

Diskriminační funkce pro dichotomický klasifikátor:

$$g_1(\mathbf{x}) = -\lambda(\omega_1|\omega_1) \cdot p(\mathbf{x}|\omega_1) \cdot P(\omega_1) - \lambda(\omega_1|\omega_2) \cdot p(\mathbf{x}|\omega_2) \cdot P(\omega_2)$$

$$g_2(\mathbf{x}) = -\lambda(\omega_2|\omega_1) \cdot p(\mathbf{x}|\omega_1) \cdot P(\omega_1) - \lambda(\omega_2|\omega_2) \cdot p(\mathbf{x}|\omega_2) \cdot P(\omega_2)$$

Výpočet hranice pomocí diskriminačních funkcí:

$$g_1(\mathbf{x}) - g_2(\mathbf{x}) = 0$$

$$-\lambda(\omega_1|\omega_1) \cdot p(\mathbf{x}|\omega_1) \cdot P(\omega_1) - \lambda(\omega_1|\omega_2) \cdot p(\mathbf{x}|\omega_2) \cdot P(\omega_2) + \lambda(\omega_2|\omega_1) \cdot p(\mathbf{x}|\omega_1) \cdot P(\omega_1) + \lambda(\omega_2|\omega_2) \cdot p(\mathbf{x}|\omega_2) \cdot P(\omega_2) = 0$$

$$(\lambda(\omega_2|\omega_1) - \lambda(\omega_1|\omega_1)) \cdot p(\mathbf{x}|\omega_1) \cdot P(\omega_1) + (\lambda(\omega_2|\omega_2) - \lambda(\omega_1|\omega_2)) \cdot p(\mathbf{x}|\omega_2) \cdot P(\omega_2) = 0$$

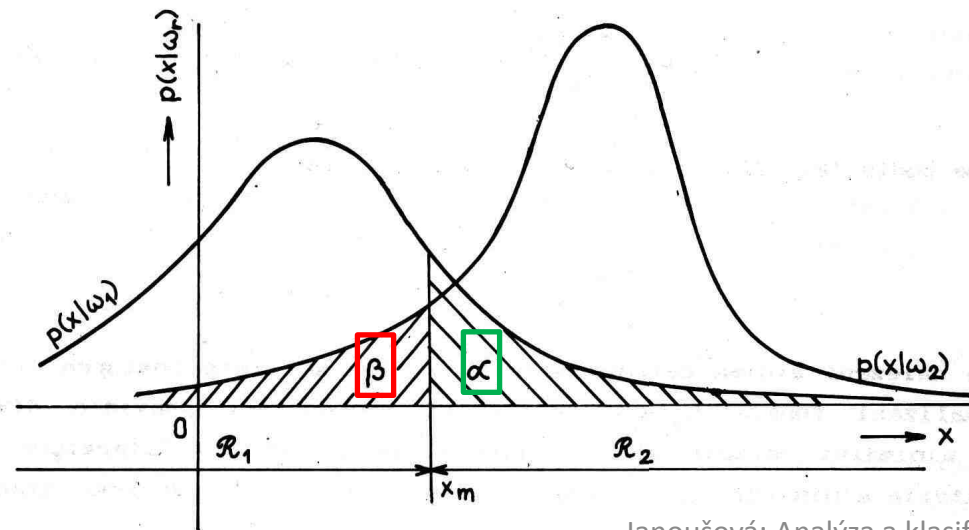
$$(\lambda(\omega_2|\omega_1) - \lambda(\omega_1|\omega_1)) \cdot p(\mathbf{x}|\omega_1) \cdot P(\omega_1) = (\lambda(\omega_1|\omega_2) - \lambda(\omega_2|\omega_2)) \cdot p(\mathbf{x}|\omega_2) \cdot P(\omega_2)$$

$$\frac{p(\mathbf{x}|\omega_1)}{p(\mathbf{x}|\omega_2)} = \frac{(\lambda(\omega_1|\omega_2) - \lambda(\omega_2|\omega_2))P(\omega_2)}{(\lambda(\omega_2|\omega_1) - \lambda(\omega_1|\omega_1))P(\omega_1)} \rightarrow \text{kritérium minimální střední ztráty}$$

Bayesův kl. – kritérium minimální střední ztráty

Celková střední ztráta v případě dvou tříd je:

$$\begin{aligned}
 J(\mathbf{a}) &= \int_{\mathcal{R}_1} \sum_{s=1}^2 \lambda(\omega_1|\omega_s) \cdot p(\mathbf{x}|\omega_s) \cdot P(\omega_s) d\mathbf{x} + \int_{\mathcal{R}_2} \sum_{s=1}^2 \lambda(\omega_2|\omega_s) \cdot p(\mathbf{x}|\omega_s) \cdot P(\omega_s) d\mathbf{x} = \\
 &= \lambda(\omega_1|\omega_1) \cdot P(\omega_1) \int_{\mathcal{R}_1} p(\mathbf{x}|\omega_1) \cdot d\mathbf{x} + \lambda(\omega_1|\omega_2) \cdot P(\omega_2) \int_{\mathcal{R}_1} p(\mathbf{x}|\omega_2) \cdot d\mathbf{x} + \\
 &+ \lambda(\omega_2|\omega_1) \cdot P(\omega_1) \int_{\mathcal{R}_2} p(\mathbf{x}|\omega_1) \cdot d\mathbf{x} + \lambda(\omega_2|\omega_2) \cdot P(\omega_2) \int_{\mathcal{R}_2} p(\mathbf{x}|\omega_2) \cdot d\mathbf{x} = \\
 &= \lambda(\omega_1|\omega_1) \cdot P(\omega_1) \cdot (1 - \alpha) + \lambda(\omega_1|\omega_2) \cdot P(\omega_2) \cdot \beta + \lambda(\omega_2|\omega_1) \cdot P(\omega_1) \cdot \alpha + \lambda(\omega_2|\omega_2) \cdot P(\omega_2) \cdot (1 - \beta)
 \end{aligned}$$



Bayesův kl. – kritérium minimální střední ztráty

$$\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{(\lambda(\omega_D|\omega_H) - \lambda(\omega_H|\omega_H))P(\omega_H)}{(\lambda(\omega_H|\omega_D) - \lambda(\omega_D|\omega_D))P(\omega_D)} \rightarrow \text{kritérium minimální střední ztráty}$$

Pro náš příklad:

Levá strana je rovna: $\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{0,078}{0,056} = 1,4$

Pravá strana je při různém nastavení vah rovna:

- $\lambda = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$ (tzn., více penalizují, pokud je pacient nesprávně zařazen do třídy kontrolních subjektů, než když je kontrolní subjekt nesprávně zařazen do třídy pacientů), pak pravá strana je rovna $\frac{(1-0) \cdot 0,5}{(2-0) \cdot 0,5} = 0,5$ a subjekt zařadím do třídy pacientů.
- $\lambda = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (penalizují shodně nesprávné zařazení do třídy kontrolních subjektů i pacientů – kritérium minimální pravděpodobnosti chybného rozhodnutí), pak pravá strana je rovna $\frac{(1-0) \cdot 0,5}{(1-0) \cdot 0,5} = 1$ a subjekt zařadím do třídy pacientů.
- $\lambda = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$ (tzn., více penalizují, pokud je kontrolní subjekt nesprávně zařazen do třídy pacientů, než když je pacient nesprávně zařazen do třídy kontrolních subjektů), pak pravá strana je rovna $\frac{(2-0) \cdot 0,5}{(1-0) \cdot 0,5} = 2$ a subjekt zařadím do třídy kontrolních subjektů.

Bayesův kl. – kritérium minimální střední ztráty

- **Poznámka:** pokud nastavíme matici ztrátových funkcí ve tvaru $\lambda = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, dostáváme kritérium minimální psti chybného rozhodnutí:

$$\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{(\lambda(\omega_D|\omega_H) - \lambda(\omega_H|\omega_H))P(\omega_H)}{(\lambda(\omega_H|\omega_D) - \lambda(\omega_D|\omega_D))P(\omega_D)}$$

$$\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{(1 - 0)P(\omega_H)}{(1 - 0)P(\omega_D)}$$

- $\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{P(\omega_H)}{P(\omega_D)} \rightarrow$ kritérium minimální psti chybného rozhodnutí

Bayesův klasifikátor – kritéria

- Kritérium maximální a posteriorní pravděpodobnosti
- Kritérium minimální pravděpodobnosti chybného rozhodnutí
- Kritérium minimální střední ztráty
- **Kritérium maximální pravděpodobnosti**

Bayesův kl. – kritérium maximální psti

- Předpoklady:
 - rovnoměrné zastoupení K tříd, tzn. $P(\omega_D) = P(\omega_H) = \frac{1}{K} = \frac{1}{2} = 0,5$
 - nulové ztráty při správném rozhodnutí, tzn. $\lambda(\omega_D|\omega_D) = \lambda(\omega_H|\omega_H) = 0$
- pak získáváme po dosazení do obecného vzorce pro výpočet věrohodnostního poměru:

$$\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{(\lambda(\omega_D|\omega_H) - 0) \cdot 0,5}{(\lambda(\omega_H|\omega_D) - 0) \cdot 0,5}$$

- $\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{\lambda(\omega_D|\omega_H)}{\lambda(\omega_H|\omega_D)} \rightarrow$ kritérium maximální pravděpodobnosti

Bayesův kl. – kritérium maximální psti

$$\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{\lambda(\omega_D|\omega_H)}{\lambda(\omega_H|\omega_D)} \rightarrow \text{kritérium maximální pravděpodobnosti}$$

Pro náš příklad:

Levá strana je rovna: $\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{0,078}{0,056} = 1,4$

Pravá strana je při různém nastavení vah rovna:

- $\lambda = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$ (tzn., více penalizuji, pokud je pacient nesprávně zařazen do třídy kontrolních subjektů, než když je kontrolní subjekt nesprávně zařazen do třídy pacientů), pak pravá strana je rovna $\frac{1}{2} = 0,5$ a subjekt zařadím do třídy pacientů.
- $\lambda = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (penalizuji shodně nesprávné zařazení do třídy kontrolních subjektů i pacientů), pak pravá strana je rovna $\frac{1}{1} = 1$ a subjekt zařadím do třídy pacientů.
- $\lambda = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$ (tzn., více penalizuji, pokud je kontrolní subjekt nesprávně zařazen do třídy pacientů, než když je pacient nesprávně zařazen do třídy kontrolních subjektů), pak pravá strana je rovna $\frac{2}{1} = 2$ a subjekt zařadím do třídy kontrolních subjektů.

Bayesův klasifikátor – kritéria

- Kritérium maximální a posteriorní pravděpodobnosti:

$$\frac{P(\omega_D|\mathbf{x})}{P(\omega_H|\mathbf{x})} = 1$$

- Kritérium minimální pravděpodobnosti chybného rozhodnutí

$$\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{P(\omega_H)}{P(\omega_D)}$$

- Kritérium minimální střední ztráty

$$\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{(\lambda(\omega_D|\omega_H) - \lambda(\omega_H|\omega_H))P(\omega_H)}{(\lambda(\omega_H|\omega_D) - \lambda(\omega_D|\omega_D))P(\omega_D)}$$

- Kritérium maximální pravděpodobnosti

$$\frac{p(\mathbf{x}|\omega_D)}{p(\mathbf{x}|\omega_H)} = \frac{\lambda(\omega_D|\omega_H)}{\lambda(\omega_H|\omega_D)}$$

Příprava nových učebních materiálů pro obor Matematická biologie

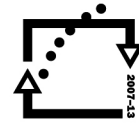
je podporována projektem OPVK

č. CZ.1.07/2.2.00/28.0043

„Interdisciplinární rozvoj studijního
oboru Matematická biologie“



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ