

# Reconhecimento de Padrões

Classificador Bayes

# Classificador Bayesiano

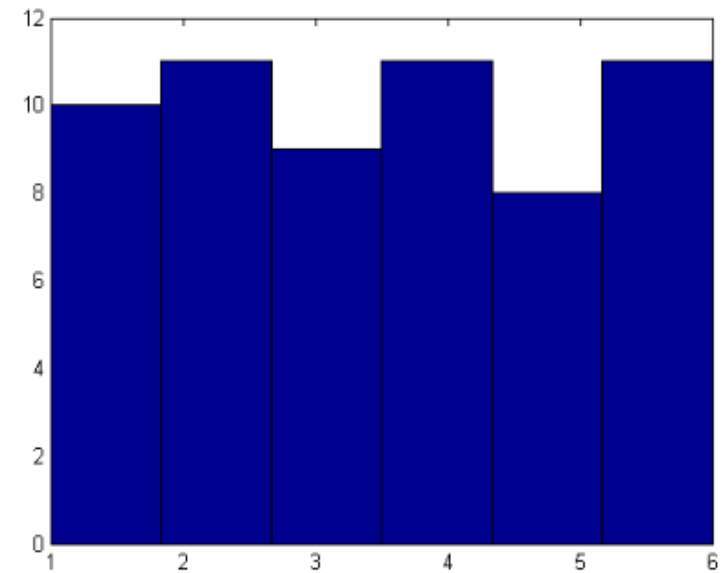
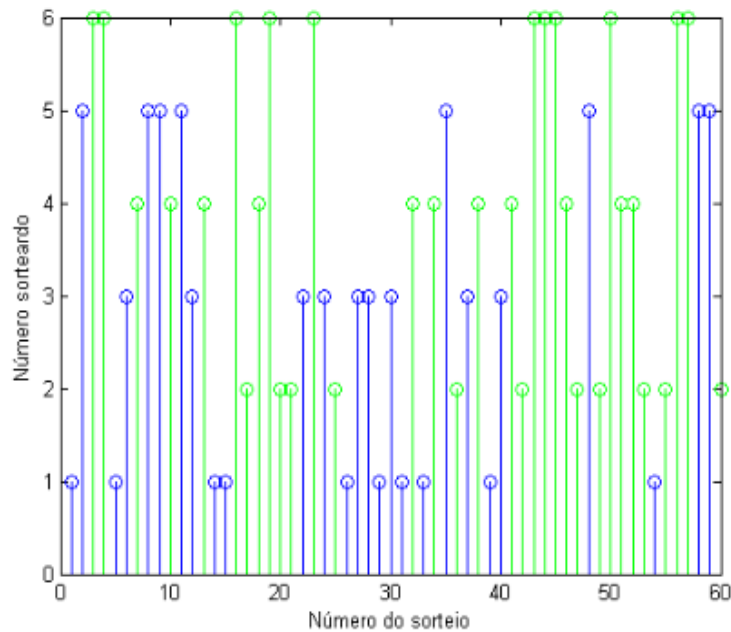
- Classificadores estatísticos
- Baseados na probabilidade de um elemento pertencer a uma classe específica
- Rápido
- Muito preciso quando volume de dados é grande

# Classificador Bayesiano simples

- Baseado em probabilidades condicionais
- Tem como hipótese que seus atributos são independentes – simplifica cálculos

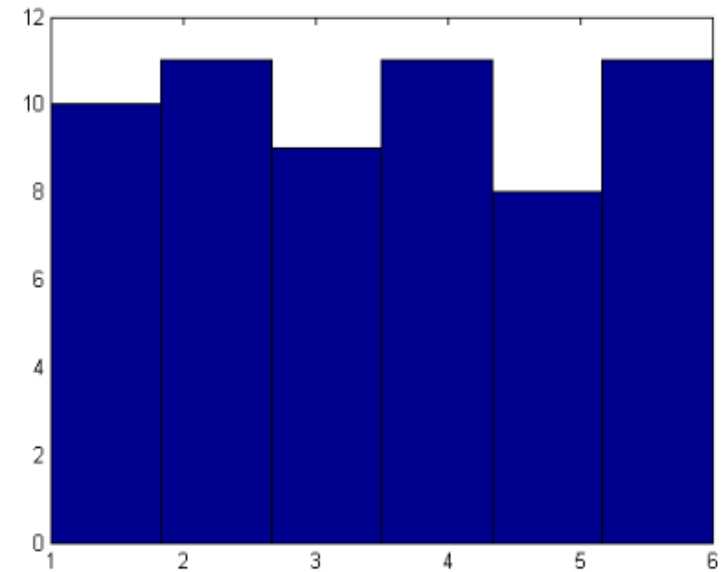
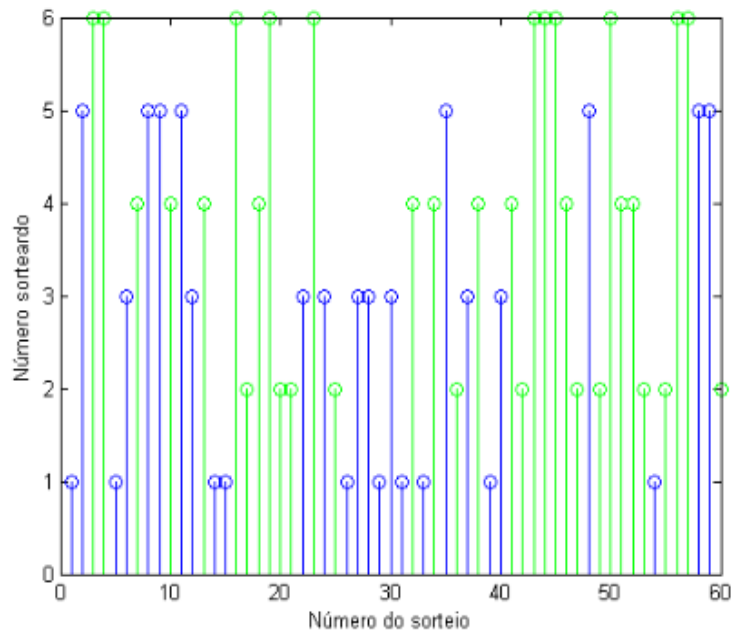
# Entendendo as probabilidades

- Problema do dado



# Entendendo as probabilidades

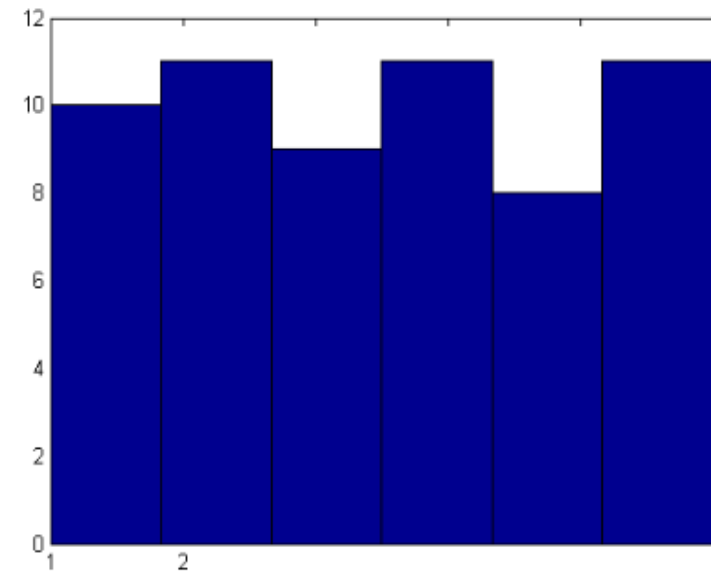
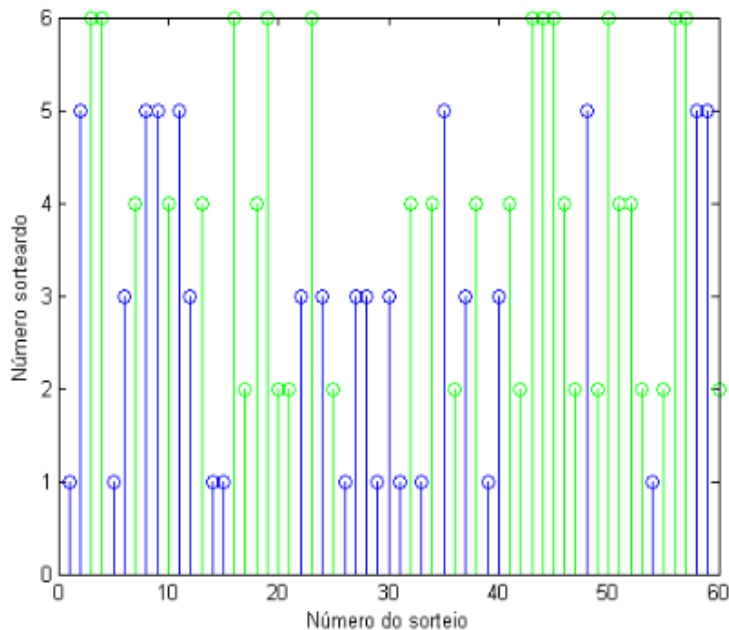
- Problema do dado



Qual a probabilidade de ocorrer o evento  $A = \{1\}$ ?

# Entendendo as probabilidades

- Problema do dado

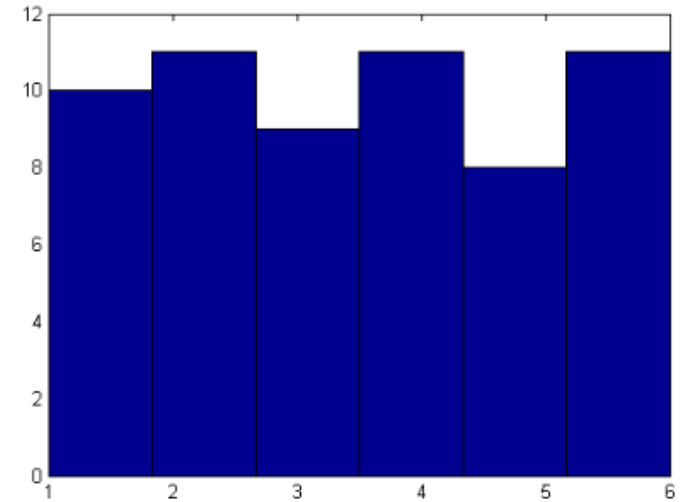


Qual a probabilidade de ocorrer o evento  $A = \{1\}$ ?

$$P[A] = \frac{10}{60} = \frac{1}{6}$$

# Entendendo as probabilidades

- Problema do dado



Qual a probabilidade de ocorrer evento  $A = \{1\}$ , dado que  $B = \{\text{ímpar}\}$  ocorreu?

$$P(A|B) \approx \frac{N_A}{N_B} \approx \frac{10}{27} \approx \frac{1}{3}$$

# Probabilidade condicional

E se temos o evento  $A = \{1, 4\}$ ? Dado que o evento B anterior ocorreu

Como B é o conjunto  
Dos números ímpares  
Temos que tomar o  
Cuidado de considerar  
apenas os números ímpares

$$\frac{N_{A \cap B}}{N_B} = \frac{\frac{N_{A \cap B}}{N_S}}{\frac{N_B}{N_S}} \approx \frac{P[A \cap B]}{P[B]}$$

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$



# Dedução da regra de Bayes

$$P[A \cap B] = P[A|B] P[B]$$

$$P[A \cap B] = P[B|A] P[A]$$

$$P[A|B] = \frac{P[B|A] P[A]}{P[B]}$$

# Classificador

$$P[C_1|x] = \frac{P[x|C_1] P[C_1]}{P[x]}$$

$$P[C_2|x] = \frac{P[x|C_2] P[C_2]}{P[x]}$$

Decisão

$$\blacktriangleright \frac{P[C_1|x]}{P[C_2|x]} \geq 1$$

# Classificador (simplificação)

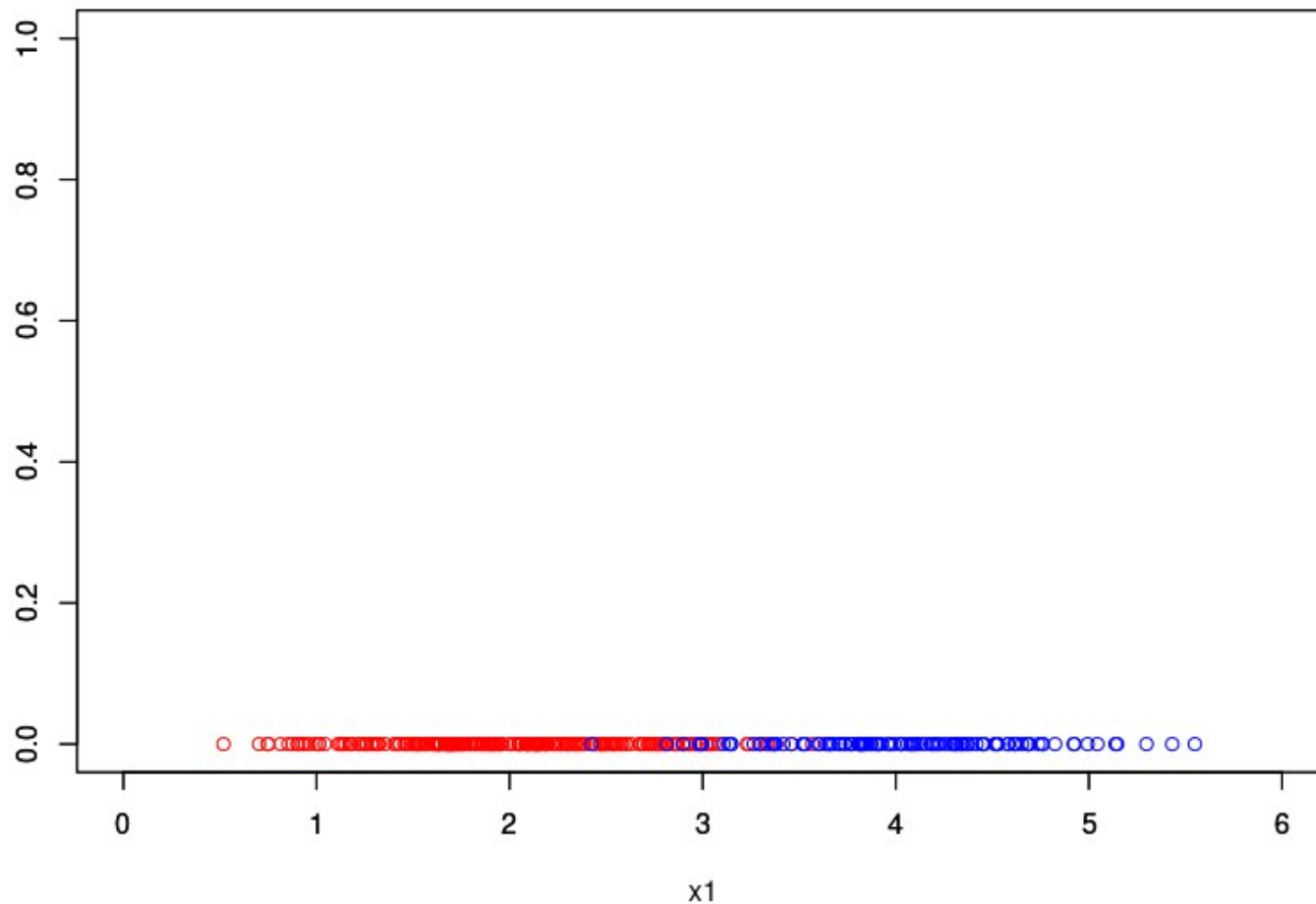
$$\frac{P[C_1|x]}{P[C_2|x]} = \frac{P[x|C_1] P[C_1]}{P[x|C_2] P[C_2]}$$

Simplificação

► Não precisa calcular  $P[x]$

# Classificador (exemplo)

Problema a ser resolvido



# Classificador (exemplo)

Cálculo das probabilidades a priori

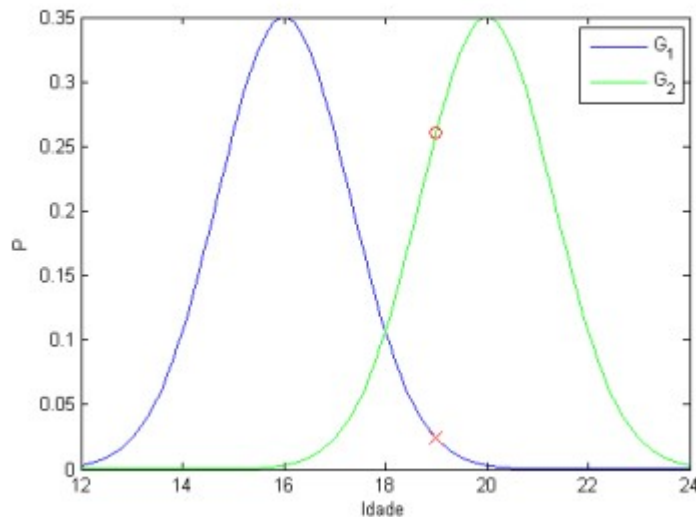
$$P[C_1] = \frac{n_1}{n_1 + n_2} = \frac{240}{240 + 120} = 0.667$$

$$P[C_2] = \frac{n_2}{n_1 + n_2} = \frac{120}{240 + 120} = 0.333$$

# Classificador (exemplo)

Cálculo das verosimilhanças

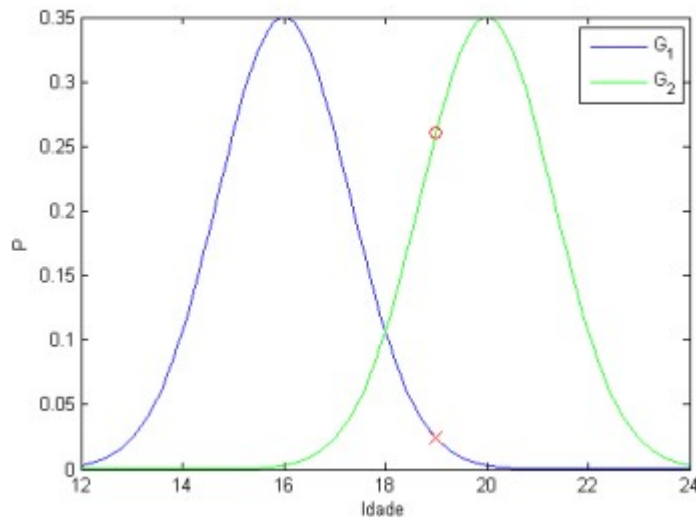
$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



# Classificador (exemplo)

Cálculo das verosimilhanças

Para o elemento selecionado

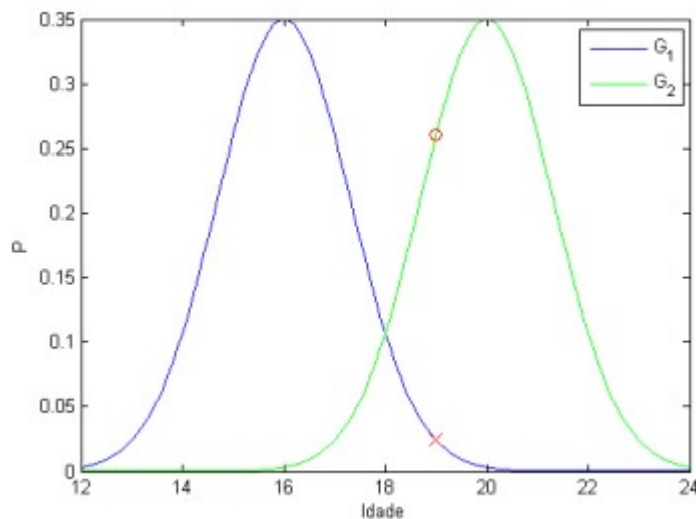


$$P[x|C1] = 0.1967$$

$$P[x|C2] = 1.45 \times 10^{-4}$$

# Classificador (exemplo)

Cálculo das verosimilhanças



$$\frac{P[C_1|x]}{P[C_2|x]} = \frac{P[x|C_1] P[C_1]}{P[x|C_2] P[C_2]}$$

$$\frac{0.1967 \times 0.667}{1.45 \times 10^{-4} \times 0.333} \approx 2698$$

Classe 1



# Observações

## Algumas observações

- ▶ Sem informação à priori - não funciona
- ▶ Quanto mais dados melhor a classificação

## Atenção

- ▶ Influência muito forte do desbalanceamento
- ▶ Engloba menor classe....
- ▶ Aplicação criteriosa!!!!