

# 3 - BAYESIAN LEARNING

Aprendizagem 2024/2025

## TEOREMA DE BAYES

- As crenças (A) podem ser suportadas por evidências (B), dada pelo seguinte teorema:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- O valor da previsão  $\hat{y}$  é dado por:

$$\hat{y} = \operatorname{argmax}_{c_i} \{p(c_i|x)\} = \operatorname{argmax}_{c_i} \left\{ \frac{p(x|c_i)p(c_i)}{p(x)} \right\}$$

## ASSUMPTIONS

- **Maximum a posteriori** (o valor do denominador vai ser igual para todos os valores calculados):

$$\hat{y}_{MAP} = \operatorname{argmax}_{c_i} \{p(c_i|x)\} = \operatorname{argmax}_{c_i} \left\{ \frac{p(x|c_i)p(c_i)}{p(x)} \right\}$$

- **Maximum likelihood** (se assumirmos que todas as crenças têm a mesma probabilidade):

$$\hat{y}_{ML} = \operatorname{argmax}_{c_i} \{p(c_i|x)\} = \operatorname{argmax}_{c_i} \left\{ \frac{p(x|c_i)p(c_i)}{p(x)} \right\}$$

# NAÏVE BAYES

- Todas as variáveis são condicionalmente independentes, e dado pela seguinte fórmula:

$$\hat{y}_{MAP} = \operatorname{argmax}_{c_i} \{p(c_i|x)\} = \operatorname{argmax}_{c_i} \left\{ p(c_i) \prod_{j=1}^d p(x_j|c_i) \right\}$$

## CÁLCULO DAS PROBABILIDADES

- Variável **categórica**: cálculo direto da probabilidade
- Variável **numérica**: distribuição normal (1 dimensão)

$$p(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

- Variáveis **numéricas**: distribuição normal (multivariada)

$$p(x|\mu, \Sigma) = \frac{1}{(2\pi)^{\frac{m}{2}} \sqrt{|\Sigma|}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}$$

## OUTRAS FÓRMULAS

- **Matriz de covariâncias** entre  $y_1$  e  $y_2$  ( $\Sigma$ ):

$$\begin{pmatrix} \sigma_{y_1}^2 & cov(y_2, y_1) \\ cov(y_1, y_2) & \sigma_{y_2}^2 \end{pmatrix}$$

$$\sigma_{y_1}^2 = \frac{1}{n-1} \sum (y_{1i} - \mu_1)^2$$

$$cov(y_1, y_2) = cov(y_2, y_1) = \frac{1}{n-1} \sum (y_{1i} - \mu_1)(y_{2i} - \mu_2)$$

# SUMÁRIO

- Ficha 3

Í