



**UNIVERSIDADE FEDERAL DE SERGIPE  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

Thiago Rocha de Melo - 201900126591

**Atividade 02  
Unidade II**

São Cristóvão

18/07/2023

# Capítulo 1

## Estimação da PDF usando Parzen com kernel Gaussiano

Para estimar a pdf utilizando Parzen, primeiro definimos o tamanho da janela. Neste experimento, utilizamos

$$h \approx \sqrt{N}$$

sendo  $N$  o tamanho dos dados utilizados. Como nossos dados possuem dimensão igual a um, o volume  $V$  é igual ao tamanho da janela  $h$ . Dessa forma, a probabilidade de  $x$  é dada pela equação

$$p(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{V\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

sendo

$$u = \frac{(x - x_i)}{h}$$

### 1.1 Classificador Bayesiano

Uma vez tendo a estimação da PDF podemos construir o classificador bayesiano definindo os "a priori". Na base de dados utilizada temos aproximadamente  $\frac{2}{3}$  dos dados para classe 0 - não diabético e  $\frac{1}{3}$  para classe 1 - diabético. Utilizaremos esses valores como "a priori", ou seja, a probabilidade da classe ( $P(Classe)$ ) enquanto que a PDF estimada será utilizada como a probabilidade de  $x$  dado a classe ( $P(X|Classe)$ ). O classificador será dado por

$$ClasseEstimada = \max[P(X|Classe0) \cdot P(Classe0), P(X|Classe1) \cdot P(Classe1)].$$

A taxa de acerto do classificador foi de 74.05%

A seguir temos o gráfico mostrando a PDF estimada de cada uma das classes.

### 1.1.1 Resultados

