



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE  
Centro de Engenharia Elétrica e Informática  
Programa de Pós-Graduação em Engenharia Elétrica

Tese de Doutorado  
**Emprego do Estimador KSG na Classificação de  
Patologias Laríngeas por Intermédio da Análise  
Dinâmica Não-Linear**

Mikaelle Oliveira Santos Gomes

Francisco Marcos de Assis, Dr.  
Orientador

Campina Grande  
Abril – 2019



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE  
Centro de Engenharia Elétrica e Informática  
Programa de Pós-Graduação em Engenharia Elétrica

## **Emprego do Estimador KSG na Classificação de Patologias Laríngeas por Intermédio da Análise Dinâmica Não-Linear**

Mikaelle Oliveira Santos Gomes

Tese de doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Campina Grande como requisito para obtenção do título de doutora em engenharia elétrica.

Área de concentração: Processamento de Informação

Dr. Francisco Marcos de Assis  
(Orientador)

Abril de 2019

G633e      Gomes, Mikaelle Oliveira Santos.  
              Emprego do estimador KSG na classificação de patologias laríngeas  
              por intermédio da análise dinâmica não-linear / Mikaelle Oliveira Santos  
              Gomes. – Campina Grande, 2019.  
              86 f. : il. color.

              Tese (Doutorado em Engenharia Elétrica) – Universidade Federal de  
Campina Grande, Centro de Engenharia Elétrica e Informática, 2019.  
              "Orientação: Prof. Dr. Francisco Marcos de Assis".  
              Referências.

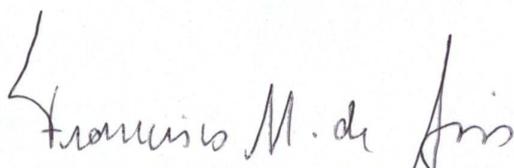
              1. Sistemas Dinâmicos. 2. Informação Mútua. 3. Patologias  
Laríngeas. 4. Estimador KSG. I. Assis, Francisco Marcos de. II. Título.

CDU 517.93(043)

**"EMPREGO DO ESTIMADOR KSG NA CLASSIFICAÇÃO DE PATOLOGIAS  
LARÍNGEAS POR INTERMÉDIO DA ANÁLISE DINÂMICA NÃO-LINEAR"**

**MIKAELLE OLIVEIRA SANTOS**

**TESE APROVADA EM 05/04/2019**



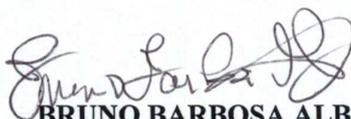
**FRANCISCO MARCOS DE ASSIS, Dr., UFCG  
Orientador(a)**

**GUILHERME DE ALENCAR BARRETO, Dr, UFC  
Examinador(a)**

**ROSÂNGELA FERNANDES COELHO, Docteur, IME  
Examinador(a)**



**RAIMUNDO CARLOS SILVÉRIO FREIRE, Dr., UFCG  
Examinador(a)**



**BRUNO BARBOSA ALBERT, D.Sc., UFCG  
Examinador(a)**

**CAMPINA GRANDE - PB**

*“A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que ninguém ainda  
pensou sobre aquilo que todo mundo vê.”  
(Arthur Schopenhauer)*

## Agradecimentos

- ★ À Deus, Senhor da vida, por tudo que eu pude vivenciar até hoje, pelas pessoas que conheci, e por tudo que ainda está por vir.
- ★ Aos meus pais, Inaldete e Adilson, por todo amor, educação, carinho e paciência para comigo.
- ★ Ao meu esposo, Ítalo Silva, pelo incentivo, companheirismo e paciência com meus momentos de ausência e aos meus irmãos, Kleiton e Kleilton, por todo apoio.
- ★ Ao Professor Francisco Marcos de Assis, meu orientador, um agradecimento carinhoso, por todos os momentos de paciência, dedicando parte do seu tempo, desde os últimos anos, para compartilhar comigo seus valiosos conhecimentos, não apenas na área acadêmica, mas também dando conselhos e ensinando valores humanos.
- ★ Aos Professores Rosângela Coelho, Guilherme Barreto, Bruno Albert e Raimundo Freire, membros da Banca, por aceitar avaliar este trabalho, de forma a compartilhar os seus valiosos conhecimentos e acrescentar mais valor a esta pesquisa.
- ★ Aos colegas do doutorado, pela torcida, pelo conhecimento compartilhado, pelas conversas e palavras de motivação.
- ★ À COPELE, em especial Ângela e Pedro, pela disponibilidade e pelos serviços prestados durante o curso deste trabalho acadêmico.
- ★ À Capes pelo fomento durante o período desta pesquisa.
- ★ Agradeço por fim à todos os meus antigos professores, amigos e colegas que contribuíram causalmente na minha formação até aqui.

## Resumo

Os distúrbios da voz e as patologias laríngeas podem afetar diferentes grupos etários. Muitos desses desvios o ser humano traz consigo desde o seu nascimento ou podem ainda ser desenvolvidos ao longo da vida. Devido a sua não estacionariedade e sua dinâmica caótica, os sinais de voz não podem ser corretamente analisados a partir de métodos tradicionais de análises de sinais experimentais. Dessa forma, a teoria do caos, uma área da teoria dos sistemas de dinâmicos não-lineares, aplicada em séries temporais não-lineares, tem sido adotada como uma nova abordagem não-linear do processamento do sinal de voz. Neste contexto, a presente tese propõe-se a investigar a aplicabilidade de um estimador de informação mútua (estimador KSG) e suas variantes (estimador KOLE), baseado nos  $k$ -vizinhos mais próximos, na estimativa de medidas da análise dinâmica não linear: atraso de reconstrução do espaço de fase em sistemas dinâmicos ( $\tau$ ), primeiro mínimo da função de informação mútua ( $PMIM$ ), entropia de Shannon ( $H$ ) e entropia de correlação ( $K_2$ ). Estudos revelaram que o estimador KSG é menos tendencioso (possui menos viés) do que o estimador ingênuo comumente usado, o que motivou a aplicá-lo para detectar a presença de distúrbios da voz causados pelas patologias edema de Reinke, paralisia das pregas vocais e nódulos nas pregas vocais. Foram utilizados dois classificadores diferentes para obtenção dos resultados: baseado em análise discriminante e um utilizando a abordagem de Máquinas de Vetor de Suporte (SVM). A comparação dos resultados obtidos com o estimador ingênuo e o com o estimador KSG, para ambos os classificadores empregados, mostraram que a utilização do estimador KSG melhora os resultados da classificação em termos de acurácia mostrando-se assim, um método de estimação eficiente.

**Palavras-chave.** Patologias Laríngeas, Sistemas Dinâmicos, Informação Mútua, Estimador KSG.

## Abstract

Voice disorders and laryngeal pathologies may affect different age groups. The human being brings many of these deviations with him/her from birth, or might still develop throughout life. Due to its non-stationarity and its chaotic dynamics, voice signals can not be correctly analyzed from traditional methods of analyzing experimental signals. Thus, the chaos theory, an area of non-linear dynamical systems theory, applied in non-linear time series, has been adopted as a new non-linear approach to voice signal processing. In this context, this thesis proposes to investigate the applicability of a estimator of mutual information (KSG estimator) and their variants (KOLE estimator), based on the nearest k-neighbors, in the estimation of measurements of non linear dynamic analysis: reconstruction delay estimation of phase space in dynamical systems ( $\tau$ ), first minimum of the mutual information function (*PMIM*), Shannon's entropy ( $H$ ) and correlation entropy( $K_2$ ). Studies have revealed that the KSG estimator is less biased than the naive estimator commonly used, which motivated us to apply it to detect the presence of voice disorders caused by laryngeal pathologies such as Reinke's edema, paralysis on the vocal folds and nodules on the vocal folds. Two types of classifiers were used to obtain results: one based on discriminant analysis and one based on a support vector approach (SVM). The comparison of the results obtained with the naive estimator and the KSG estimator, for both classifiers used, shows that the KSG estimator had an improved performance in terms of accuracy, sensitivity and specificity, thus revealing an efficient estimation method.

**Keywords.** Laryngeal Pathologies, Dynamic Systems, Mutual Information, KSG Estimator.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>2</b>
1.1	Motivação . . . . .	3
1.2	Revisão Bibliográfica e Estado da Arte . . . . .	4
1.3	Principais Contribuições . . . . .	8
1.4	Organização da Tese . . . . .	9
<b>2</b>	<b>Análise Acústica dos Sinais de Voz</b>	<b>10</b>
2.1	O Processo de Produção da Voz . . . . .	10
2.2	Análise Acústica dos Sinais de Voz . . . . .	12
2.3	Patologias Laríngeas . . . . .	14
2.3.1	Nódulos nas Pregas Vocais . . . . .	14
2.3.2	Edema de Reinke . . . . .	15
2.3.3	Paralisia das Pregas Vocais . . . . .	16
2.4	Considerações Finais do Capítulo . . . . .	17
<b>3</b>	<b>Fundamentação Teórica</b>	<b>18</b>
3.1	Sistemas Dinâmicos e Processamento de Voz . . . . .	19
3.2	Sistemas Dinâmicos . . . . .	19
3.2.1	O Exemplo do Pêndulo Simples . . . . .	20
3.2.2	Sistemas Dinâmicos não-lineares . . . . .	22
3.2.3	O método da Informação Mútua Média . . . . .	23
3.3	Medidas da Análise Dinâmica Não-linear . . . . .	25
3.3.1	Primeiro Mínimo da Função de Informação Mútua (PMIM) e atraso de Reconstrução $\tau$ . . . . .	25
3.3.2	Entropia de Shannon . . . . .	28
3.3.3	Entropia de Correlação . . . . .	32
3.4	Considerações Finais do Capítulo . . . . .	33
<b>4</b>	<b>Materiais e Métodos</b>	<b>34</b>
4.1	Base de Dados . . . . .	34
4.2	Metodologia . . . . .	34
4.3	Descrição dos Classificadores . . . . .	35
4.3.1	Análise Discriminante . . . . .	35
4.3.2	Máquinas de Vetor de Suporte - SVM . . . . .	37
4.4	Avaliação e Interpretação . . . . .	38
4.5	Considerações Finais do Capítulo . . . . .	39

<b>5</b>	<b>Resultados</b>	<b>40</b>
5.1	Teste dos Estimadores . . . . .	40
5.1.1	Medida $\tau$ . . . . .	40
5.1.2	Entropia de Shannon . . . . .	44
5.2	Classificação de Sinais de Vozes Saudáveis <i>Versus</i> Patológicas Usando Medidas da Análise Não-Linear . . . . .	47
5.2.1	Classificação entre Vozes Saudáveis <i>versus</i> Vozes Afetadas por Patologias Laríngeas . . . . .	47
5.2.2	Classificação entre Vozes Saudáveis <i>versus</i> Vozes Afetadas pela Patologia Edema de Reinke . . . . .	50
5.2.3	Classificação entre Vozes Saudáveis <i>versus</i> Vozes Afetadas pela Patologia Nódulo nas Pregas Vocais . . . . .	52
5.2.4	Classificação entre Vozes Saudáveis <i>versus</i> Vozes Afetadas pela Patologia Paralisia das Pregas Vocais . . . . .	55
5.3	Classificação de Sinais de Vozes Patológicas Usando Medidas da Análise Não-Linear . . . . .	59
5.3.1	Classificação entre Vozes Afetadas por Paralisia das Pregas Vocais <i>versus</i> Vozes Afetadas por Nódulo nas Pregas Vocais . . . . .	60
5.3.2	Classificação entre Vozes Afetadas por Paralisia das Pregas Vocais <i>versus</i> Vozes Afetadas pela Patologia Edema de Reike . . . . .	63
5.3.3	Classificação entre Vozes Afetadas por Edema de Reike nas Pregas Vocais <i>versus</i> Vozes Afetadas por Nódulo nas Pregas Vocais . . . . .	65
<b>6</b>	<b>Conclusões e Trabalhos Futuros</b>	<b>70</b>
6.1	Resultados dos testes Experimentais . . . . .	70
6.2	Classificação de Sinais de Vozes Saudáveis <i>Versus</i> Patológicas Usando Medidas da Análise Não-Linear . . . . .	71
6.3	Classificação de Sinais de Vozes Patológicas Usando Medidas da Análise Não-Linear . . . . .	71
6.4	Perspectiva de Trabalhos Futuros . . . . .	72
<b>A</b>	<b>Conceitos de Probabilidade e Estatística</b>	<b>81</b>
<b>B</b>	<b>Testes Estatísticos</b>	<b>84</b>
<b>C</b>	<b>Lista de Publicações do Doutorado</b>	<b>86</b>

# Lista de Figuras

2.1	Diagrama esquemático do mecanismo de produção de fala humana. . . .	11
2.2	Pregas vocais em: (a) adução e (b) abdução - visão endoscópica. . . . .	11
2.3	Imagens da laringe adulta, obtidas por telelaringoscopia. A. Durante a respiração. B. Durante a fonação. . . . .	12
2.4	Diagrama de blocos da produção da voz humana. . . . .	13
2.5	Nódulos bilaterais nas pregas vocais. . . . .	15
2.6	Edema de Reinke bilateral . . . . .	16
2.7	Paralisia unilateral da prega direita durante a fonação. . . . .	16
3.1	O Pêndulo Simples. . . . .	20
3.2	Diagrama de fases para o pêndulo simples. . . . .	21
3.3	Atrator de Lorenz, com parâmetros $\sigma = 10$ , $r = 28$ e $b = 8/3$ . . . . .	22
3.4	(a) Curva de Informação Mútua para obtenção do $\tau$ ótimo e PMIM para um sinal normal e (b) Espaço de fase reconstruído para o sinal da curva (a), com $\tau$ ótimo = 6 e dimensão de imersão $m = 2$ . . . . .	24
3.5	Exemplos de espaço de fase reconstruído (projeção em duas dimensões) para sinais de vozes: (A) Saudável, (B) Com Edema de Reinke, (C) Com Nódulo e (D) com Paralisia. . . . .	24
3.6	Determinação de $\epsilon_n$ , $n_x$ , e $n_y$ para o algoritmo KSG. . . . .	27
3.7	Representação gráfica da interpretação do estimador Kozachenko-Leonenko considerando $k = 3$ então têm-se pelo menos uma amostra com distância $r \in [\epsilon/2, \epsilon/2 + d\epsilon/2]$ até $x_i$ (em azul), $k - 1$ amostras com distâncias menores que $r$ (em vermelho) e, $N - k - 1$ amostras com distâncias maiores que $r$ (branco). . . . .	31
4.1	Diagrama em blocos da metodologia empregada. . . . .	35
4.2	Função discriminante linear em um espaço de características arbitrário. .	36
4.3	Função discriminante quadrática em um espaço de característica arbitrário. 36	
5.1	Distribuição média dos valores de $\tau$ para os sinais de vozes saudáveis (SDL) e patológicos com as patologias edema de Reinke (EDM), nódulo (NDL) e paralisia das pregas vocais (PRL), estimadas com o estimador ingênuo. .	41
5.2	Distribuição média dos valores de $\tau$ para os sinais de vozes saudáveis (SDL) e patológicos com as patologias edema de Reinke (EDM), nódulo (NDL) e paralisia das pregas vocais (PRL), estimadas com o estimador KSG. . . .	41
5.3	Estimativas de informações mútuas usando o estimador ingênuo. . . . .	42
5.4	Estimativas de informações mútuas usando o estimador KSG/KOLE. . .	43

5.5	Distribuição média dos valores de entropia de Shannon para os sinais de vozes saudáveis (SDL) e patológicos com as patologias edema de Reinke (EDM), nódulo (NDL) e paralisia das pregas vocais (PRL), estimadas com o estimador ingênuo. . . . .	44
5.6	Distribuição média dos valores de entropia de Shannon para os sinais de vozes saudáveis (SDL) e patológicos com as patologias edema de Reinke (EDM), nódulo (NDL) e paralisia das pregas vocais (PRL), estimadas com o estimador KOLE. . . . .	45
5.7	Estimativas de entropia de Shannon usando os estimadores KOLE e ingênuo.	45
5.8	Classificação Saudável x Nódulo - Medidas Individuais . . . . .	58
5.9	Classificação Saudável x Paralisia - Medidas Individuais . . . . .	59
5.10	Classificação Saudável x Nódulo - Medidas Combinadas . . . . .	59
5.11	Classificação Edema x Nódulo - Medidas Combinadas . . . . .	69

# Lista de Tabelas

1.1	Taxas de Classificação (%) Vozes Saudáveis x Patológicas - $K_2$ . . . . .	8
4.1	Matriz de confusão em um teste de detecção da presença/ausência de doença.	38
5.1	Resultados dos testes estatísticos utilizando <i>Wilcoxon signed-rank test</i> . . .	43
5.2	Resultados dos testes estatísticos utilizando <i>Wilcoxon ranksum test</i> . . .	44
5.3	Resultados dos testes estatísticos utilizando <i>Wilcoxon signed-rank test</i> . . .	46
5.4	Resultados dos testes estatísticos utilizando <i>Wilcoxon ranksum test</i> . . .	46
5.5	Taxas de Classificação (%) Saudável x Patológico - Estimador Ingênuo. . .	47
5.6	Taxas de Classificação (%) Saudável x Patológico - Estimador KSG/KOLE.	48
5.7	Taxa de Classificação (%) Saudável x Patológico - Estimador Ingênuo. . .	48
5.8	Taxa de Classificação (%) Saudável x Patológico - Estimador KSG/KOLE.	48
5.9	Taxas de Classificação (%) Saudável x Patológico. Medidas Combinadas - Estimador Ingênuo. . . . .	49
5.10	Taxas de Classificação (%) Saudável x Patológico. Medidas Combinadas - Estimador KSG/KOLE. . . . .	49
5.11	Taxa de Classificação (%) Saudável x Patológico. Medidas Combinadas - Estimador Ingênuo. . . . .	49
5.12	Taxa de Classificação (%) Saudável x Patológico. Medidas Combinadas - Estimador KSG/KOLE. . . . .	50
5.13	Taxas de Classificação (%) Saudável x Edema - Estimador Ingênuo. . . .	50
5.14	Taxas de Classificação (%) Saudável x Edema - Estimador KSG/KOLE.	50
5.15	Taxa de Classificação (%) Saudável x Edema - Estimador Ingênuo. . . . .	51
5.16	Taxa de Classificação (%) Saudável x Edema - Estimador KSG/KOLE. . .	51
5.17	Taxas de Classificação (%) Saudável x Edema. Medidas Combinadas - Estimador Ingênuo. . . . .	51
5.18	Taxas de Classificação (%) Saudável x Edema. Medidas Combinadas - Estimador KSG/KOLE. . . . .	52
5.19	Taxa de Classificação (%) Saudável x Edema. Medidas Combinadas - Estimador Ingênuo. . . . .	52
5.20	Taxa de Classificação (%) Saudável x Edema. Medidas Combinadas - Estimador KSG/KOLE. . . . .	52
5.21	Taxas de Classificação (%) Saudável x Nódulo - Estimador Ingênuo. . . .	53
5.22	Taxas de Classificação (%) Saudável x Nódulo - Estimador KSG/KOLE.	53
5.23	Taxa de Classificação (%) Saudável x Nódulo - Estimador Ingênuo. . . .	53
5.24	Taxa de Classificação (%) Saudável x Nódulo - Estimador KSG/KOLE. . .	54
5.25	Taxas de Classificação (%) Saudável x Nódulo. Medidas Combinadas - Estimador Ingênuo. . . . .	54

5.26	Taxas de Classificação (%) Saudável x Nódulo. Medidas Combinadas - Estimador KSG/KOLE. . . . .	54
5.27	Taxa de Classificação (%) Saudável x Nódulo. Medidas Combinadas - Estimador Ingênuo. . . . .	55
5.28	Taxa de Classificação (%) Saudável x Nódulo. Medidas Combinadas - Estimador KSG/KOLE. . . . .	55
5.29	Taxas de Classificação (%) Saudável x Paralisia - Estimador Ingênuo. . .	56
5.30	Taxas de Classificação (%) Saudável x Paralisia - Estimador KSG/KOLE.	56
5.31	Taxa de Classificação (%) Saudável x Paralisia - Estimador Ingênuo. . . .	56
5.32	Taxa de Classificação (%) Saudável x Paralisia - Estimador KSG/KOLE.	56
5.33	Taxas de Classificação (%) Saudável x Paralisia. Medidas Combinadas - Estimador Ingênuo. . . . .	57
5.34	Taxas de Classificação (%) Saudável x Paralisia. Medidas Combinadas - Estimador KSG/KOLE. . . . .	57
5.35	Taxa de Classificação (%) Saudável x Paralisia. Medidas Combinadas - Estimador Ingênuo. . . . .	57
5.36	Taxa de Classificação (%) Saudável x Paralisia. Medidas Combinadas - Estimador KSG/KOLE. . . . .	58
5.37	Taxas de Classificação (%) Paralisia x Nódulo - Estimador Ingênuo. . . .	60
5.38	Taxas de Classificação (%) Paralisia x Nódulo - Estimador KSG/KOLE.	61
5.39	Taxas de Classificação (%) Paralisia x Nódulo - Estimador Ingênuo. . . .	61
5.40	Taxa de Classificação (%) Paralisia x Nódulo - Estimador KSG/KOLE. .	61
5.41	Taxas de Classificação (%) Paralisia x Nódulo. Medidas Combinadas - Estimador Ingênuo. . . . .	62
5.42	Taxas de Classificação (%) Paralisia x Nódulo. Medidas Combinadas - Estimador KSG/KOLE. . . . .	62
5.43	Taxa de Classificação (%) Paralisia x Nódulo. Medidas Combinadas - Estimador Ingênuo. . . . .	62
5.44	Taxa de Classificação (%) Paralisia x Nódulo. Medidas Combinadas - Estimador KSG/KOLE. . . . .	62
5.45	Taxas de Classificação (%) Paralisia x Edema - Estimador Ingênuo. . . .	63
5.46	Taxas de Classificação (%) Paralisia x Edema - Estimador KSG/KOLE. .	63
5.47	Taxa de Classificação (%) Paralisia x Edema - Estimador Ingênuo. . . .	64
5.48	Taxa de Classificação (%) Paralisia x Edema - Estimador KSG/KOLE. .	64
5.49	Taxas de Classificação (%) Paralisia x Edema. Medidas Combinadas - Estimador Ingênuo. . . . .	64
5.50	Taxas de Classificação (%) Paralisia x Edema. Medidas Combinadas - Estimador KSG/KOLE. . . . .	65
5.51	Taxa de Classificação (%) Paralisia x Edema. Medidas Combinadas - Estimador Ingênuo. . . . .	65
5.52	Taxa de Classificação (%) Paralisia x Edema. Medidas Combinadas - Estimador KSG/KOLE. . . . .	65
5.53	Taxas de Classificação (%) Edema x Nódulo - Estimador Ingênuo. . . . .	66
5.54	Taxas de Classificação (%) Edema x Nódulo - Estimador KSG/KOLE. .	66
5.55	Taxa de Classificação (%) Edema x Nódulo - Estimador Ingênuo. . . . .	67
5.56	Taxa de Classificação (%) Edema x Nódulo - Estimador KSG/KOLE. . .	67
5.57	Taxas de Classificação (%) Edema x Nódulo. Medidas Combinadas - Estimador Ingênuo. . . . .	67

5.58	Taxas de Classificação (%) Edema x Nódulo. Medidas Combinadas - Estimador KSG/KOLE. . . . .	68
5.59	Taxa de Classificação (%) Edema x Nódulo. Medidas Combinadas - Estimador Ingênuo. . . . .	68
5.60	Taxa de Classificação (%) Edema x Nódulo. Medidas Combinadas - Estimador KSG/KOLE. . . . .	68

# Lista de Siglas e Símbolos

$Ac$  – Medida de acurácia

EDM – Vozes com a patologia edema de Reinke nas pregas vocais

$Esp$  – Medida de especificidade

$H$  – Entropia de Shannon

$k_2$  - Entropia de correlação

KOLE – Estimador de Kozachenko-Leonenko

KSG – Estimador de Kraskov et. al.

LDA – Análise discriminante linear

LPC – Coeficiente de predição linear

MNL – Medidas de análise não-linear

MQR – Medidas de quantificação de recorrência

NDL – Vozes com a patologia nódulo nas pregas vocais

$p$  – Nível de significância em um teste de hipóteses

PMIM – Primeiro mínimo da informação mútua

PRL – Vozes com a patologia paralisia das pregas vocais

PTL – Vozes patológicas

SDL – Vozes saudáveis

$Sen$  – Medida de sensibilidade

SVM – Máquinas de vetor de suporte

$\tau$  –Atraso de reconstrução da dinâmica não-linear

# Notação e Terminologia

A fim de permitir um melhor entendimento desta tese, uma notação padrão foi adotada. As variáveis aleatórias são escritas com letras maiúsculas e valores particulares que elas assumem com letras minúsculas. Processos aleatórios são escritos com letras em maiúsculo e em negrito.

Séries de variáveis aleatórias são indexadas com o primeiro índice no subscrito e o último índice no sobrescrito, por exemplo, a série de variáveis  $X_1, X_2, X_3, X_4, X_5 = X_1^5$ . É possível também que o subscrito seja omitido quando o primeiro índice da série da variável aleatória for 1, ou seja, no exemplo anterior,  $X_1^5 = X^5$ . O símbolo  $\mathbb{E}X$  denota esperança de uma variável aleatória  $X$ . Uma barra sobre uma variável aleatória denota sua média amostral, isto é:

$$\bar{X} = \frac{1}{N} \sum_{n=1}^N X_n$$

O símbolo  $var(X)$  denota a variância de  $X$ . A função logarítmica, quando escrita como  $\log$ , está na base 2 (caso em que as medidas de informação são feitas em bits). Quando a função logarítmica for escrita  $\ln$ , ela está na base neperiana (caso em que as medidas de informação são feitas em nats).

# Capítulo 1

## Introdução

O homem utiliza diversos meios de comunicação para desenvolver a sua capacidade intelectual e o seu meio social. A fala é a principal ferramenta de comunicação entre as pessoas, pois com ela é possível expressar os sentimentos e ideias, além de possibilitar a troca de informações.

Do ponto de vista fisiológico, a voz humana é o resultado da interação de órgãos de diferentes sistemas do corpo humano[1], um conjunto de estruturas do trato vocal, cujas partes mais intimamente associadas à produção do som são os pulmões, a traqueia, a laringe, a faringe, as cavidades nasais e a cavidade oral [2].

O sistema vocal, apesar de pequeno, comparado com os outros sistemas do corpo humano, possui uma capacidade de produção complexa e potente. Sua representação máxima está focada nas pregas vocais. O trato vocal atua como um filtro e suas frequências de ressonância designam-se por formantes. As vogais são reconhecidas pelos seus formantes, que são produzidos em nível glótico e modificados pelos ajustes específicos do trato vocal [3]. O tamanho e o formato do trato vocal são fatores determinantes das características do som a ser emitido e dependem diretamente da idade e sexo do falante [4].

A voz é considerada uma função multidimensional e, por isso, em sua avaliação, devem ser considerados aspectos fisiológicos, perceptivo-auditivos, acústicos e aerodinâmicos [5]. O sinal vocal é considerado um produto complexo da interação não linear entre as propriedades aerodinâmicas e biomecânicas do sistema de produção da voz, havendo, então, uma grande correspondência entre a fisiologia laríngea e as medidas acústicas [6].

Os distúrbios da voz e as patologias laríngeas, podem afetar diferentes grupos etários. Muitos deles o ser humano traz consigo desde o seu nascimento, sendo diagnosticados ainda na infância, por meio da identificação de dificuldades respiratórias ou choro anormal ou de forma tardia, por meio de manifestações sutis que ocorrem ao longo do crescimento [7, 8].

No decorrer dos últimos anos, a análise da qualidade vocal por intermédio do uso de medidas da análise acústica, empregando técnicas de processamento digital de sinais, vem sendo amplamente estudada, na busca por uma medida, ou um conjunto de medidas, que sejam capazes de revelar ou predizer as condições anatômicas e fisiológicas das estruturas laríngeas [9, 10].

Devido a sua não estacionariedade e sua dinâmica caótica, os sinais de voz não podem ser corretamente analisados a partir de métodos tradicionais de análises de sinais experimentais como, a função de autocorrelação e espectro de potência [11]. A teoria do caos, uma área da teoria dos sistemas dinâmicos não-lineares, aplicada em séries temporais não-lineares, foi recentemente adotada como uma nova abordagem não linear do

processamento do sinal de voz [12].

Nesta proposta, as medidas não-lineares abordadas foram: ( $\tau$ ), primeiro mínimo da função de informação mútua (PMIM), entropia de Shannon ( $H$ ) e entropia de correlação ( $K_2$ ). Para se obter essas medidas faz-se necessário conhecer inicialmente a distribuição de probabilidade. Entretanto, muitas vezes essa distribuição e/ou seus parâmetros não são conhecidos, fazendo-se necessário sua estimação.

O método de estimação comumente empregado utiliza o modelo baseado em histogramas, mais conhecido na literatura como estimador ingênuo, que apesar de ser bastante difundido e amplamente utilizado, apresenta uma forte dependência da dimensionalidade dos dados e erros sistemáticos, produzindo um viés de estimação, que em algumas situações, difere do valor de informação mútua obtido analiticamente [91] [92] [93].

Nesta proposta apresentamos o modelo de estimação desenvolvido por Kozachenko e Leonenko [94] para entropia, mais tarde adaptado por Kraskov et. al. [38] para cálculo da informação mútua. Os estimadores propostos são baseados na distribuição média de probabilidade dos vizinhos de cada ponto da amostra, produzindo estimativas de medidas de informação mais próximas do valor analiticamente calculado e influenciando diretamente nos resultados das classificações entre patologias laríngeas [91] [92].

Cinco classes de sinais foram consideradas neste estudo: saudável (SDL), paralisia (PRL), edema (EDM), nódulos (NDL) e todas as patologias juntas (PTL). São considerados sete casos de discriminação: SDL  $\times$  PTL, SDL  $\times$  PRL, SDL  $\times$  EDM, SDL  $\times$  NDL, PRL  $\times$  EDM, PRL  $\times$  NDL, e EDM  $\times$  NDL. Foram utilizados também dois classificadores distintos, um baseado em análise discriminante (LDA) e outro em máquinas de vetor de suporte (SVM).

Entre os resultados obtidos podemos citar a classificação EDM  $\times$  NDL, onde a combinação das medidas  $\tau$  e  $PMIM$  com o estimador KSG/KOLE e classificador SVM, que apresentou taxa de acurácia de 80.7692%, contra 53.8461% com o estimador ingênuo. Na classificação SDL  $\times$  NDL, utilizando a medida  $K_2$  individualmente com o estimador KSG/KOLE e classificador LDA, foi obtida uma taxa de acurácia de 84.64%, contra 41.96% com o estimador ingênuo.

## 1.1 Motivação

Existe uma grande variedade de patologias laríngeas que causam modificações na voz. Essas patologias podem ser tanto de origem orgânica como nódulos, cistos ou edemas, quanto de origem neurológica, tal como paralisia nas pregas vocais [13, 14]. Algumas dessas patologias, a exemplo dos nódulos e edema de Reinke, podem ser adquiridas devido a hábitos sociais não-saudáveis, como alcoolismo e tabagismo, além do abuso vocal, como gritos excessivos, fala excessiva, entre outros. No caso de paralisia, as principais causas incluem traumas cirúrgicos ou acidentais, doenças cardiovasculares e neurológicas [15].

Estudos realizados em diversos países, incluindo o Brasil, com profissionais que trabalham com a voz, tais como professores, telefonistas, cantores, atores, leiloeiros, políticos, religiosos, atendentes de telemarketing e bancários, indicam que as patologias nódulos, edema de Reinke e paralisia nas pregas vocais estão entre as patologias de maior incidência em adultos, tanto do gênero masculino quanto do gênero feminino [16, 17]. Entretanto, a literatura aponta os professores como o grupo mais vulnerável [18].

O diagnóstico da qualidade vocal inicialmente é feito pelo otorrinolaringologista, que inclui a anamnese, seguido de exames físicos e visual da laringe, a exemplo da videolarin-

goscopia direta, videoestroboscopia e eletromiografia, exames esses de caráter invasivo, que podem trazer desconforto ao paciente [19].

A videolaringoscopia direta é um exame realizado pelo médico com o objetivo de visualizar a laringe utilizando uma microcâmera. A videoestroboscopia permite a visualização do comportamento vibratório das pregas vocais, e a eletromiografia é um método de registro dos potenciais elétricos gerados nas fibras musculares em ação. Essas técnicas visuais resultam em uma avaliação qualitativa, de resultados difíceis de serem quantificados, e necessitam do conhecimento e da experiência do avaliador [2] [20].

Devido a esse fato, técnicas baseadas na análise acústica vocal tem despertado o interesse por parte dos pesquisadores, tanto para o acompanhamento na melhoria da qualidade vocal, bem como o auxílio ao diagnóstico médico. Essas, são técnicas automáticas de auxílio ao diagnóstico, menos invasivas e de baixo custo, quando comparadas àquelas baseadas em exames videolaringoscópicos [19], possibilitando, ainda, um tratamento à distância - caso em que a voz do paciente pode ser, por exemplo, gravada em um computador pessoal e enviada para um médico em outra localidade, e este avaliar a qualidade da voz, efetuar pré-diagnóstico de patologias laríngeas, bem como a evolução de um tratamento medicamentoso ou pós-cirúrgico.

Ao longo das últimas décadas, observações dos modelos computacionais das pregas vocais [21, 22, 23, 24], experimentos com laringes excitadas [25, 26] e análise dinâmica não-linear de vozes humanas [27, 28, 2, 29] estabeleceram a existência do caos na produção da voz.

Para que se tenha um sistema computacional de auxílio a diagnósticos eficiente que, seja capaz de discriminar corretamente entre tipos de patologias diferentes, faz-se necessário a escolha das características e/ou parâmetros que melhor representem as desordens da voz ocasionadas por uma dada patologia [11]. A literatura ainda não traz um consenso acerca das medidas de maior acurácia na avaliação dessas desvios vocais. Por isso, se fazem necessários estudos que possam revelar o poder de discriminação das medidas acústicas de maneira isolada e/ou combinadas para serem empregadas na discriminação entre vozes saudáveis/patológicas.

Uma classificação eficiente pode auxiliar o terapeuta a avaliar o quanto a terapia está sendo efetiva, de forma objetiva. Para tanto, é necessário que a técnica proposta tenha confiabilidade e apresente as informações das mudanças ocorridas no sinal antes e após a terapia vocal, necessárias para um diagnóstico mais preciso.

Apesar de haver muitos trabalhos relacionados à identificação de distúrbios da voz, não há uma confirmação precisa de uma medida ou um conjunto de medidas que seja capaz de encontrar os parâmetros mais adequados para modelagem de uma patologia em particular [20] [29] [12] [30].

A discriminação entre patologias laríngeas ainda é objeto de investigação por parte dos pesquisadores. Portanto, o estudo de técnicas de análise acústica não-linear é uma área bastante promissora, uma vez que a interdisciplinaridade dos procedimentos pode proporcionar a investigação com mais precisão de patologias laríngeas [28].

## 1.2 Revisão Bibliográfica e Estado da Arte

Nesta seção, são apresentados os trabalhos relevantes encontrados na literatura que utilizam a análise não linear, na discriminação de vozes patológicas e que serviram como embasamento para o desenvolvimento desta tese.

Apesar de seu uso relativamente recente, a teoria do caos, uma área da teoria dos sistemas de dinâmicos não-lineares, aplicada em séries temporais não-lineares, tem sido adotada como uma nova abordagem não linear do processamento do sinal de voz e tem apresentado resultados significativos [27, 31, 11, 12, 32, 33, 34].

Diversos métodos têm sido propostos na literatura com a tarefa de classificar patologias laríngeas empregando análise dinâmica não-linear. No entanto, observa-se que determinado método ou característica pode apresentar um bom desempenho para classificar entre vozes saudáveis e patológicas, porém não apresentam igual desempenho na diferenciação entre vozes patológicas. Além disso, os métodos utilizados empregam um modelo de estimação ingênuo que pode produzir resultados pouco precisos e/ou enviesados.

Desta forma, a busca por características e métodos mais precisos e eficientes para uma discriminação entre patologias laríngeas com níveis de precisão mais confiáveis ainda é fruto de diversas pesquisas. Nesta seção, será apresentada uma revisão bibliográfica dos trabalhos que também utilizam a análise dinâmica não-linear na classificação e discriminação de desordens vocais e patologias laríngeas com fins de diagnóstico.

Pinho *et al.* [27], em seu trabalho utilizam informações das imagens do espaço de fase reconstruído dos sinais de voz, para classificação entre vozes normais e vozes afetadas por patologias laríngeas (edema de Reinke e paralisia nas pregas vocais). São empregadas sete características obtidas pelos métodos da contagem de caixas e da diferença, considerando atrasos diferentes e quatro dimensões de caixas: 40x40, 20x20, 10x10 e 5x5. A classificação é realizada com as medidas extraídas individualmente e de forma combinada, através de redes neurais artificiais. Os resultados obtidos mostram que o uso de características extraídas do espaço de fase para diagnosticar a presença de patologias nas pregas vocais é bastante interessante e promissor. Os autores observaram também que o tamanho da grade influenciou nos resultados, sendo que a grade 10x10 foi a que obteve melhor desempenho, comparada às outras grades. Quanto aos métodos, observou-se que o método da diferença foi mais eficiente para discriminar as classes do que o método da contagem de caixas, apresentando acurácia média acima de 94%, na discriminação entre vozes normais e patológicas. Na discriminação entre patologias (Edema vs. Paralisia), o desempenho foi inferior ao que foi obtido entre vozes saudáveis e patológicas. Os autores concluem ainda que o uso de outros o uso de outros classificadores, a exemplo de svm (Support Vector Machine – SVM) podem ser utilizados no futuro com o intuito de melhorar o desempenho na classificação.

Costa *et al.* [31], em seu trabalho são associadas abordagens lineares e não lineares para avaliar o desempenho na classificação utilizando-se 8 medidas oriundas da análise dinâmica não linear (dimensão de correlação, quatro medidas de entropia, expoente de Hurst, maior expoente de Lyapunov e o primeiro mínimo da função de informação mútua), além de coeficientes LPC, obtidos a partir da análise preditiva linear, de forma individual, e a partir da combinação das mesmas. Os resultados obtidos, constataam que as medidas não lineares conseguem capturar bem as desordens vocais provocadas por patologias nas pregas vocais, como nódulos, edemas e paralisia. A consideração das não linearidades inerentes ao processo de produção da fala e sua consequente análise utilizando medidas não lineares é eficiente em distinguir um sinal saudável de um sinal patológico, comprovadamente com as patologias consideradas no trabalho. Entretanto, os autores observam que a abordagem pelo modelo linear de produção da fala, em que os coeficientes LPC modelam o trato vocal, consegue distinguir melhor as especificidades existentes nas patologias analisadas, notadamente quando combinadas à abordagem não linear. As mudanças nas

estruturas anatômicas do trato vocal e da fonte (laringe) foram bem retratadas com a combinação das duas abordagens, sendo a principal contribuição deste trabalho. Com a abordagem híbrida, o desempenho na distinção entre as patologias obteve um aumento considerável. Isto denota que o método empregado é bastante promissor na tarefa de discriminar entre vozes saudáveis e vozes afetadas por patologias laríngeas, bem como entre patologias. Por fim, os autores concluem que os resultados obtidos sugerem a viabilidade da técnica empregada para a discriminação entre vozes saudáveis e patológicas em geral, como também entre vozes afetadas por patologias laríngeas distintas como edema, nódulos e paralisia nas pregas vocais.

Costa [11], em sua tese, emprega a análise dinâmica não linear (e teoria do caos), como também por meio a análise de quantificação de recorrência, na caracterização e a classificação de sinais de vozes saudáveis e vozes afetadas por diferentes patologias laríngeas (edema, paralisia e nódulos nas pregas vocais). A partir da combinação de características dos conjuntos das medidas de análise não linear (MNL) e das medidas de quantificação de recorrência (MQR), ao todo 15 medidas, as médias da taxa de acurácia obtidas variaram nos intervalos de confiança: [95,44%; 100%] para a classificação entre vozes saudáveis e patológicas; [94,75%; 100%] entre vozes saudáveis e afetadas por edema, e entre saudáveis e nódulos. Também são avaliados os efeitos do uso de vetores híbridos formados por características MNL, MQR e coeficientes extraídos da análise preditiva linear (LPC). Utilizando apenas a medida Primeiro Mínimo da função de informação mútua, as taxas de acurácia obtidas foram: 62,55% entre vozes saudáveis e patológicas. Entre vozes patológicas a maior taxa de acurácia foi de 66,43% entre vozes com paralisia e vozes com nódulo.

Henríquez *et al.* [12], propõem a caracterização de vozes saudáveis e patológicas através de medidas baseadas na análise dinâmica não linear. Eles usam seis medidas caóticas não-lineares na discriminação entre dois níveis de qualidade da voz: saudável e patológica. As medidas empregadas são entropia de Rényi de primeira e segunda ordem, entropia de correlação, dimensão de correlação, primeiro mínimo da função de informação mútua e entropia de Shannon. Foram utilizadas duas bases de dados diferentes, a *multiquality database* e *commercial database - MEEI Voice Disorders*. Um classificador baseado em redes neurais foi implementado a fim de avaliar as medidas propostas. Um estudo estatístico prévio foi realizado para verificar a discriminação das características dos dois bancos de dados. Para a base de dados *multiquality database* o estudo estatístico mostrou diferenças notáveis entre vozes saudáveis e patológicas e até mesmo entre os três diferentes níveis de patologias para cada um dos características estudadas. A dimensão de correlação e o valor do primeiro mínimo da função de informação mútua foram as características que melhor discriminaram entre as diferentes qualidades de voz da *commercial database - MEEI Voice Disorders*. As taxas de sucesso global obtidas foram: 82,47% *multiquality database* e 99,69% *commercial database - MEEI Voice Disorders*. As medianas das vozes saudáveis e patológicas diferem ao nível de significância de 5% em ambas as bases de dados e em todas as características consideradas. Os autores concluem que as medidas estudadas podem ser utilizadas para documentar a evolução de pacientes, além de também poderem ser usadas em sistemas de ajuda para diagnóstico de patologia no sistema de produção da fala.

Scalassara *et al.* [32], apresentam um estudo sobre o uso de medidas de entropia para dois grupos de sinais de voz: um grupo composto por sinais de vozes saudáveis e outro composto por sinais de vozes afetadas pela patologia nódulo nas pregas vocais. O método de estimação de entropia desenvolvido por (Moddemeijer, 1989) foi utilizado para esti-

mar a entropia de cada sinal de voz. Os resultados obtidos através de testes estatísticos (Teste T) mostraram que as classes eram claramente separáveis utilizando a medida de entropia. O grupo dos sinais afetados por nódulo nas pregas vocais apresentou maior valor de entropia quando comparado ao grupo de sinais de vozes saudáveis. Os resultados obtidos foram comparados com os valores de *jitter*, *shimmer* e *pitch* das amostras, que foram obtidos com um software comercial (*Voice Analysis version 6.0*). A variabilidade dos parâmetros para o grupo de sinais de vozes afetados por nódulo foi significativamente maior que para o grupo saudável, apresentando, portanto, um comportamento que se compara favoravelmente ao obtido pelo método da entropia. Os sinais de vozes saudáveis e afetados por nódulo foram analisados por meio de da técnica de reconstrução do espaço de fase, o que mostrou diferentes padrões visuais para cada grupo. Para determinar as características do padrão visual, três tipos de comportamento dinâmico das órbitas eram observado: a) número de voltas, b) regularidade do percurso do atrator, e c) distribuição das trajetórias dos atratores (divergência e convergência das trajetórias das órbitas dos atratores). Os autores concluem ainda que, a análise do espaço de fases ajuda a representar o padrão de vogais de maneira dinâmica. Esta técnica permite visualizar a dinâmica diferencial entre vozes saudáveis e vozes com nódulos de pregas vocais.

Falcão *et al.* [35], apresentam um estudo sobre a aplicação da entropia de Shannon, entropia relativa e a entropia de Tsallis para a classificação automática de vozes saudáveis ou patológicas, com edema, cistos, nódulos nas cordas vocais e paralisia dos nervos laríngeos. Os resultados obtidos mostraram que a entropia de Tsallis apresentou-se mais robusta na separação de vozes saudáveis e patológicas sendo, no geral, o modelo mais eficiente, pois apresenta uma maior probabilidade de correta detecção da patologia. Porém, a entropia relativa se revelou mais precisa na correta rejeição da presença de patologias, ou seja, apresenta um melhor diagnóstico de vozes saudáveis, com o menor índice de falsos positivos. Os resultados para a entropia de Shannon também foram significativos, principalmente no correto diagnóstico da presença de patologias. Segundo a avaliação dos autores, os resultados obtidos mostram que o uso da entropia para diagnosticar a presença de patologias nas pregas vocais é bastante interessante e promissor. As medidas de entropia utilizadas não se mostraram eficazes na separação dos diferentes tipos de patologia abordados.

Jiang *et al.* [33] realizaram um estudo comparativo entre a análise dinâmica não linear (dimensão de correlação) e análise de perturbação (*jitter* e *shimmer*) em termos dos efeitos do comprimento do sinal analisado, taxa de amostragem e nível de ruído. Nos experimentos foram utilizadas vogais sustentadas de vozes saudáveis e de vozes afetadas por patologias na laringe. Segundo os autores, as grandes variações nos valores das medidas de perturbação para os casos dos sinais aperiódicos e caóticos tornam inapropriado o uso dessas medidas para análise de tais sinais. Em contrapartida, a análise dinâmica não linear pode quantificar sinais caóticos. Além disso, a dimensão de correlação oferece uma medida de análise mais confiável para sinais quase periódicos de menor comprimento, taxa de amostragem mais baixa e maiores níveis de ruído.

Torres *et al.* [36], em seu trabalho utilizam diferentes medidas de complexidade a fim de caracterizar os sinais de vozes patológicas. Os autores realizaram um estudo comparativo da complexidade destes sinais com a de sinais de vozes saudáveis. As medidas utilizadas foram a medida de complexidade de Ziv-Lempel, a entropia de Shannon, a entropia de Tsallis, a entropia relativa de Kullback e suas derivadas (a entropia generalizada e o divergente). Os experimentos foram realizados considerando a estratificação das vozes em relação ao gênero. Os resultados de testes estatísticos mostram que as medidas

utilizadas não oferecem o mesmo potencial discriminativo para ambos os sexos. A entropia de Shannon e a medida de complexidade de Ziv-Lempel não apresentaram diferenças estatisticamente significativas em nenhum dos casos. A entropia de Tsallis e as entropias derivadas da entropia relativa apresentaram diferenças estatisticamente para ambos os grupos ( $P < 0,001$ ). A entropia relativa apresentou diferenças significativas apenas no caso do grupo de vozes masculinas.

Todos os trabalhos supracitados utilizam o método de estimação ingênuo já discutido no capítulo anterior. O trabalho de Costa [11] se relaciona de forma particular com esta tese por utilizar a mesma base de dados utilizada neste trabalho, bem como algumas medidas da análise dinâmica não linear. Na tabela 1.1 são apresentados alguns resultados obtidos por Costa [11] utilizando o estimador ingênuo comparando com os resultados obtidos nesta tese utilizando o estimador KSG:

Tabela 1.1: Taxas de Classificação (%) Vozes Saudáveis x Patológicas -  $K_2$ .

Vozes Saudáveis x Patológicas - $K_2$ .		
Casos de Classificação	<i>KSG</i>	[11]
Saudável × Patológica	73,75 ± 3,65	52,88 ± 5,09
Saudável × Edema	75,22 ± 2,79	60,97 ± 10,32
Saudável × Nódulo	84,64 ± 4,39	57,38 ± 8,81
Saudável × Paralisia	80,90 ± 3,10	61,96 ± 11,981

### 1.3 Principais Contribuições

Nesta seção são apresentadas as principais contribuições desta tese:

1. O uso de um estimador mais acurado, o KSG e suas variantes, na estimação das medidas da análise dinâmica não linear: atraso de reconstrução da dinâmica não-linear ( $\tau$ ), primeiro mínimo da função de informação mútua (PMIM), entropia de Shannon (H) e entropia de correlação ( $k_2$ );
2. Uso da medida ( $\tau$ ) na classificação de vozes saudáveis e patológicas de maneira individual e combinada às outras medidas consideradas nesta tese;
3. Comparação entre métodos de estimação de medidas de informação em termos de acurácia.
4. Utilização das medidas acima citadas de maneira individual e combinada na classificação das patologias laríngeas investigadas.

O primeiro item investiga o papel do estimador o KSG e suas variantes, na estimação das medidas da análise dinâmica não linear: atraso de reconstrução da dinâmica não-linear ( $\tau$ ), primeiro mínimo da função de informação mútua (PMIM), entropia de Shannon (H) e entropia de correlação ( $k_2$ ). O objetivo principal é observar, se o uso de um estimador mais robusto (menos enviesado) interfere no resultado da aplicação considerada nesta tese (discriminação entre vozes saudáveis e patológicas).

O segundo item ainda não foi abordado na literatura (não foram encontrados trabalhos que utilizem o parâmetro  $\tau$  combinado a outras medidas na classificação de vozes saudáveis e patológicas e entre patologias). O parâmetro  $\tau$  é utilizado em sistemas dinâmicos para permitir a reconstrução da dinâmica do sistema em estudo a partir da variável que se tem conhecimento, no caso dos sinais de vozes, a partir do conhecimento da série temporal, é possível reconstruir um espaço de fase difeomórficamente<sup>1</sup> equivalente ao original.

O terceiro item compara o desempenho de dois métodos de estimação de medidas de informação: o KSG e suas variantes desenvolvido por [38], baseado no método dos vizinhos mais próximos e o estimador ingênuo, comumente utilizado na literatura, baseado na técnica de segmentação e contagem.

Por fim, o quarto item tem a finalidade de investigar o poder discriminatório das medidas da análise dinâmica não linear ( $\tau$ ), primeiro mínimo da função de informação mútua (PMIM), entropia de Shannon (H) e entropia de correlação ( $k_2$ ) na discriminação entre sinais de vozes patológicas. A discriminação entre vozes saudáveis e patológicas já é bastantes difundida na literatura, no entanto, ainda não se tem um consenso sobre a medida ou o conjunto de medidas capazes de separar diferentes classes patológicas de maneira eficiente.

## 1.4 Organização da Tese

Este documento está organizado da seguinte forma: o Capítulo 2 trata da análise acústica dos sinais de vozes, descrevendo o mecanismo de produção da fala e as patologias laríngeas consideradas nesta tese. No Capítulo 3 é apresentada a teoria na qual se baseia esta pesquisa, são introduzidos os conceitos de sistemas dinâmicos e apresentados os estimadores utilizados. No Capítulo 4 é apresentada a metodologia empregada na pesquisa. No Capítulo 5, estão apresentados os resultados obtidos da classificação entre sinais de vozes saudáveis e sinais de vozes afetados por patologias, bem como entre sinais afetados por patologias, e por fim no Capítulo 6 encontra-se as conclusões e trabalhos futuros.

---

<sup>1</sup>Na matemática, um difeomorfismo é um isomorfismo na categoria das variedades diferenciáveis. Ele é uma função invertível que leva uma variedade diferenciável em outra, de modo que tanto a função quanto sua inversa sejam suaves. Duas variedades diferenciáveis dizem-se difeomorfas se existir uma aplicação entre essas variedades que seja diferenciável, invertível e a sua inversa seja diferenciável [37].

# Capítulo 2

## Análise Acústica dos Sinais de Voz

Do ponto de vista fisiológico, a voz humana é o resultado da interação de órgãos de diferentes sistemas do corpo humano [1], um conjunto de estruturas do trato vocal, cujas partes mais intimamente associadas à produção do som são os pulmões, a traqueia, a laringe, a faringe as cavidades nasais e a cavidade oral [2].

O trato vocal possui uma capacidade de produção complexa e potente. Sua representação máxima está focada nas pregas vocais. A voz é utilizada tanto para comunicação, quanto para expressar emoções, pensamentos e sentimentos, para satisfazer suas necessidades, além de representar a identidade de cada indivíduo, sendo considerada tão pessoal quanto à impressão digital.

Neste capítulo, são apresentados diversos aspectos da voz, tais como o processo de produção, os conceitos de voz saudável e voz patológica, análise acústica dos sinais de voz, as medidas acústicas do sinal de voz mais comumente utilizadas e as patologias laríngeas consideradas nesta proposta.

### 2.1 O Processo de Produção da Voz

A fonação é uma função neurofisiológica inata, mas a voz forma-se ao longo da vida, baseada nas características anatomofuncionais do indivíduo, bem como nos aspectos emocionais de sua história pessoal. Assim sendo, a voz é o resultado da fonação acrescida de ressonância [1].

A Figura 2.1 ilustra a anatomia do aparelho fonador. Os pulmões, brônquios e traqueia produzem o “ar”, matéria prima da produção vocal; a laringe (onde se encontram as pregas vocais) produz a energia da fala e, a faringe, fossas nasais e boca são responsáveis pela ressonância.

Os sons sonoros ocorrem quando o fluxo de ar sai dos pulmões e atinge a traqueia até alcançar a laringe, produzindo uma vibração nas pregas vocais. Diferente dos sons sonoros, os sons surdos não provocam vibrações, pois quando o fluxo de ar atinge a traqueia as pregas vocais estão relaxadas.

Na produção de sons orais, o véu palatino está levantado e o fluxo de ar é irradiado pela boca e na produção de sons nasais o véu palatino está abaixado e a cavidade oral fechada (lábios, dentes, palato), assim, o fluxo de ar é radiado pelas narinas [40].

A laringe é um órgão tubular, um arcabouço esquelético membranoso, situada no plano mediano e anterior superficial do pescoço. Comunica-se inferiormente com a traqueia e superiormente com a faringe [41]. As funções básicas da laringe, em ordem de importância são proteção das vias aéreas inferiores, respiração e fonação.

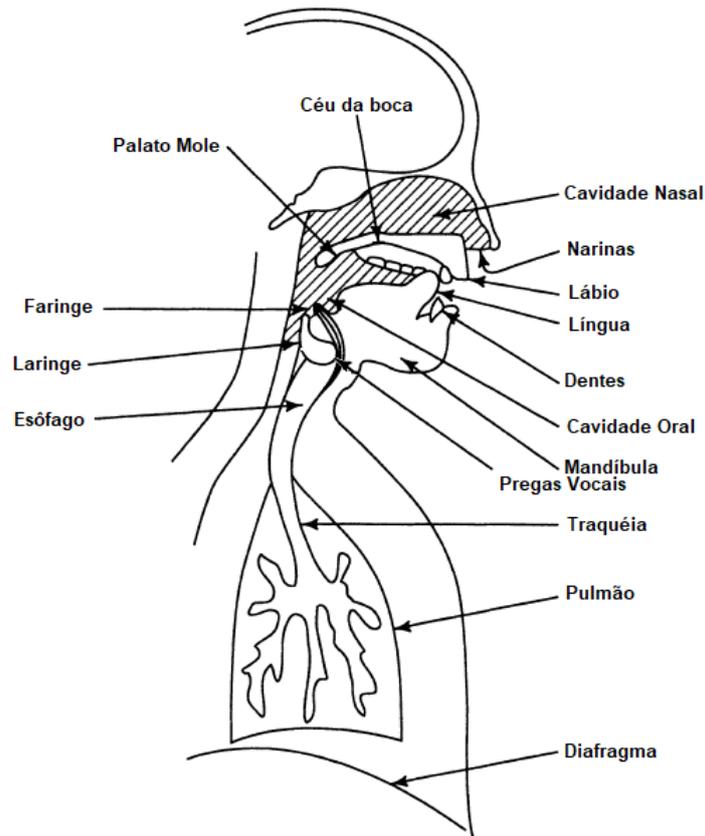


Figura 2.1: Diagrama esquemático do mecanismo de produção de fala humana [39].

As pregas vocais são duas dobras de músculos, ligamentos e mucosas que se estendem horizontalmente na laringe. Na Figura 2.2, são ilustrados os processos de abdução (afastamento) e adução (fechamento) das pregas vocais que ocorrem durante a fonação. Uma desordem nesse movimento, pode acarretar o surgimento de alguns tipos de desordens vocais [42].

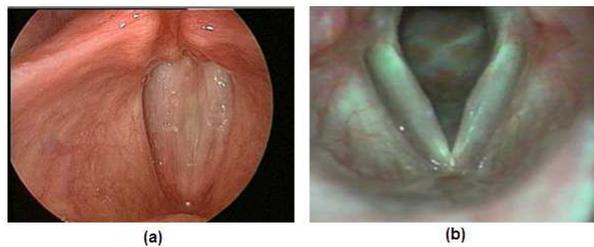


Figura 2.2: Pregas vocais em: (a) adução e (b) abdução - visão endoscópica. Fonte: [43].

No início da vida, a laringe se apresenta muito alta e seguindo ao desenvolvimento orgânico, ela inicia sua descida em relação à posição no pescoço, o que continua por toda a vida, permanecendo na mesma posição entre os 15 e 20 anos e segue descendo discretamente durante a terceira idade. A consequência direta a esse fato é o alongamento do tubo de ressonância que pode amplificar melhor as frequências graves [44]. Na Figura 2.3 é apresentado a imagem de uma laringe obtidas por telelaringoscopia.

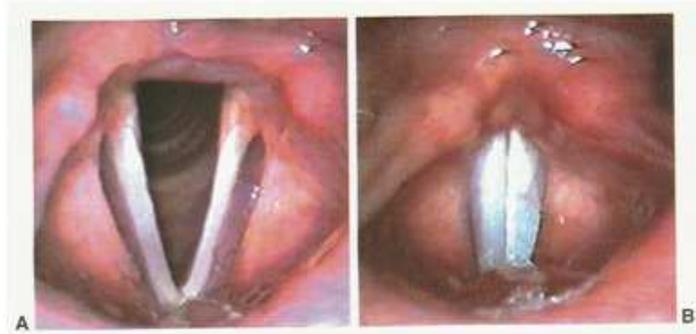


Figura 2.3: Imagens da laringe adulta, obtidas por telarlaringoscopia. A. Durante a respiração. B. Durante a fonação. Fonte: [1].

## 2.2 Análise Acústica dos Sinais de Voz

A análise acústica de sinais de voz tem como objetivo quantificar e caracterizar um sinal sonoro, possibilitando a integração de dados fornecidos pela avaliação perceptivo-auditiva com o plano fisiológico. Tal método permite um detalhamento do processo de geração do sinal sonoro, fornecendo uma estimativa indireta dos padrões vibratórios das pregas vocais, bem como dos formatos do trato vocal e das modificações nestes formatos [42].

Quando usada no âmbito do estudo da voz, a análise acústica permite, de forma não invasiva, comparada aos exames laringoscópicos usuais, determinar e quantificar a qualidade vocal do indivíduo através dos diferentes parâmetros acústicos que compõem o sinal: periodicidade, amplitude, duração e composição espectral. Constituído-se, assim, um método de avaliação objetiva que permite, entre outras utilidades, um diagnóstico precoce de problemas vocais.

Por muito tempo, a abordagem tradicional para modelagem da voz tem sido o modelo linear (fonte-filtro), no qual as pregas vocais são consideradas a fonte sonora e o trato vocal, o filtro. O sinal resultante é o sinal acústico da voz [45].

Várias pesquisas têm sido realizadas utilizando técnicas baseadas no modelo linear de produção da fala, por meio de medidas paramétricas ou não-paramétricas, tais como: codificação por predição linear, análise cepstral e suas derivadas (coeficientes cepstrais, deltacepstrais, cepstrais ponderados e delta-cepstrais ponderados), mel-cepstral (delta mel e delta-delta mel) com aplicações em sistemas de reconhecimento de fala e de locutor, melhoramento do sinal de voz, classificação de vozes patológicas, entre outras [46, 47, 48, 2, 49].

Por meio da análise acústica, os atributos físicos da voz são analisados no domínio do tempo, da frequência e da intensidade, além de outras medidas complexas, que conjugam do cruzamento de tais domínios [3].

Historicamente, o século XX marca o período moderno da análise acústica. As primeiras análises iniciaram-se com o oscilógrafo, em 1920, que produzia gráficos relacionando a amplitude do som e o tempo [1]. Na década de 40, foi desenvolvido o espectrógrafo sonoro, aparelho que teve implicação revolucionária, por permitir um registro tridimensional do sinal sonoro, integrando os aspectos de tempo, frequência e intensidade num único gráfico de dois eixos, chamado de espectrograma [50].

Somente no início dos anos 70, começaram a operar os primeiros processadores digitais de sinais, com definições mais acuradas e mais claras [51], possibilitando o armazenamento

digital [1].

As medidas obtidas na análise acústica correspondem a medidas físicas definidas. O sinal glótico (sinal da fonte) sofre efeitos ao longo do trato vocal supraglótico até a saída deste para o meio externo (ação de filtro). Há uma somatória das ondas sonoras provenientes da fonte glótica com outras refletidas ao longo do trato vocal, sendo a resultante final (sinal de saída), o sinal irradiado pelos lábios [52] [53] como pode ser observada na Figura 2.4.

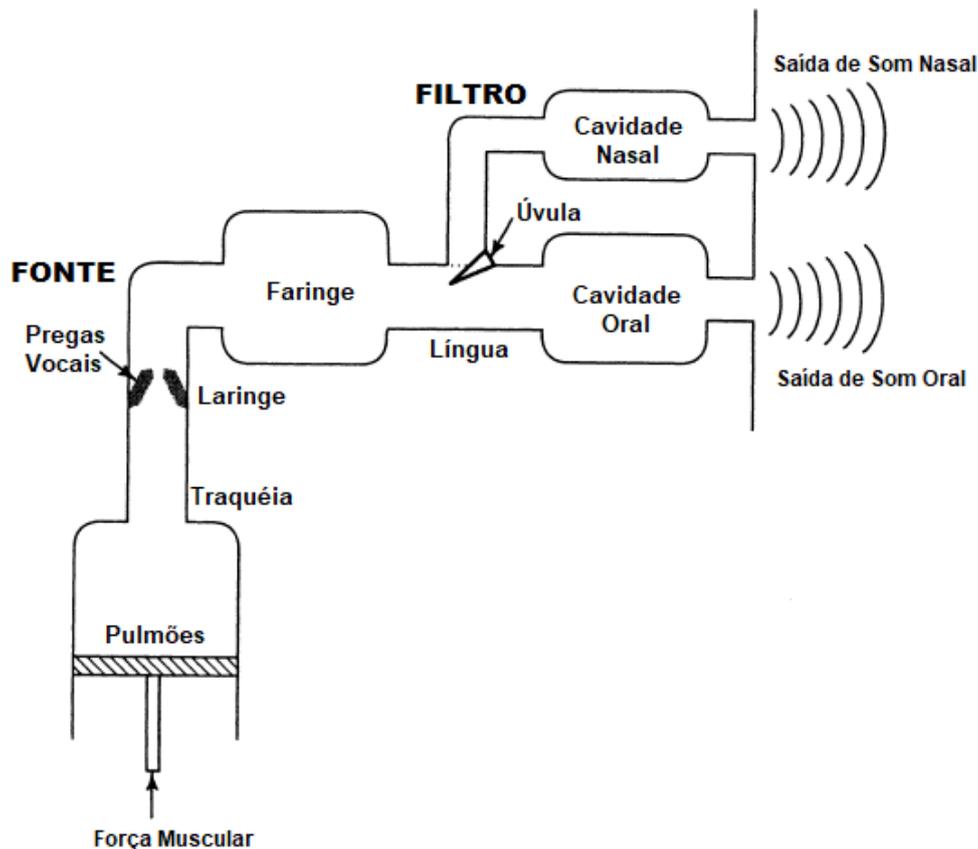


Figura 2.4: Diagrama de blocos da produção da voz humana. [39].

A análise acústica não fornece medidas diretas da fonte glótica, uma vez que o sinal de fala registrado é o sinal de saída, que é a somatória do sinal glótico mais os efeitos dos filtros. Por este motivo, os instrumentais de análise realizam análises indiretas, a partir de procedimentos matemáticos que permitem, por exemplo, eliminar do sinal vocal de saída os efeitos da atividade supraglótica e apresentar medidas relacionadas à atividade glótica.

Contudo, existem fortes evidências teóricas e experimentais [54] [55] [34] [33] para a existência de importantes fenômenos aerodinâmicos não-lineares durante o processo de produção da voz, seja na vibração das pregas vocais ou na geração da onda glótica [56]. Estas não linearidades se devem às fortes forças de restauração na colisão das pregas vocais, turbulência no fluxo de ar devido a uma constrição do trato vocal durante a produção de sons surdos, e ao acoplamento não-linear entre a fonte e o trato vocal durante a pronúncia de alguns sons sonoros [48]. Devido a tais fatores, os métodos clássicos de

análise de dados baseados em um modelo linear têm sido enriquecidos com novos métodos que são derivados da teoria dos sistemas dinâmicos não-lineares [33].

Do ponto de vista dos sistemas dinâmicos, um sinal de voz pode ser considerado como uma observação do sistema de produção da fala, que pode ser usada para revelar e modelar sua dinâmica usando as técnicas não-lineares. Ao longo das últimas duas décadas, pesquisas relacionadas aos mecanismos de produção da voz têm sido realizadas considerando as técnicas da dinâmica não-linear e da teoria do caos com objetivos variados, dentre os quais podem ser destacados: classificação de fonemas [34], reconhecimento automático de locutor [57], discriminação entre vozes saudáveis e patológicas, diagnóstico de patologias laríngeas e avaliação de efeitos de tratamentos clínicos [27, 58, 59, 60, 12, 2].

## 2.3 Patologias Laríngeas

Alterações na voz podem ou não estar associadas a uma patologia orgânica ou física. A produção de uma voz variável e irregular pode ser a primeira manifestação de um distúrbio neurológico. Por outro lado, um problema orgânico sério, como câncer em fase inicial, pode não produzir qualquer alteração perceptível na qualidade da voz. Como exemplos de doenças por fatores neurológicos, que causam distúrbios na voz, podem ser citadas: doença cerebrovascular, paralisia cerebral, paralisia da prega vocal, refluxo, parkinsonismo e esclerose múltipla. Quanto às doenças orgânicas que causam distúrbios na voz estão: laringite crônica, pólipos de prega vocal, nódulos vocais e edema de Reinke [61].

A busca por métodos que proporcionem uma avaliação de distúrbios vocais causados por patologias laríngeas tem sido objeto de estudo de diversos pesquisadores. A comunidade científica procura, cada vez mais, proporcionar ferramentas de apoio ao diagnóstico médico de patologias laríngeas que sejam eficazes no pré-diagnóstico, bem como no acompanhamento de tratamentos medicamentosos e de terapias vocais. A detecção precoce da patologia pode aumentar significativamente a eficiência dos tratamentos médicos [11].

Entretanto, a escolha das características adequadas que representem as distúrbios vocais associadas a patologia, ainda é um campo de estudo em aberto. Muito se tem investigado em discriminar vozes afetadas por patologias laríngeas de vozes saudáveis. Mas, muito pouco, ainda, em se tratando da identificação de parâmetros acústicos que melhor discriminem uma patologia específica dentre outras.

Entre as patologias da laringe que afetam particularmente as pregas vocais são consideradas nesta proposta duas de origem orgânica: edema de Reinke e nódulos, e uma de origem neurológica: paralisia das pregas vocais.

### 2.3.1 Nódulos nas Pregas Vocais

Durante muito tempo pesquisadores tem estudado as semelhanças e diferenças de lesões confinadas ao espaço de Reinke [62, 63, 64]. Existem uma grande quantidade de trabalhos na literatura, que tentam fornecer critérios discretos que permitiriam a diferenciação histopatológica entre pólipos, nódulos e edema de Reinke. No entanto, todos os autores concluem que pólipos, nódulos e edema de Reinke têm características comuns, o que dificulta ao patologista fornecer um diagnóstico definitivo.

Os nódulos são pequenas lesões, aparentemente simétricas, confinadas na junção dos terços anterior e médio das pregas vocais [65] como pode ser visto na Figura 2.5. Essas

lesões geralmente diminuem ou desaparecem quando o uso excessivo da área é interrompido.



Figura 2.5: Nódulos bilaterais nas pregas vocais.  
[65]

Os nódulos vocais são as lesões benignas mais comuns das pregas vocais, e podem acometer tanto crianças quanto adultos. Nos adultos, o público mais afetado são as mulheres, principalmente as profissionais que usam a voz em demasia. O principal motivo que poderia levar à formação de nódulos é o comportamento vocal inadequado por mau uso e abuso vocal [1].

A massa aumentada dos nódulos nas pregas vocais contribui para uma frequência de voz mais grave e com maior período (julgada como rouquidão) [11]. Quanto as características acústicas, na presença dos nódulos o jitter (medida da variação na frequência fundamental de um sinal de voz) e o shimmer (medida da variação na amplitude de um sinal de voz) apresentam valores elevados quando comparados ao valores de uma voz normal (sem patologia laríngea). Outro fato são as evidências de ruído no espectro, de cuja intensidade depende a severidade da rouquidão e o tamanho da lesão [19].

### 2.3.2 Edema de Reinke

Dijkers e Nikkels propuseram definições clínicas para o edema de Reinke [66] com base na descrição de Kleinsasser [67]. O edema de Reinke é um inchaço unilateral ou bilateral das pregas vocais, preenchido com fluido de viscosidade variável, móvel durante a fonação, como pode ser visto na Figura 2.6. Embora um diagnóstico clínico seja viável com o uso da estroboscopia, isso nem sempre é o caso, pois as formas intermediárias com aparência macroscópica mista são frequentemente observadas.

O edema de Reinke tem sido associado a 2 fatores etiológicos: tabagismo e abuso vocal. Recentemente, o refluxo gastroesofágico foi identificado como outro fator de risco potencial. Embora esses fatores tenham sido amplamente aceitos, seu efeito quantificado, particularmente no uso do tabaco e na exposição ao fumo, na gravidade da lesão recebeu pouca atenção na literatura [62].

Partindo para a análise acústica, observa-se uma frequência fundamental mais baixa que a esperada para sexo e idade. Em um estudo realizado com sinais de voz da vogal sustentada /a/ com edema, foi observado que a frequência fundamental para um grupo de 32 mulheres com edema varia em torno de 160 Hz, enquanto para um grupo de 11



Figura 2.6: Edema de Reinke bilateral.  
[65].

homens, varia em torno de 90 Hz. Para as vozes saudáveis, os valores femininos variam em torno de 240 Hz, enquanto os masculinos variam em torno de 140 Hz [2].

Quanto ao shimmer os valores apresentam-se geralmente elevados, o que pode ser caracterizado por uma movimentação mais lenta da vibração das pregas vocais e do ruído espectral, devido à rouquidão característica deste tipo de patologia. A proporção harmônico-ruído pode apresentar valores dentro da normalidade, porém a energia de ruído glótico apresenta-se aumentada. Há uma variação no tempo de fonação, podendo ser mais longo, quando o edema é mais severo ou mais curto em casos de edemas discretos e muito flácidos [1].

### 2.3.3 Paralisia das Pregas Vocais

A paralisia da prega vocal é uma lesão nervosa, que pode atingir apenas uma prega (paralisia unilateral) ou ambas (paralisia bilateral). Essas lesões resultam em uma paresia (fraqueza) ou paralisia dos músculos [68]. Na Figura 2.7 é apresentada a paralisia unilateral da prega direita, de modo que as pregas vocais não fecham completamente durante a fonação.



Figura 2.7: Paralisia unilateral da prega direita durante a fonação.  
[69].

A paralisia das pregas vocais pode ocorrer acidentalmente após cirurgias comuns como cirurgias de coração e pulmão, mesmo que não haja um corte no nervo. O nervo também pode parar de funcionar se esticado ou comprimido, e às vezes depois de um simples manuseio, como por exemplo no ato da colocação do tubo de respiração durante a anestesia geral [70, 15].

A paralisia da prega vocal unilateral é a imobilidade de uma prega vocal por causa da disfunção nervosa. Embora existam outras razões para que uma dobra vocal seja imutável, eles são incomuns. Este nervo pode ser danificado por distúrbios cerebrais (por exemplo, tumores cerebrais e acidentes vasculares cerebrais), traumas cirúrgicos, e por infecções virais dos nervos [68].

A causa principal da paralisia bilateral das pregas vocais em adultos tem sido a tireoidectomia. Entre outras causas, podem ser citadas, ainda, doenças malignas do pescoço ou mediastino, pós-intubação endotraqueal, trauma cervical, doença neurológica ou alguma causa desconhecida. Os sintomas perceptuais mais comuns da paralisia unilateral são a soproidade e a rouquidão [11].

## 2.4 Considerações Finais do Capítulo

Neste capítulo foram apresentados os aspectos inerentes a produção da voz, trazendo uma abordagem acerca dos sinais de vozes saudáveis e patológicas, análise acústica dos sinais de voz, bem como as patologias laríngeas consideradas nesta proposta.

No capítulo seguinte, será introduzida a fundamentação teórica que serviu de base para o desenvolvimento dessa tese. Será apresentado um paralelo entre sistemas dinâmicos e sua aplicação na análise de sinais de voz, bem como as medidas de classificação utilizadas.

# Capítulo 3

## Fundamentação Teórica

Nas últimas décadas tem-se verificado um grande desenvolvimento no estudo dos fenômenos não-lineares com a introdução de novas abordagens e conceitos no tratamento de sistemas dinâmicos. O estudo de propriedades topológicas de atratores caóticos reconstruídos a partir de séries temporais experimentais tem permitido, uma interpretação alternativa dos processos turbulentos, onde as estruturas dissipativas coerentes desempenham um papel relevante [71].

A análise dinâmica de sinais pode ser feita por meio de uma modelagem matemática, onde se faz necessária a construção de um modelo que possa ser usado para se obter informações úteis associadas à produção do fenômeno em estudo, ou ainda, por meio da análise de séries temporais, em que se considera simplesmente uma série temporal escalar, em geral, associada com uma aquisição experimental para entender o comportamento dinâmico do sistema. O ponto essencial dessa análise é que uma série temporal contém informações sobre variáveis não observáveis do sistema, o que permite a reconstrução do espaço de estados [11, 72].

De acordo com Kantz e Schreiber [73], a ligação mais direta entre a teoria do caos e o mundo real é a análise dos dados de séries temporais, nos termos da dinâmica não linear. É devido, principalmente, ao desenvolvimento de uma base matemática coerente para a descrição do caos, em dinâmicas determinísticas não-lineares, que se tem dado uma significativa atenção à análise dinâmica não-linear e ao modelamento determinístico dos dados de uma série temporal [74, 75].

Segundo Savi [76] uma série temporal pode ser definida como um conjunto discreto numerável de valores de uma variável de estado de um sistema dinâmico. O não conhecimento das equações que regem o sistema de produção dessas séries são algumas das dificuldades encontradas no estudo da dinâmica associada a séries temporais experimentais. A inevitável presença de ruído em sinais experimentais dificulta ainda o seu estudo ainda, podendo acarretar interpretações incorretas dos resultados. Por exemplo, caos determinístico, processos estocásticos e processos multiperiódicos com períodos incomensuráveis dão origem a séries temporais irregulares e aperiódicas que parecem, à primeira vista, caóticas [71]. Dessa forma, o uso de técnicas adequadas é extremamente importante para a análise de séries temporais.

Quando se deseja determinar se a dinâmica que deu origem a uma série temporal é determinística (ou não), há de se lançar mão de outros métodos. Se for determinística, pode-se ter condições de desenvolver um modelo capaz de descrever a dinâmica. No caso de comportamento estocástico, ao contrário, trata-se de um processo a muitos graus de liberdade, caso em que só se pode esperar uma descrição estatística do sistema [71].

Enquanto a definição das invariantes dinâmicas refere-se ao sistema dinâmico em questão, suas estimativas a partir da série temporal tornam-se possíveis por meio de um conjunto de teoremas formulados por Takens [77]. Efetivamente, esses teoremas estabelecem que as invariantes dinâmicas estimadas das séries temporais observadas a partir da reconstrução de sua trajetória no espaço de estados usando coordenadas defasadas serão as mesmas do sistema dinâmico em análise sob certas condições.

### 3.1 Sistemas Dinâmicos e Processamento de Voz

Os métodos tradicionalmente empregados no processamento da voz para avaliação de patologias laríngeas assumem o modelo de produção linear da voz, com medidas derivadas a partir da estimação da forma de onda do fluxo glotal (hansen1998). Diversos estudos tem mostrado que a produção da voz é um processo fortemente não-linear envolvendo efeitos biomecânicos e aerodinâmicos [54, 78].

A teoria dos sistemas dinâmicos não-lineares, através do estudo do caos determinístico, oferece uma alternativa para a análise do sinal de voz. Caos é um termo que descreve um comportamento pseudo-aleatório gerado por um sistema que é ao mesmo tempo determinístico e não-linear [33].

A saída de um sistema caótico é não-predizível e extremamente sensível a pequenas diferenças nas condições iniciais. Contudo, métodos baseados na dinâmica não-linear são capazes de descrever de forma quantitativa o comportamento caótico.

Medidas advindas da análise dinâmica não-linear tem sido empregadas no processamento de séries temporais associadas a sinais de voz em aplicações que vão da caracterização deste tipo de sinal [57, 79] a aplicações na análise da voz apresentado algum tipo de patologia na laringe [12, 78, 80].

Observações clínicas demonstram que a maioria das doenças na laringe causa mudanças na qualidade da voz, o que resulta em irregularidades e instabilidades vocais. Métodos de dinâmica não-linear são capazes de analisar comportamentos irregulares e podem ser valiosos em várias áreas do estudo da voz, incluindo avaliação de tratamentos clínicos, classificação de vozes de acordo com a taxonomia da desordem proposto por Titze [81], diferenciação de vozes saudáveis e patológicas, e possibilidade de diagnóstico de patologias da laringe [33].

### 3.2 Sistemas Dinâmicos

Um sistema dinâmico pode ser definido como um modelo matemático, comumente empregado para representar sistemas físicos, biológicos, financeiros, entre outros [82] [76]. A dinâmica dos sistemas é caracterizada por estados (grandezas) que mudam com o tempo [83].

A quantidade de variáveis de estado corresponde à dimensão do sistema dinâmico, ou seja, ao número de graus de liberdade do sistema [82] [76]. A partir das variáveis de estado, pode-se investigar o comportamento da dinâmica dos sistemas, por meio do chamado espaço de fase.

Um espaço de fase ou espaço de estados é um espaço abstrato no qual representamos a evolução de um sistema dinâmico. Através do espaço de fase, podemos investigar o comportamento dinâmico do sistema a partir das variáveis de estado [84].

Um sistema dinâmico pode ser descrito por um sistema de equações diferenciais[85].

$$\frac{d\mathbf{v}(t)}{dt} = \mathbf{F}(\mathbf{v}(t)) \quad (3.1)$$

com  $d \geq 3$  graus de liberdade,  $\mathbf{v}(t) = [x_1(t), x_2(t), \dots, x_d(t)]$ .

### 3.2.1 O Exemplo do Pêndulo Simples

Poincaré [86], em seu trabalho premiado sobre a estabilidade do sistema solar, usou o pêndulo simples para exemplificar alguns conceitos e ideias. O pêndulo, é de fato, um sistema paradigmático no estudo de comportamentos dinâmicos [83].

Seja o sistema constituído por um corpo de massa  $m$  e uma haste rígida de comprimento  $l$  (de massa desprezível) que pode se mover livremente em um plano vertical, como mostrado na Figura 3.1

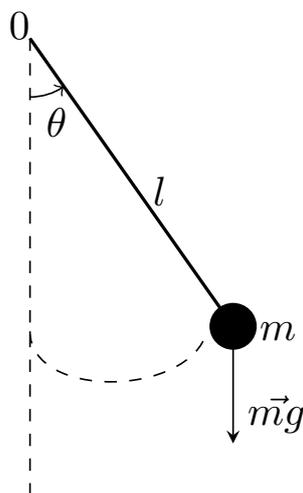


Figura 3.1: O Pêndulo Simples.

A equação de movimento associada é:

$$ml^2\ddot{\theta} + mgl \sin \theta = 0 \quad (3.2)$$

Não é possível se obter uma solução para essa equação em termos de funções elementares. Entretanto, é possível identificar as principais características de suas soluções e compreender de modo qualitativo os possíveis movimentos deste sistema físico utilizando um diagrama de fases.

A equação 3.2 pode ser escrita como

$$\begin{aligned} \dot{\theta} &= \varphi = f(\theta, \varphi), \\ \dot{\varphi} &= -(g/l) \sin \theta = g(\theta, \varphi). \end{aligned} \quad (3.3)$$

Para construir as trajetórias no espaço de fases  $(\theta, \dot{\theta} = \varphi)$  faz-se

$$\frac{d\theta}{d\varphi} = \frac{f}{g} = \frac{\varphi}{-(\frac{g}{l} \sin \theta)}, \quad (3.4)$$

ou seja,

$$-\left(\frac{g}{l}\right) \sin \theta d\theta = \varphi d\varphi, \quad (3.5)$$

Integrando-a obtém-se

$$\varphi^2 - a \cos \theta = C, \quad (3.6)$$

em que  $a = 2g/l$  e  $C = cte$ .

No chamado espaço de fases tem-se uma curva  $h(\theta, \varphi)$  diferente para cada valor da constante  $C$ . As curvas representadas na Figura 3.2 são chamadas trajetórias de fase ou linhas de fluxo. O conjunto de todas as curvas é chamado diagrama de fases ou retrato de fases do sistema. Um dado par de valores  $(\theta, \dot{\theta})$  é definido como estado do sistema. O diagrama de fases apresentado na Figura 3.2 mostra como esse estado evolui à medida que o tempo passa.

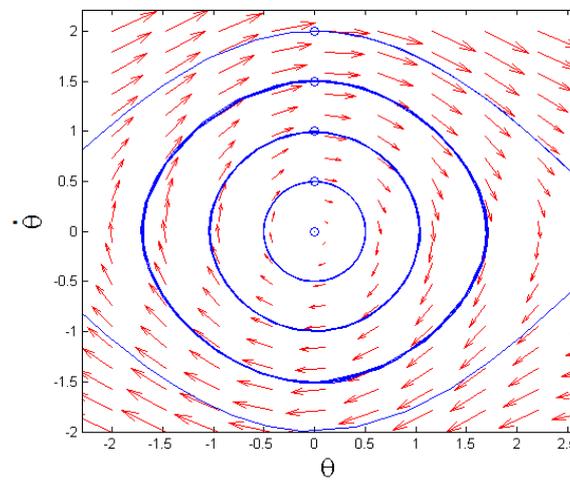


Figura 3.2: Diagrama de fases para o pêndulo simples.

Embora a tradição dos estudos em sistemas dinâmicos remonte a Henri-Poincaré [86], que inspirado por problemas em mecânica celeste, percebeu a utilidade do estudo das estruturas topológicas no espaço de fases de trajetórias dinâmicas, é no domínio dos sistemas dissipativos, fluxos e mapas para os quais o volume associado na espaço de fases contraem-se, que a ruptura epistemológica ocorre, com a introdução dos conceitos de atrator estranho (caótico) e do seu corolário, à sensibilidade as condições iniciais.

Um atrator estranho, fornece uma descrição global do comportamento assintótico de um sistema dinâmico. Diversas informações podem ser extraídas a partir do estudo dos atratores. A caracterização das propriedades dos atratores, com a utilização das teorias do caos determinístico e sistemas dinâmicos, permite a análise de séries experimentais com comportamento caótico e consiste uma das motivações centrais dessa teoria [71].

A constatação pioneira de que movimentos intrinsecamente caóticos podem ocorrer em sistemas determinísticos dissipativos é devida a Lorenz [87], que, com a finalidade de estudar problemas da previsão meteorológica para tempos longos, analisou as equações associadas a processos físicos envolvendo a convecção térmica bidimensional, concluindo pela impraticabilidade de tal previsão devido às imprevisões na determinação das condições iniciais [71].

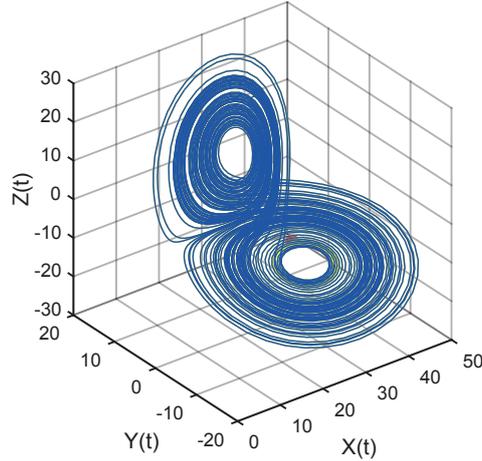


Figura 3.3: Atrator de Lorenz, com parâmetros  $\sigma = 10$ ,  $r = 28$  e  $b = 8/3$ .

O sistema de Lorenz possui três graus de liberdade, correspondendo a três variáveis de estado  $x(t)$ ,  $y(t)$  e  $z(t)$  [87]:

$$\frac{dx(t)}{dt} = -\sigma x(t) + \sigma y(t), \quad (3.7)$$

$$\frac{dy(t)}{dt} = -x(t)z(t) + rx(t) - y(t), \quad (3.8)$$

$$\frac{dz(t)}{dt} = x(t)y(t) - bz(t), \quad (3.9)$$

em que  $\sigma$ ,  $r$  e  $b$  são parâmetros de controle [87]. Cada estado desse sistema é representado por um ponto na trajetória, que é parametrizado pela variável escalar  $t$ , tempo. Uma projeção do atrator de Lorenz é apresentada na Fig. 3.3.

Os sistemas dinâmicos também podem ser descritos por mapas de tempo discretos[85].

Podemos concentrar nossa análise em mapas de tempo discretos, uma vez que toda a amostragem ocorre em tempo discreto. Nesse caso, nós amostramos fluxos contínuos no tempo de início  $t_0$  em amostras de tempo  $\tau_s$ , para obter:

$$\mathbf{v}(t_0 + (n + 1)\tau_s) \approx \mathbf{v}(t_0 + n\tau_s) + \tau_s \mathbf{F}(\mathbf{v}(t_0 + n\tau_s)). \quad (3.10)$$

### 3.2.2 Sistemas Dinâmicos não-lineares

Enquanto o comportamento futuro do sistema não-linear pode ser determinado se as condições iniciais forem conhecidas, o mesmo não ocorre com um sistema aleatório. Embora um sistema não-linear evolua no tempo com um comportamento instável e aperiódico, tal comportamento é determinístico, pois seu estado futuro pode ser conhecido, desde que se tenha conhecimento do seu estado atual.

O estado futuro pode, porém, ser radicalmente modificado a partir de pequenas mudanças no estado atual. A dificuldade de se conhecer o estado presente com exatidão leva à necessidade de modelar o sistema não-linear como aleatório, em algumas situações, quando os detalhes do comportamento não são de interesse, embora ele seja, na realidade, determinístico.

A teoria dos sistemas dinâmicos tem fornecido novas ferramentas para se analisar séries temporais caóticas obtidas em experimentos. Em geral, nem todas as variáveis de um sistema podem ser observadas, dispondo-se apenas da medida de um único observável  $x(t)$ . Os sinais de voz, por exemplo, podem apresentar a evolução no tempo de apenas uma variável de estado,  $x(t)$  (um sinal escalar). Dessa forma, a dinâmica pode ser apresentada como [85]:

$$x(t_0 + n\tau_s) \rightarrow x(t_0 + (n + 1)\tau_s), \quad (3.11)$$

ou simplesmente

$$x(n) \rightarrow x(n + 1). \quad (3.12)$$

Os sistemas dinâmicos subjacentes, como o processo de produção da voz, é muito complexo e suas equações não são conhecidas. No entanto, o teorema de imersão de Takens [77] estabelece que, a partir do conhecimento de apenas uma variável do sistema, no caso da voz a série temporal, é possível reconstruir um espaço de fase difeomorficamente equivalente ao original. O teorema de imersão de Takens permite converter uma sequência de medidas escalares  $x(n)$  em vetores de estado, recuperando a dinâmica do sistema [73] [77]. O espaço de fase reconstruído com  $m$  dimensões é então formada por vetores  $\mathbf{x}(n)$ :

$$\mathbf{x}(n) = [x(n), x(n - \tau), \dots, x(n - (m - 1)\tau)] \quad (3.13)$$

em que  $x(n)$  é o sinal de voz,  $\tau$  é o atraso de reconstrução, ou seja, um número inteiro múltiplo do tempo de amostragem  $\tau_s$ , e  $m$  é a dimensão de imersão do espaço de fase reconstruído. O teorema de Takens [77] é estritamente um teorema de existência e não sugere como determinar o atraso de reconstrução  $\tau$  e a dimensão de imersão  $m$ . A chamada dimensão de imersão  $m$  pode ser obtida por vários métodos, por exemplo, o método dos vizinhos falsos. Nesta proposta, não entraremos em detalhes sobre o parâmetro  $m$ . Em vez disso, nos concentraremos em nosso objeto de estudo, o atraso de reconstrução  $\tau$ .

Além de ser um múltiplo do tempo de amostragem, o atraso de reconstrução  $\tau$  deve obedecer a algumas características que o tornam adequado para recuperar a dinâmica do sistema. De fato, um atraso de reconstrução muito curto não capturará a dinâmica dos dados, enquanto um atraso de reconstrução muito grande o tornará completamente independente em um sentido estatístico [85]. Fraser e Swinney [88] propuseram uma das técnicas conhecidas e populares para determinar o atraso de reconstrução adequado, técnica essa advinda da teoria da informação, chamada de informação mútua média.

### 3.2.3 O método da Informação Mútua Média

O método da informação mútua média foi proposto por A. M. Fraser e H. L. Swinney [88]. De acordo com os autores, ao se escolher o passo  $\tau$   $\tau'$ , em que  $\tau'$  é o intervalo de tempo que minimiza a informação mútua e é determinado pelo ponto em que ocorre o primeiro mínimo na curva de informação mútua, que representa a autocorrelação da série, medindo a dependência do sinal consigo mesmo, pode-se garantir a reconstrução de vetores com o menor nível de informação redundante (estatisticamente independentes), mas ainda correlacionados.

A teoria da informação procura identificar o quanto de informação se pode ter, de uma medida realizada em um determinado instante de tempo  $t$ , quando se observa outra medida, do mesmo sinal, em um tempo posterior  $t + \tau$ . A informação mútua média é uma alternativa à função de autocorrelação, uma vez que é capaz de medir eficientemente a informação tanto para sistemas lineares quanto para sistemas não-lineares [76]. Na Figura 3.4 é apresentada a curva de informação mútua para obtenção do  $\tau$  ótimo e o PMIM em um sinal normal (sem patologia) e o espaço de fase reconstruído a partir do  $\tau$  ótimo e dimensão de imersão  $m = 2$  e na Figura 3.5 são ilustrados exemplos de espaços de fase reconstruídos com o  $\tau$  ótimo para cada tipo de voz que será investigada nesta tese.

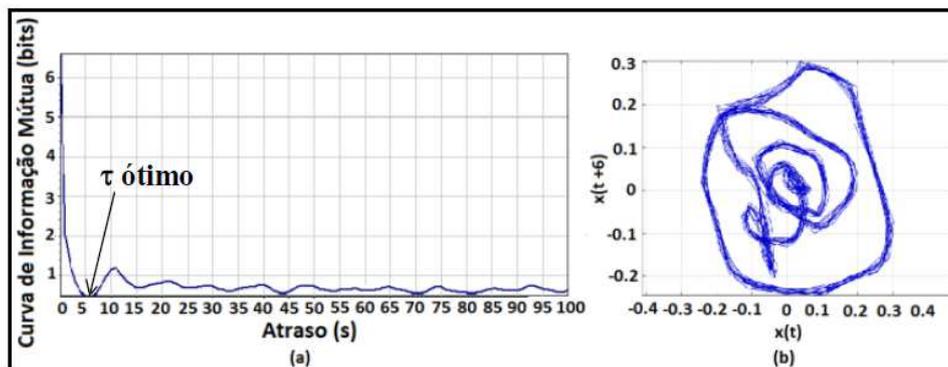


Figura 3.4: (a) Curva de Informação Mútua para obtenção do  $\tau$  ótimo e PMIM para um sinal normal e (b) Espaço de fase reconstruído para o sinal da curva (a), com  $\tau$  ótimo = 6 e dimensão de imersão  $m = 2$ . Fonte: [27]

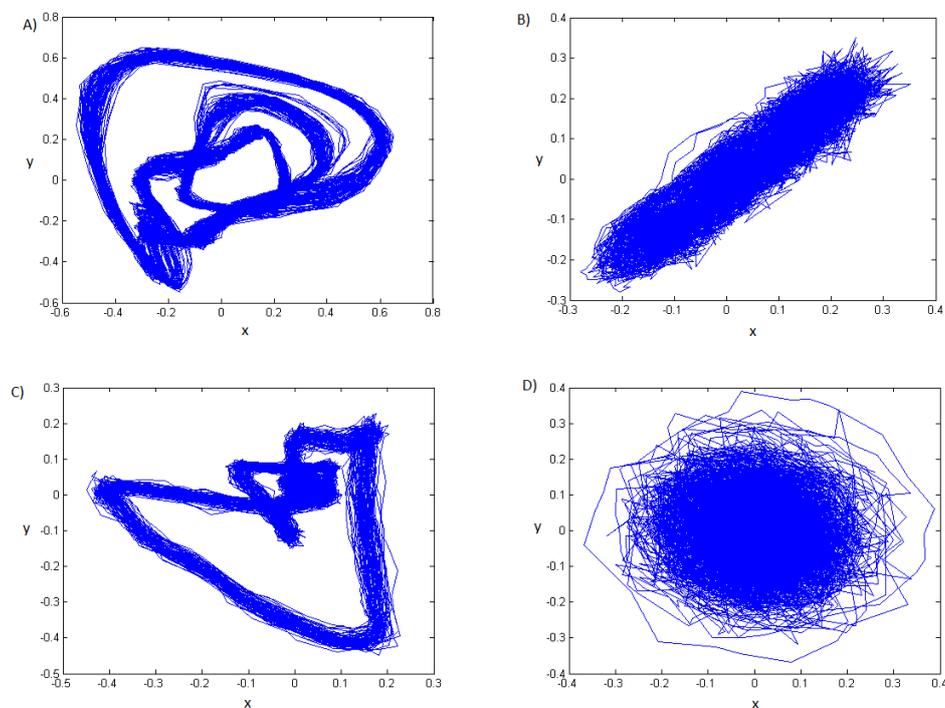


Figura 3.5: Exemplos de espaço de fase reconstruído (projeção em duas dimensões) para sinais de vozes: (A) Saudável, (B) Com Edema de Reinke, (C) Com Nódulo e (D) com Paralisia.

### 3.3 Medidas da Análise Dinâmica Não-linear

#### 3.3.1 Primeiro Mínimo da Função de Informação Mútua (PMIM) e atraso de Reconstrução $\tau$

Como mencionado anteriormente, para determinarmos o atraso de reconstrução  $\tau$  adequado para cada sistema utiliza-se comumente ferramentas, técnicas da teoria da informação, como o método de informação mútua média proposto por [88]. De acordo com este método, pode-se garantir vetores de reconstrução com o menor nível de informação redundante, mas ainda correlacionados. A teoria da informação tem como objetivo identificar o quanto de informação se pode ter de uma medida realizada em um determinado instante de tempo  $t$ , quando se observa outra medida, do mesmo sinal, em um tempo posterior  $t + \tau$  [73].

#### Informação Mútua

De maneira geral, a informação mútua indica o quanto duas ou mais variáveis aleatórias são dependentes. Ela fornece as mesmas informações que a função de autocorrelação fornece em sistemas lineares sendo, na realidade, um tipo de generalização para sistemas não-lineares [89]. Mais precisamente, a informação mútua é uma medida confiável da dependência entre variáveis aleatórias medindo o quanto o conhecimento de uma variável aleatória diminui a incerteza sobre outra, como por exemplo entre as variáveis aleatórias discretas  $X$  e  $Y$ , sendo dada pela equação 3.14 [90]:

$$I(X; Y) = - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (3.14)$$

$$= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x) + \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x|y) \quad (3.15)$$

Na equação 3.16,  $H(X)$  é a entropia de Shannon da variável aleatória  $X$ , ao passo que  $H(X|Y)$  é a entropia, da variável aleatória  $X$  quando a variável aleatória  $Y$  é conhecida.

$$I(X; Y) = H(X) - H(X|Y) \quad (3.16)$$

As medidas acima citadas são definidas da seguinte forma:

$$H(X) = - \sum_{x \in \mathcal{X}} p(x) \log p(x) \quad (3.17)$$

$$H(X|Y) = - \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x, y) \log p(x|y) \quad (3.18)$$

Pela regra da cadeia, a informação mútua pode ser estendida a mais de duas variáveis aleatórias, como dada pela equação 3.19:

$$I(X^N; Y^N) = H(Y^N) - H(Y^N|X^N) \quad (3.19)$$

$$= - \sum_{n=1}^N I(Y_n; X^N|Y^{n-1}) \quad (3.20)$$

em que  $H(Y^N)$  é a entropia conjunta, das variáveis aleatórias  $Y_1, \dots, Y_N$ , e  $H(Y^N|X^N)$  é a entropia condicional de  $Y^N$  em relação a  $X^N$ .

A informação mútua entre duas variáveis será baixa quando elas forem fracamente correlacionadas e, alta, na situação contrária. Quando a informação mútua é 0, significa que as duas variáveis são independentes, sem nenhuma correlação entre si. A seguir, tem-se outras relações entre a informação mútua e a entropia [90]:

$$I(Y; X) = H(Y) - H(Y|X) \quad (3.21)$$

$$I(X; Y) = H(X) + H(Y) - H(X, Y) \quad (3.22)$$

$$I(X; Y) = I(Y; X) \quad (3.23)$$

$$I(X; X) = H(X) \quad (3.24)$$

Para o caso de variáveis aleatórias contínuas, a informação mútua é definida por:

$$I(X; Y) = - \int f(x, y) \log \frac{f(x, y)}{f(x)f(y)} dx dy \quad (3.25)$$

$$= h(X) - h(X|Y) \quad (3.26)$$

em que  $h(X)$  é a entropia diferencial da variável aleatória  $X$  e  $h(X|Y)$  é a entropia diferencial condicional.

Para se calcular a entropia ou informação mútua entre variáveis aleatórias discretas, faz-se necessário conhecer inicialmente sua distribuição de probabilidade. Entretanto, muitas vezes essa distribuição e/ou seus parâmetros não são conhecidos, fazendo-se necessário sua estimação.

A estimação da entropia e da informação mútua, envolvendo apenas variáveis aleatórias discretas, é simples, com aplicação direta da definição de Shannon. No entanto, quando uma das variáveis envolvidas é contínua, torna-se necessária a aplicação de um método de estimação de densidade.

### Estimação dos Parâmetro PMIM e $\tau$ através do método ingênuo

A informação mútua média entre  $x(t)$  e sua versão defasada  $x(t + \tau)$  é obtida a partir de um histograma de  $b$  intervalos de classe (bins), criado para estimar a distribuição de probabilidade dos dados do sinal  $x(t)$  [28] [11]:

$$I(\tau) = - \sum_{i=1}^b \sum_{j=1}^b p_{i,j}(x(t), x(t + \tau)) \log \left[ \frac{p_{i,j}(x(t), x(t + \tau))}{p_i(x(t))p_j(x(t + \tau))} \right], \quad (3.27)$$

em que  $p_i(x(t))$  é a probabilidade de que o sinal  $x(t)$  assuma um valor dentro do  $i$ -ésimo intervalo do histograma,  $p_j(x(t + \tau))$  é a probabilidade de que  $x(t + \tau)$  esteja no  $j$ -ésimo intervalo e  $p_{i,j}(x(t), x(t + \tau))$  é a probabilidade de que, simultaneamente,  $x(t)$  esteja no  $i$ -ésimo intervalo e  $x(t + \tau)$  no  $j$ -ésimo intervalo [73]. O tempo de defasagem, então, é o valor de  $\tau$  para o qual a função de informação mútua média atinge seu primeiro mínimo local (PMIM) [88].

Apesar do método de estimação baseado em histogramas, mais conhecido na literatura como estimador ingênuo, ser bastante difundido e amplamente utilizado, ele apresenta

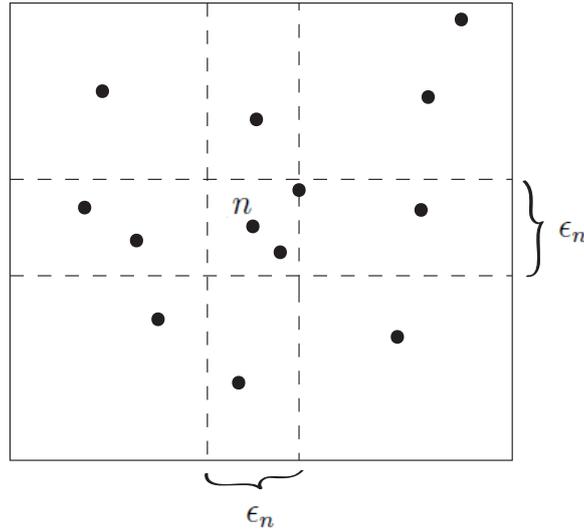


Figura 3.6: Determinação de  $\epsilon_n$ ,  $n_x$ , e  $n_y$  para o algoritmo KSG, para  $k = 2$  e um  $n$  fixado. Nesse exemplo,  $n_x = 3$ ,  $n_y = 5$ , e  $N = 13$ .

uma forte dependência da dimensionalidade dos dados e erros sistemáticos, produzindo um viés de estimação, que em algumas situações, difere do valor de informação mútua obtido analiticamente [91] [92] [93]. Ou seja, o valor de informação mútua estimado pode não representar um valor verdadeiro de informação mútua, nem mesmo um valor próximo do valor da informação mútua verdadeiro. Isso acontece porque essa estimativa depende fortemente do número  $b$  de bins (segmentos) utilizados.

### Estimação dos Parâmetro PMIM e $\tau$ através do Método KSG

Em contraste com o método ingênuo acima citado, Kraskov et. al. [38] desenvolveram um estimador de informação mútua, que é aqui chamado estimador KSG. O estimador do KSG baseia-se no trabalho de Kozachenko e Leonenko [94], que estima a entropia ( $H(X)$ ) baseado na distribuição média de probabilidade dos vizinhos de cada ponto da amostra. A informação mútua pode ser escrita como na Equação 3.28:

$$I(X; Y) = H(X) + H(Y) - H(X, Y) \quad (3.28)$$

A ideia básica do estimador KSG é usar diferentes vizinhos para estimar as entropias marginais  $H(X)$  e  $H(Y)$  e estimar a entropia conjunta  $H(X, Y)$  para cancelar o viés do estimador. Existem dois estimadores ligeiramente diferentes desenvolvidos por Kraskov et. Al [38] com desempenhos semelhantes, um deles é adotado aqui. Este estimador considera a distância de cada ponto para o seu  $k$ -ésimo vizinho mais próximo, projeta a distância em relação a  $X$  e  $Y$  e considera o espaçamento mais largo dos dois:  $\epsilon/2 = \max\{\epsilon_x/2, \epsilon_y/2\}$ . Com essas distâncias, é possível contar o número de pontos  $n_x$  e  $n_y$  em relação a  $X$  e  $Y$  que estão a uma distância estritamente inferior ao espaçamento  $\epsilon/2$ , Como ilustrado na Fig. 3.6.

A estimativa KSG é dada como na Equação 3.29:

$$I(X, Y) = \psi(k) - \langle \psi(n_x + 1) + \psi(n_y + 1) \rangle + \psi(N) \quad (3.29)$$

em que  $\langle \rangle$  denota a média aritmética,  $k$  é o número de vizinhos considerado,  $N$  é o tamanho da amostra e  $\psi$  é a função digamma.

A função digamma é definida como a primeira derivada do logaritmo da função gamma ( $\Gamma(x)$ ) [95]:

$$\psi(x) = \frac{d}{dx} \log(\Gamma(x)) \quad (3.30)$$

$$= \frac{\Gamma'(x)}{\Gamma(x)} \quad (3.31)$$

A função  $\Gamma(x)$  é dada por:

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt, \quad (3.32)$$

para  $x > 0$ .

Se as distribuições forem muito distorcidas e/ou irregulares, os autores do método sugerem transformá-las para se tornarem mais uniformes (ou pelo menos mais ou menos simétricas). Nesse caso, o estimador KSG apresentou melhora significativa nos resultados depois de transformar as variáveis em  $x' = \log x$  e  $y' = \log y$ .

### 3.3.2 Entropia de Shannon

A entropia é uma das medidas mais comuns e bem conhecidas da teoria da informação, ciência originalmente proposta em [96] por Claude Shannon. Em geral, a entropia refere-se ao grau de desordem, incerteza ou quantidade de informação. Em se tratando de sistemas dinâmicos a entropia,  $H$ , é a medida de informação necessária para localizar um sistema em um determinado estado, isso significa que  $H$  é a medida da incerteza sobre o sistema físico [36].

As medidas de entropia avaliam o grau de desordem de um sinal, podendo ser empregadas para avaliar as desordens nos sinais de vozes afetados por patologias na laringe [97][32].

Embora o conceito de entropia seja relativamente simples e matematicamente bem objetivo, na prática, sua estimação pode ser um processo complexo, pois nem sempre se conhece o conjunto de probabilidades necessários para sua obtenção, fazendo-se necessário o uso de estimadores. Os estimadores comumente utilizados para variáveis contínuas são baseados em histograma e em funções de suavização de kernel, no entanto, quando aplicados para estimação de outras medidas podem produzir viés [38] [92]. Tais problemas são minimizados utilizando estimadores baseados nas distâncias de vizinhos tanto para estimação de entropia quanto de outras mediadas [98, 38].

### Entropia Diferencial

De acordo com Shannon, a entropia é uma medida do número de bits necessários, em média, para descrever variáveis aleatórias discretas e, a entropia diferencial está relacionada ao logaritmo do volume efetivo no qual variáveis aleatórias contínuas estão inseridas. Ao contrário da entropia discreta, a entropia diferencial pode ser negativa e ocorre quando esse volume é menor que um.

A entropia diferencial está relacionada ao grau de dispersão de uma variável aleatória. Ou seja, para pequenos valores de entropia temos a variável aleatória confinada em um

pequeno volume e, à medida que o valor da entropia aumenta, indica que o volume no qual a variável aleatória está inserida também está aumentando, amplificando assim, sua dispersão.

Um teorema importante da entropia para variáveis aleatórias contínuas está na propriedade de equipartição assintótica (AEP). As definições e teoremas que formalizam a AEP permitem definir um conjunto típico e caracterizar o comportamento de sequências típicas.

De modo geral, é possível dividir o conjunto de todas as sequências em dois, o conjunto típico (a entropia da amostra está próxima da entropia verdadeira) e o conjunto não-típico (demais sequências). Qualquer propriedade comprovada para as sequências típicas é verdadeira, com alta probabilidade, e determina o comportamento médio de uma amostra grande [90].

**Teorema 1.** *Seja  $X_1, X_2, \dots, X_n$  uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) com função densidade de probabilidade  $f(x)$  [90]. Então,*

$$-\frac{1}{n} \log f(X_1, X_2, \dots, X_n) \rightarrow E[-\log f(X)] = h(X), \quad (3.33)$$

converge em probabilidade isto é, para todo  $\epsilon > 0$ ,

$$Pr \left\{ \left| -\frac{1}{n} \log f(X_1, X_2, \dots, X_n) - h(X) \right| > \epsilon \right\} \rightarrow 0. \quad (3.34)$$

A prova do Teorema 1 é uma consequência direta da lei fraca dos grandes números, na qual, para variáveis aleatórias independentes e identicamente distribuídas (i.i.d.),  $\frac{1}{n} \sum_{i=1}^n X_i$  se aproxima do valor esperado  $E[X]$  ao considerar grandes valores de  $n$ .

**Definição 1.** *Para  $\epsilon > 0$  e um  $n$  qualquer, um conjunto típico  $A_\epsilon^{(n)}$  em relação à função densidade de probabilidade  $f(x_1, x_2, \dots, x_n)$  é definido como:*

$$A_\epsilon^{(n)} = \left\{ (x_1, x_2, \dots, x_n) \in \llbracket X \rrbracket : \left| -\frac{1}{n} \log f(x_1, x_2, \dots, x_n) - h(X) \right| \leq \epsilon \right\} \quad (3.35)$$

em que  $(x_1, x_2, \dots, x_n)$  são realizações independentes e identicamente distribuídas da variável aleatória  $X$  e  $f(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i)$ .

**Teorema 2.** *O conjunto típico  $A_\epsilon^{(n)}$  possui as seguintes propriedades:*

1.  $Pr \left( A_\epsilon^{(n)} \right) > 1 - \epsilon$  para  $n$  suficientemente grande.
2.  $Vol \left( A_\epsilon^{(n)} \right) \leq 2^{n(h(X)+\epsilon)}$  para todo  $n$ .
3.  $Vol \left( A_\epsilon^{(n)} \right) \geq (1 - \epsilon) 2^{n(h(X)-\epsilon)}$  para  $n$  suficientemente grande.

## Estimadores de Entropia Diferencial

Como citado anteriormente, nem sempre temos o conhecimento do conjunto de probabilidades necessário para o cálculo da entropia, fazendo-se necessário sua estimação. O processo de estimação de entropia para variáveis contínuas requer atenção, e soluções aparentemente simples podem comprometer esse processo, por exemplo, a discretização

dos dados pode resultar em estimativas enviesadas. Neste trabalho, dois estimadores para entropia diferencial serão apresentados, um baseado em histogramas, comumente abordado na literatura como ingênuo, e o outro proposto por Kozachenko-Leonenko [94] que neste trabalho chamaremos de KOLE, baseado nos  $k$ -vizinhos mais próximos.

## Histograma

A abordagem mais simples e comum para estimar entropia de uma variável aleatória contínua é a técnica baseada em histograma, que representa a distribuição de frequência (absoluta ou relativa) por meio de retângulos, cujas larguras representam intervalos e cujas áreas são proporcionais às frequências.

Na técnica de histograma por frequência relativa, o suporte no qual  $N$  realizações de uma variável aleatória estão inseridas, é quantizado em intervalos de comprimento  $\Delta$  e, portanto, assume-se que a função densidade de probabilidade é constante em cada intervalo. O número de amostras presentes no mesmo intervalo de quantização é contado e a estimação da função densidade de probabilidade da variável aleatória, para qualquer valor real de  $x$ , é dada por:

$$\hat{f}(x) = \frac{\text{número de observações no mesmo intervalo}}{N\Delta}. \quad (3.36)$$

Após determinação da função densidade de probabilidade, estima-se a entropia diferencial utilizando a relação entre variáveis aleatórias discretas e contínuas [90],

$$\hat{H}(X) + \log \Delta \rightarrow h(X), \quad (3.37)$$

em que  $\hat{H}(X)$  é a entropia da variável quantizada,

$$\hat{H}(X) = - \sum \Delta \hat{f}(x) \log \hat{f}(x) - \log \Delta, \quad (3.38)$$

sendo uma estimação verdadeira quando a função densidade de probabilidade é integrável segundo Riemann [99].

## Estimador Kozachenko-Leonenko - KOLE

Em contrapartida ao método baseado em histogramas, o estimador proposto por Kozachenko-Leonenko [94] é um método direto, isto é, não estima necessariamente a função densidade de probabilidade das amostras pra estimar a entropia.

O estimador KOLE considera que a distribuição de probabilidade da distância entre  $x_i$  e seu  $k$ -ésimo vizinho é igual a chance de ter:

1. 1 amostra com distância  $r \in [\epsilon/2, \epsilon/2 + d\epsilon/2]$  até  $x_i$ ;
2.  $k - 1$  outras amostras com distâncias menores que  $\epsilon/2$ , e;
3.  $N - k - 1$  amostras com distâncias maiores que  $\epsilon/2$ .

Portanto, seja  $p_i$  a massa da bola centrada em  $x_i$ ,  $p_i(\epsilon) = \int_{\|\xi - x_i\| < \epsilon/2} f(\xi) d\xi$ , modela-se a seguinte distribuição trinomial:

$$P_k(\epsilon)d\epsilon = \frac{(N-1)!}{1!(k-1)!(N-k-1)!} \times \frac{dp_i(\epsilon)}{d\epsilon} d\epsilon \times p_i(\epsilon)^{k-1} \times (1-p_i(\epsilon))^{N-k-1}. \quad (3.39)$$

O estimador do logaritmo da função densidade de probabilidade,  $\widehat{\log f(x)}$ , é obtido assumindo que  $f(x)$  é constante no entorno da amostra  $x_i$ , assim:

$$p_i(\epsilon) \approx c_d \epsilon^d f(x_i), \quad (3.40)$$

em que  $d$  é a dimensão da variável aleatória  $X$  e  $c_d$  é o volume da bola unitária  $d$ -dimensional (para norma Euclideana  $c_d = \pi^{d/2}/\Gamma(1 + d/2)$ ).

Calculando o valor esperado do logaritmo em ambos os lados da equação 3.40, têm-se:

$$E \log f(x_i) \approx \log p_i(\epsilon) - dE \log \epsilon - \log c_d, \quad (3.41)$$

mas,

$$\begin{aligned} E(\log p_i(\epsilon)) &= \int_0^\infty P_k(\epsilon) \log p_i(\epsilon) d\epsilon \\ &= k \binom{N-1}{k} \int_0^1 p_i(\epsilon)^{k-1} (1-p_i(\epsilon))^{N-k-1} \log p_i(\epsilon) dp_i(\epsilon) \\ &= \psi(k) - \psi(N), \end{aligned} \quad (3.42)$$

em que  $p_i$  pode ser interpretada como uma variável aleatória beta-distribuída na qual, a função densidade de probabilidade é:

$$f(p_i; k, N-k) = \frac{p_i^{k-1} (1-p_i)^{N-k-1}}{\mathbf{B}(k, N-k)}. \quad (3.43)$$

Desta forma, obtém-se o seguinte estimador para entropia diferencial:

$$\widehat{h}(X) = -\psi(k) + \psi(N) + \log c_d + \frac{d}{N} \sum_{i=1}^N \log \epsilon(i) \quad (3.44)$$

em que  $\psi(\cdot)$  é a função *digamma*<sup>1</sup>,  $N$  é o número total de amostras analisadas,  $c_d$  é o volume da bola unitária  $d$ -dimensional (para norma Euclideana  $c_d = \pi^{d/2}/\Gamma(1 + d/2)$ ) e  $\epsilon(i)$  é duas vezes a distância de  $x_i$  ao seu  $k$ -ésimo vizinho.

**Exemplo 3.1.** Considerando o vetor de realizações de uma variável aleatória desconhecida  $\mathbf{x} = [-6 \quad -4,2 \quad -1 \quad 0,5 \quad 1 \quad 3]$ , a representação gráfica para interpretação do estimador Kozachenko-Leonenko é expressa na Figura 3.7. Nesse caso,  $N = 6$ ,  $d = 1$  e  $c_d = 2$ . Deve-se fixar o  $k$ -ésimo vizinho analisado assim,

- se  $k = 2$ , então para  $x_i = -1$  têm-se  $\epsilon(i) = 4$  e para  $x_i = 0,5$  têm-se  $\epsilon(i) = 3$ ;
- se  $k = 3$ , então para  $x_i = -1$  têm-se  $\epsilon(i) = 6,4$  e para  $x_i = 0,5$  têm-se  $\epsilon(i) = 5$ .

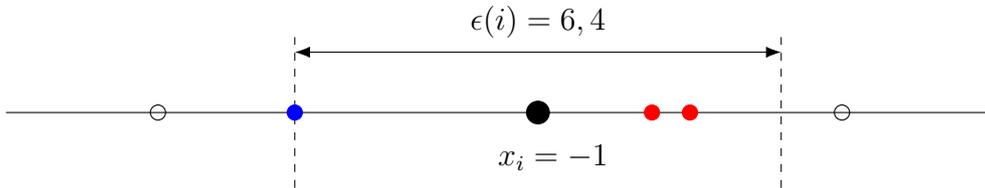


Figura 3.7: Representação gráfica da interpretação do estimador Kozachenko-Leonenko considerando  $k = 3$  então têm-se pelo menos uma amostra com distância  $r \in [\epsilon/2, \epsilon/2 + d\epsilon/2]$  até  $x_i$  (em azul),  $k - 1$  amostras com distâncias menores que  $r$  (em vermelho) e,  $N - k - 1$  amostras com distâncias maiores que  $r$  (branco).

<sup>1</sup>A função *digamma* é definida como a derivada do logaritmo da função *gamma*:  $\psi(x) = \frac{d}{dx} \log(\Gamma(x))$ .

### 3.3.3 Entropia de Correlação

O atrator reconstruído possui uma estrutura geométrica, dotada de uma medida relacionada às frequências relativas com as quais diferentes partes do atrator são visitadas. A medida associada ao atrator de um sistema dinâmico é invariante sob o operador de evolução, e, conseqüentemente, é chamada de medida invariante.

As dimensões e entropias se tornaram populares para caracterizar atratores e são invariantes no âmbito do processo de reconstrução. Elas permitem caracterizar as medidas de probabilidade, associadas à dimensão de imersão do atrator reconstruído,  $m$ , de uma série temporal estacionária.

A entropia de correlação ( $K_2$ ), em particular, é comumente utilizada na caracterização de sistemas dinâmicos, uma vez que pode ser estimada diretamente a partir da integral de correlação  $C_m(\varepsilon)$ . A integral de correlação pode ser definida como a probabilidade de que dois pontos independentes do atrator, estejam a uma certa distância um do outro e essa distância seja menor que  $\varepsilon$ , isto é:

$$C_m(\varepsilon) = \iint \Theta(\varepsilon - \|x - y\|) d\mu_m(y) d\mu_m(x) \quad (3.45)$$

em que a função  $\Theta\{\cdot\}$

$$\Theta\{s\} = \begin{cases} 0, & \text{se } s < 0 \\ 1, & \text{se } s \geq 0, \end{cases} \quad (3.46)$$

e  $\|\cdot\|$  é uma norma. Usualmente são consideradas a norma do supremo  $\|x_1, \dots, x_m\| = \sup_i |x_i| = \max(|x_1|, \dots, |x_m|)$  ou a norma euclidiana  $(\sum_{i=1}^m x_i^2)^{1/2}$ .

Para séries temporais determinísticas, a integral de correlação, dado um valor pequeno de  $\varepsilon$  e uma dimensão de imersão  $m$  grande, comporta-se de acordo com a relação

$$C_m(\varepsilon) \sim e^{-m\tau K_2 r^{D_2}} \quad (3.47)$$

em que  $\tau$  é o atraso de reconstrução,  $D_2$  é a dimensão de correlação e  $K_2$  é a entropia de correlação.

A entropia de correlação  $K_2$  mede o aumento exponencial da incerteza sobre os valores futuros, dado o passado com precisão finita. Mais precisamente, é um limite inferior na soma de expoentes positivos de Lyapunov e é, portanto, uma medida de “quão caótico” é o sistema [100]. Sistemas lineares, por exemplo, devem ter  $K_2 = 0$  enquanto a dimensão de correlação pode ser arbitrariamente grande.

Grassberger & Procaccia(1983)[101] desenvolveram um método para estimar a integral de correlação dada na equação 3.45 a partir da soma de correlação

$$\hat{C}_m(\varepsilon) = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \Theta(\varepsilon - \|X_i - X_j\|) \quad (3.48)$$

A partir desse método, a entropia de correlação  $K_2$  pode ser obtida pela equação

$$K_2(\varepsilon, m) = \frac{1}{\tau} \ln \frac{C_m(\varepsilon)}{C_{m+1}(\varepsilon)} \quad (3.49)$$

Experimentos realizados por [102] mostram que, o pequeno aumento da estimativa da entropia de correlação após a incorporação da dimensão de imersão acima de  $m = 10$ ,

pode ser atribuído à presença de uma pequena quantidade de ruído de medição. Dessa forma, nos experimentos realizados no desenvolvimento dessa tese, os valores da dimensão de imersão  $m$  variam de  $m = 1, \dots, 10$ . Similarmente ao que foi apresentado na subseção 3.3.1, a entropia de correlação foi estimada para valores de  $\tau$  obtidos tanto com o método ingênuo quanto com o estimador KSG.

## 3.4 Considerações Finais do Capítulo

Neste capítulo foram apresentados os conceitos base para o desenvolvimento desta tese. Foram introduzidos os conceitos de sistemas dinâmicos, espaço de fases, como se dá a análise dinâmica de sinais, além de um paralelo entre sistemas dinâmicos e sinais de voz (aplicação considerada nesta tese). Foram apresentadas ainda algumas medidas da análise dinâmica não-linear, sua aplicação em análise de sinais de vozes, além do detalhamento das medidas consideradas nesta tese para a classificação das patologias laríngeas já discutidas no capítulo anterior.

No próximo capítulo serão apresentados os trabalhos que serviram de base para o desenvolvimento desta tese e como eles se relacionam com o presente trabalho, com um breve comparativo dos resultados apresentados na literatura e os obtidos nesta tese.

# Capítulo 4

## Materiais e Métodos

### 4.1 Base de Dados

A base de dados utilizada foi a base de dados Disordered Voice Database, Model 4337, da Kay Elemetrics, gravada pelo Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab [103]. Dessa base de dados, foram utilizados 53 sinais de voz de locutores com laringes saudáveis e 114 sinais de voz de locutores afetados por patologias laríngeas (52 sinais de voz de laringes afetadas por paralisia nas pregas vocais, 44 sinais de voz de laringes afetadas por edema de Reinke, e 18 sinais de voz de laringes afetadas por nódulos vocais). Os sinais são da vogal sustentada /a/. Os sinais de voz de laringes saudáveis, originalmente amostrados a 50.000 amostras/s, foram sub-amostrados a 25.000 amostras/s para equiparar à taxa de amostragem dos sinais de voz de laringes patológicas.

Cinco classes de sinais foram consideradas neste estudo: saudável (SDL), paralisia (PRL), edema (EDM), nódulos (NDL) e todas as patologias juntas (PTL). São considerados sete casos de discriminação:  $SDL \times PTL$ ,  $SDL \times PRL$ ,  $SDL \times EDM$ ,  $SDL \times NDL$ ,  $PRL \times EDM$ ,  $PRL \times NDL$ , e  $EDM \times NDL$ .

### 4.2 Metodologia

A metodologia empregada segue o modelo de diagrama em blocos representado na Figura 4.1.

Os sinais de voz selecionados da base de dados são analisados por meio da estimação das medidas  $\tau$ , PMIM, Entropia de Shannon e Entropia de Correlação, utilizando-se o estimador clássico e o estimador KSG. Essas características foram extraídas diretamente da série temporal que representa o sinal de voz.

Como ferramentas para extração das medidas são empregados os seguintes softwares: TISEAN - Time Series Analysis v. 3.01 [104] e o CRP (Cross Recurrence Plot) [105], além de elaboração de rotinas (scripts) e uso de funções pré-definidas no ambiente Matlab®v.8.0.0.783 .

Em seguida, é realizada a classificação dos sinais como saudáveis ou patológicos - e entre sinais de vozes apresentando diferentes patologias. Como o objetivo principal desta tese é mostrar o quanto a escolha de um modelo estatístico mais robusto pode influenciar diretamente nos resultados da classificação, e não somente apresentar medidas que apresentem um alto valor de acurácia na separação entre as classes (uma vez que grande parte das medidas já foi amplamente utilizada em outros trabalhos) utilizamos dois

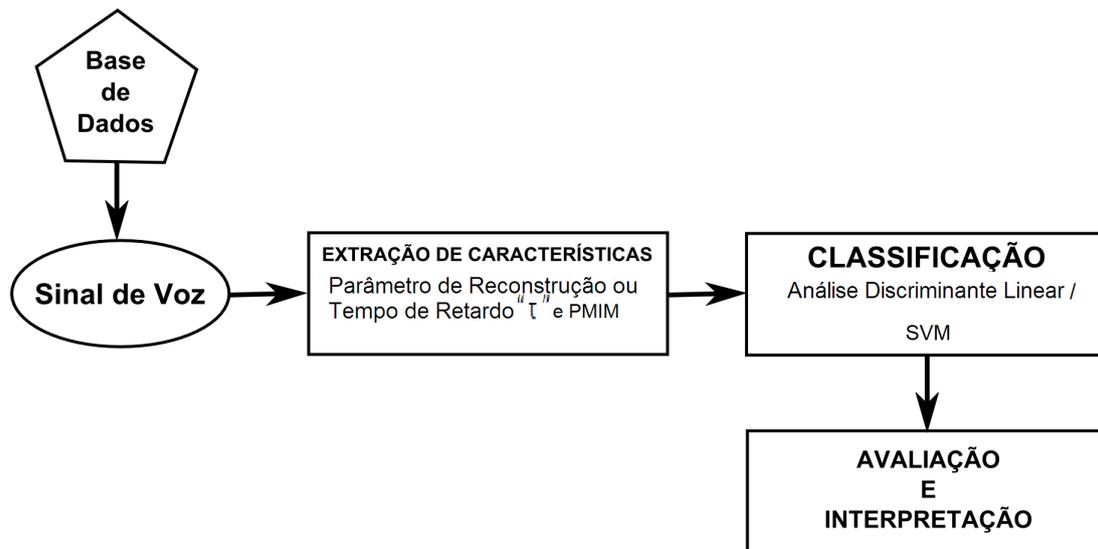


Figura 4.1: Diagrama em blocos da metodologia empregada.

classificadores distintos: um baseado em análise de discriminante e o outro em máquinas de vetor de suporte (SVM).

Nesta proposta empregou-se a análise discriminante linear (LDA), por apresentar melhor desempenho para as classes consideradas, para detectar a presença de desordens vocais causadas pelas patologias edema de Reinke, paralisia e nódulos nas pregas vocais e os resultados obtidos foram comparados (resultados obtidos com estimador clássico x resultados obtidos com estimador KSG).

Inicialmente, é feita a classificação com cada uma das medidas consideradas individualmente e depois, com o objetivo de melhorar os resultados individuais, as medidas são combinadas entre elas.

## 4.3 Descrição dos Classificadores

### 4.3.1 Análise Discriminante

A análise discriminante é uma técnica da estatística multivariada utilizada para classificar objetos em dois ou mais grupos. A ideia básica é a obtenção de uma combinação linear das características observadas que apresente maior poder de discriminação entre populações. Esta combinação linear é denominada função discriminante [20] [2] [106].

Entretanto, em geral, a média e a variância das populações não são conhecidas, resultando na necessidade de estimação desses parâmetros. Dessa forma, pode-se assumir que as populações têm uma mesma matriz de covariâncias ou não. Quando a regra de classificação assume que as matrizes de covariância são iguais, as funções discriminantes são ditas lineares (LDA) e, quando não, são funções discriminantes quadráticas (QDA) [20] [2] [106].

Na Figura 4.2, é apresentado um exemplo da aplicação da análise discriminante linear em um espaço de duas características. A função discriminante linear realiza a transformação dos dados de ambas as classes para o sub-espaço LDA, no qual é traçado um hiperplano de separação entre as mesmas [20] [2] [1].

Na Figura 4.3, é apresentado um exemplo da aplicação da análise discriminante

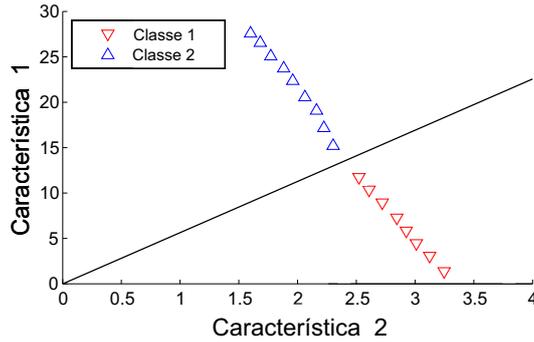


Figura 4.2: Função discriminante linear em um espaço de características arbitrário. Fonte: [28].

quadrática em um espaço de duas características. A função discriminante quadrática busca uma curva não linear que proporcione a maior separabilidade entre as classes.

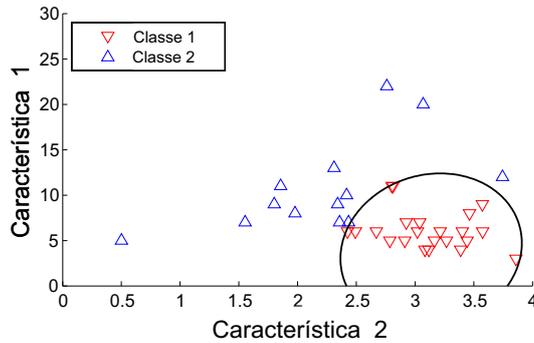


Figura 4.3: Função discriminante quadrática em um espaço de característica arbitrário. Fonte: [28].

Visando utilizar a maior quantidade de dados possível em ambas as fases do processo de classificação, é criada uma partição aleatória para definir os conjuntos de treinamento e teste para a validação do modelo estatístico usando a validação cruzada com  $k$  subconjuntos ( $k$ -fold cross-validation) estratificada.

O termo “estratificada” significa que todos os grupos envolvidos no processo de classificação estarão representados, de forma proporcional, nos conjuntos de teste e de treino. Nesta pesquisa é utilizado um valor de  $k$  igual a 10 no processo de validação cruzada. Segundo Witten et al. [107], extensivos testes sobre vários conjuntos de dados usando diferentes técnicas de aprendizado automático, têm mostrado que  $k = 10$  é o número de subconjuntos (“folds”) que tem apresentado as melhores estimativas de erro, além de existir algumas evidências teóricas, embora não conclusivas, que sugerem esse valor. Por isso, a validação cruzada com 10 subconjuntos tem se tornado padrão em termos práticos [108].

Inicialmente, é analisado o desempenho de classificação considerando apenas dois grupos de sinais: um de vozes patológicas - formado conjuntamente por vozes com edema, paralisia e nódulos - e o outro de vozes saudáveis. Posteriormente, é analisado o desempenho de classificação entre cada dois dos quatro diferentes grupos:  $SDL \times PRL$ ,  $SDL$

× EDM, SDL × NDL, PRL × EDM, PRL × NDL, e EDM × NDL.

### 4.3.2 Máquinas de Vetor de Suporte - SVM

O SVM (do inglês *Support Vector Machines*) constitui um método de aprendizado de máquina supervisionado proposto por Vapnik [109], amplamente utilizado em classificação de dados [110]. O objetivo da classificação por meio de vetor de suporte é encontrar o hiperplano de separação ótimo, em um espaço de características de alta dimensão, de uma forma computacionalmente eficiente.

O SVM é aplicado tanto em problemas de ordem linear quanto problemas de ordem não-linear e usa uma função denominada kernel para mapear os dados de entrada em um espaço de característica de alta dimensão no qual o problema se torna linearmente separável para o classificador [111]. A boa capacidade de generalização, robustez em espaços de grande dimensão, convexidade da função objetivo e uma teoria de aprendizado bem definida matemática e estatisticamente são importantes vantagens do classificador SVM [112].

Para realizar a classificação, esta técnica transforma vetores de entrada em um espaço de características de alta dimensão, usando uma transformação não linear (com uma função kernel), e então realiza uma separação linear neste espaço de características. A partir de um conjunto de treinamento  $Q_\xi$ ,  $\{Q_\xi\}_{\xi=1}^N = \{(\mathbf{x}_\xi, L_\xi)\}_{\xi=1}^N$ ,  $L_\xi \in \{+1, -1\}$  representa o estado efetivo  $L$  do sinal acústico representado por  $\xi$ . Assim, o classificador é um hiperplano definido como:

$$g(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b, \quad (4.1)$$

em que  $\mathbf{w}$  é o vetor normal que é perpendicular ao hiperplano, e  $b$  é o deslocamento do hiperplano a partir da origem. Dessa forma,

$$\begin{aligned} \mathbf{w}^T \mathbf{x} + b &\geq 0, & L_\xi &= +1 \\ \mathbf{w}^T \mathbf{x} + b &< 0, & L_\xi &= -1 \end{aligned} \quad (4.2)$$

Então, o hiperplano é escolhido por meio da solução do seguinte problema de otimização:

$$\text{minimize } \frac{1}{2} \mathbf{w}^T \mathbf{w}, \quad (4.3)$$

que está sujeito a

$$L_\xi(\mathbf{w}^T \mathbf{x} + b) \geq 1, \quad \xi = 1, 2, \dots, N. \quad (4.4)$$

Neste trabalho foi utilizada a biblioteca LIBSVM<sup>1</sup> em ambiente Linux, utilizando os parâmetros padrão, Kernel função de base radial ( $\exp(-\text{gamma} * |u - v|^2)$ ), com os comandos executados diretamente no terminal. A proporção de dados utilizados para treino e teste se deu da seguinte forma: 30% dos dados foram utilizados para treino e 70% para teste.

<sup>1</sup><https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>

## 4.4 Avaliação e Interpretação

Para mensurar a precisão dos classificadores empregados em cada estudo de caso, a medida de acurácia, comumente empregada, foi utilizada. No caso da análise discriminante, outras duas medidas foram utilizadas: sensibilidade e especificidade. Essas medidas estão relacionadas à capacidade de um classificador em diagnosticar uma doença em um paciente doente (Verdadeiro Positivo - VP) ou saudável (Falso Positivo - FP), ou, ainda, diagnosticar um estado saudável em um paciente saudável (Verdadeiro Negativo - VN) ou doente (Falso Negativo - FN) [11]. Cada um desses parâmetros aparecem na chamada matriz de confusão (Tabela 4.1), que representa o resultado obtido no classificador.

Tabela 4.1: Matriz de confusão em um teste de detecção da presença/ausência de doença.

Resultado	Doença	
	Presente	Ausente
Positivo	Verdadeiro Positivo (VP)	Falso Positivo (FP)
Negativo	Falso Negativo (FN)	Verdadeiro Negativo (VN)

A acurácia (Ac) é a taxa de classificação correta global, refletindo a capacidade do classificador de identificar corretamente quando há e quando não há a presença da patologia. A acurácia é definida como a relação entre o número de casos corretamente classificados e todos os casos apresentados ao classificador [11]:

$$Ac = \frac{VP + VN}{VP + VN + FP + FN}. \quad (4.5)$$

A sensibilidade (Sen) é a capacidade do classificador em identificar a presença da patologia quando ele de fato existe, sendo definida pela relação entre o número de casos corretamente classificados como presença da patologia e a quantidade total de casos com a patologia:

$$Sen = \frac{VP}{VP + FN}. \quad (4.6)$$

A especificidade (Esp) é a capacidade do classificador em identificar corretamente a ausência da patologia quando de fato ele não existe, sendo definida pela relação entre o número de casos corretamente classificados como saudáveis e a quantidade total de casos de estado saudável:

$$Esp = \frac{VN}{VN + FP}. \quad (4.7)$$

O classificador apresenta bom desempenho caso seja capaz de obter altos valores para acurácia, sensibilidade e especificidade. Quando se trata da discriminação entre uma classe saudável e uma classe patológica, a representação das medidas de sensibilidade e especificidade é mais clara. No entanto, quando a discriminação é entre classes patológicas, é necessário que seja definido, no classificador, qual grupo de sinais terá sua correta classificação medida pela sensibilidade e qual grupo terá sua correta classificação medida pela especificidade [28].

## 4.5 Considerações Finais do Capítulo

Neste capítulo foi apresentada a metodologia empregada no desenvolvimento desta tese. Foram apresentadas a base de dados utilizada, com o detalhamento da quantidade de sinais para cada tipo de patologia considerada e os casos de classificação avaliados. Foram apresentados também os classificadores utilizados e como são obtidas as medidas de acurácia, sensibilidade e especificidade.

No capítulo seguinte serão apresentados comparativos entre os estimadores ingênuo e KSG/KOLE. Serão utilizados sinais gaussianos a fim de observar como convergem os valores de entropia de Shannon e informação mútua dado o emprego de cada estimador.

# Capítulo 5

## Resultados

Neste capítulo estão reunidos todos os resultados obtidos da classificação entre vozes saudáveis e patológicas, bem como entre sinais de vozes patológicas utilizando os estimadores ingênuo e KSG/KOLE/KOLE e os classificadores LDA e SVM. Ao final de cada seção os resultados são discutidos e os melhores resultados obtidos são comparados graficamente.

São apresentados ainda um teste de desempenho entre os estimadores utilizando sinais gaussianos com o objetivo de apresentar como os estimadores convergem para o valor analiticamente calculado para as medidas de entropia e informação mútua

### 5.1 Teste dos Estimadores

Nesta seção iremos testar o comportamento dos estimadores apresentados no capítulo 3 na estimativa das medidas entropia de Shannon e informação mútua, para isso utilizaremos sinais gaussianos com média zero e variância unitária. O objetivo principal é comparar o desempenho dos estimadores KSG/KOLE/KOLE e ingênuo na estimação das medidas de informação quanto a sua convergência em relação aos valores calculados analiticamente. A escolha por distribuições gaussianas se deu por serem distribuições conhecidas e de fácil cálculo analítico das medidas de entropia e informação mútua.

São apresentadas ainda a distribuição dos valores médios das medidas utilizadas nesta tese para os 53 sinais de vozes saudáveis (SDL) e 114 sinais de vozes patológicas (PTL) utilizadas nos casos de classificação. No gráfico de caixas (*boxplot*) apresentados nas Figuras 5.1, 5.2, 5.5 e 5.6 a linha interior em cada caixa (retângulo) indica o valor da mediana, as bases do retângulo mostram o 25-ésimo e o 75-ésimo percentil, os traços horizontais exteriores às caixas (*whiskers*) delimitam as extremidades do intervalo dos dados quando não há pontos fora de uma determinada faixa de valores e as pequenas cruces acima destas são valores atípicos (*outliers*).

#### 5.1.1 Medida $\tau$

Inicialmente estimamos o valor de  $\tau$  utilizando o estimador ingênuo e o estimador KSG/KOLE, a fim de comparar, qual método era capaz de estimar os menores valores de  $\tau$ , uma vez que se deseja estimar o menor valor de  $\tau$  que permita que os vetores sejam estatisticamente independentes. Nas Figuras 5.1 e 5.2, estão ilustradas a distribuição dos valores médios de  $\tau$  para sinais de vozes saudáveis (SDL) e patológicas com paralisia (PRL), edema de

Reinke (EDM) e nódulos (NDL), utilizando-se o estimador ingênuo e o estimador KSG respectivamente.

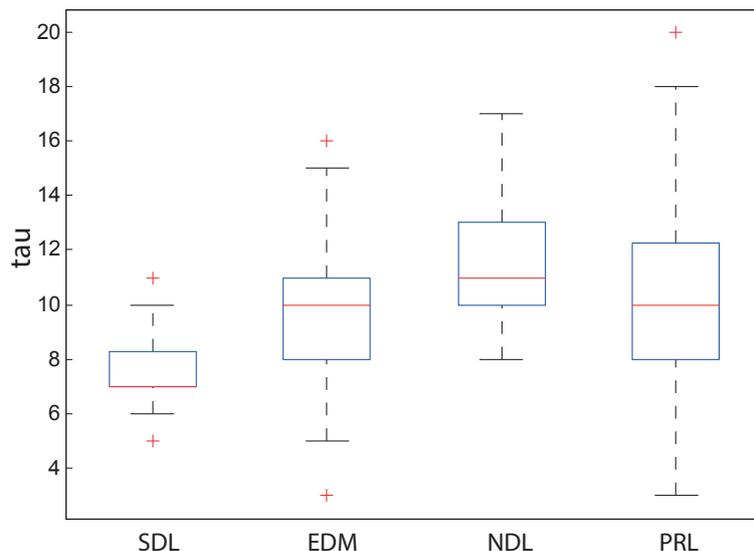


Figura 5.1: Distribuição média dos valores de  $\tau$  para os sinais de vozes saudáveis (SDL) e patológicos com as patologias edema de Reinke (EDM), nódulo (NDL) e paralisia das pregas vocais (PRL), estimadas com o estimador ingênuo.

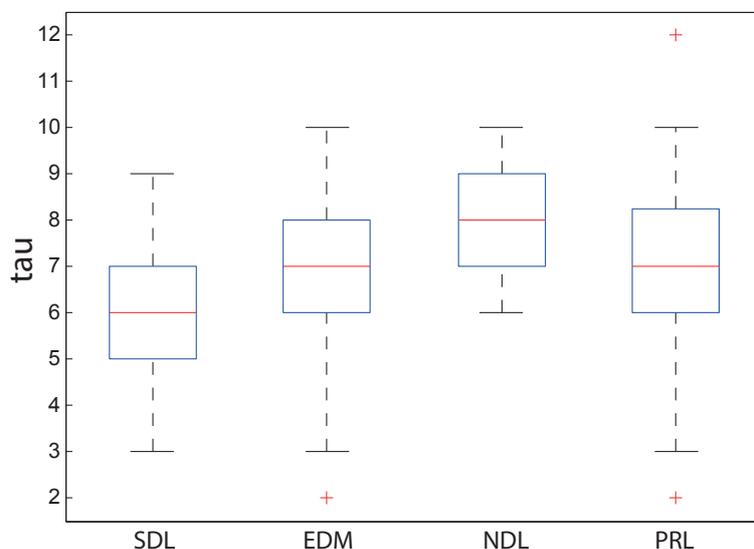


Figura 5.2: Distribuição média dos valores de  $\tau$  para os sinais de vozes saudáveis (SDL) e patológicos com as patologias edema de Reinke (EDM), nódulo (NDL) e paralisia das pregas vocais (PRL), estimadas com o estimador KSG.

Os resultados apresentados nas Figuras 5.1 e 5.2 revelaram diferenças significativas nos valores de  $\tau$  estimados com o estimador ingênuo e com o estimador KSG para todas classes consideradas. Em alguns casos os valores de  $\tau$  caem mais que a metade, como por exemplo: para a classe edema de Reinke, com o estimador ingênuo estimamos um valor de  $\tau = 16$  e com o estimador KSG, para o mesmo sinal de voz, estimamos um valor de  $\tau = 8$ , para a classe nódulo, com o estimador ingênuo estimamos um valor de  $\tau = 15$

e com o estimador KSG, para o mesmo sinal de voz, estimamos um valor de  $\tau = 7$ , para a classe paralisia, com o estimador ingênuo estimamos um valor de  $\tau = 20$  e com o estimador KSG, para o mesmo sinal de voz, estimamos um valor de  $\tau = 9$ .

### Teste dos Estimadores Ingênuo e KSG

O estimador KSG tem se mostrado menos tendencioso do que o estimador ingênuo, em termos de enviesamento, apresentando valores de informação mútua muito próximos ao valor analítico [91] [113] [92]. Para testar o desempenho dos estimadores foram feitas simulações com conjuntos de dados gaussianos bivariados, com média zero, variância unitária e valores determinados para coeficientes de correlação -  $r = 0$ ,  $r = 0.3$ ,  $r = 0.6$  e  $r = 0.9$ . Para esses casos, existe um valor analítico bem estabelecido para a informação mútua, que pode ser usado para comparar as estimativas:

$$I(X; Y) = -\frac{1}{2} \log(1 - r^2). \quad (5.1)$$

As Figuras 5.3 e 5.4 ilustram a simulação com o estimador ingênuo e o estimador KSG.

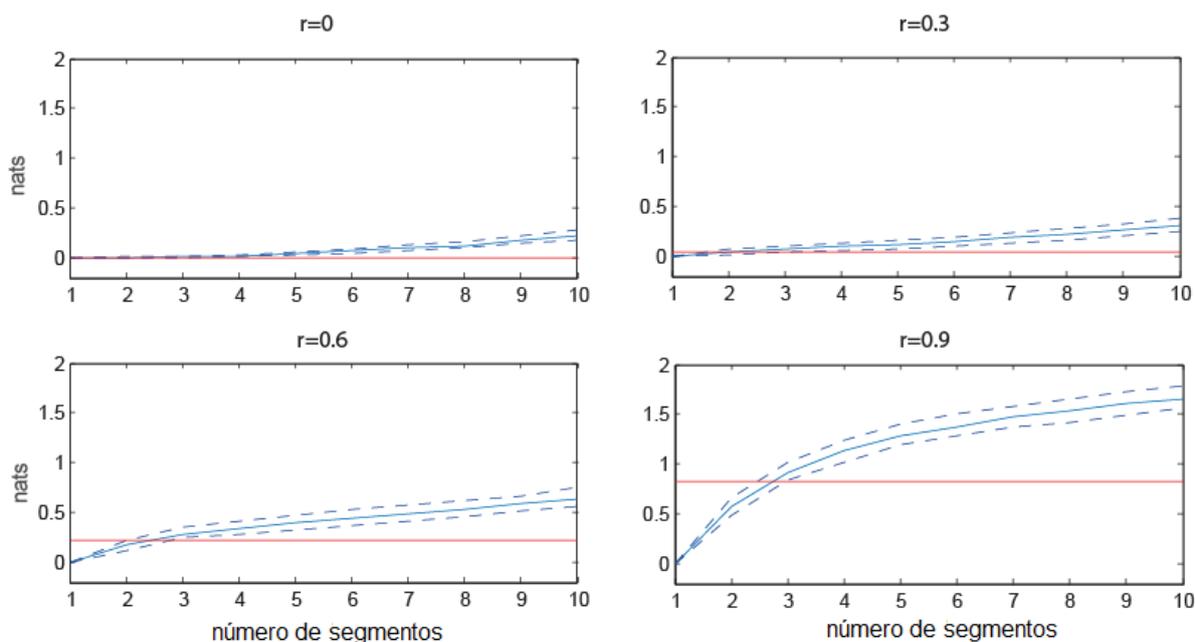


Figura 5.3: Estimativas de informações mútuas usando o estimador ingênuo. As estimativas são traçadas como uma função do número de bins utilizados no procedimento. Os dados possuíam distribuições gaussianas com média zero, variância unitária e coeficientes de correlação  $r = 0$ ,  $r = 0.3$ ,  $r = 0.6$  e  $r = 0.9$ , em cada quadro. As linhas azuis sólidas indicam valores médios em mais de 100 ensaios, cada teste com tamanho de 400 amostras, as linhas azuis tracejadas indicam 10% a 90% dos valores, enquanto as linhas vermelhas indicam o valor analítico de informação mútua.

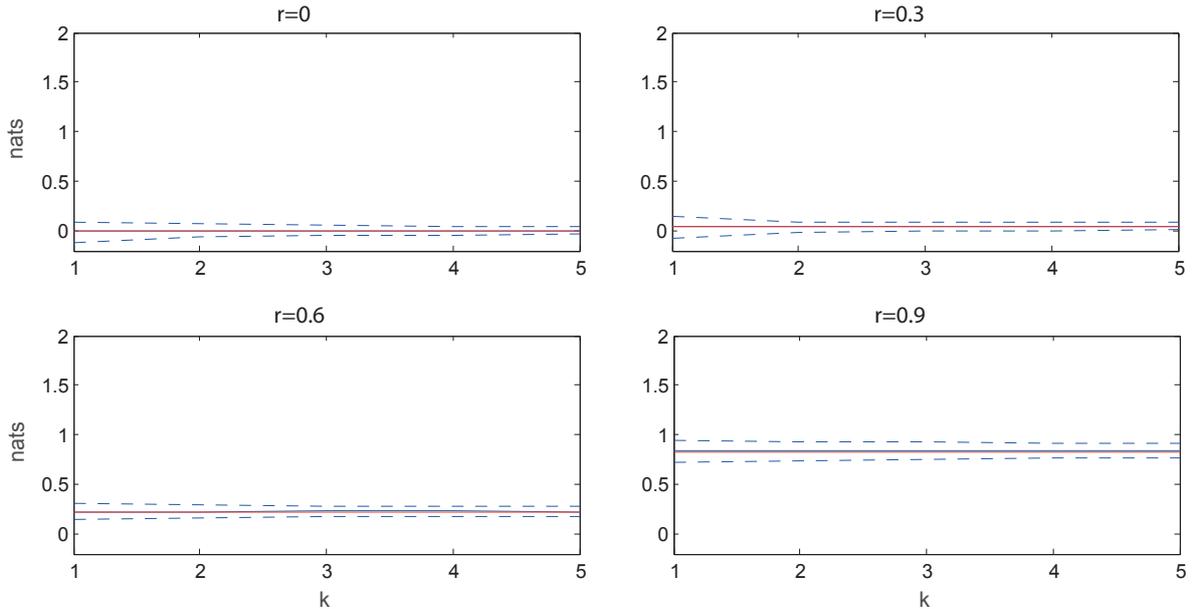


Figura 5.4: Estimativas de informações mútuas usando o estimador KSG. As estimativas são traçadas como uma função do número de bins utilizados no procedimento. Os dados possuíam distribuições gaussianas com média zero, variância unitária e coeficientes de correlação  $r = 0$ ,  $r = 0.3$ ,  $r = 0.6$  e  $r = 0.9$ , em cada quadro. As linhas azuis sólidas indicam valores médios em mais de 100 ensaios, cada teste com tamanho de 400 amostras, as linhas azuis tracejadas indicam 10% a 90% dos valores, enquanto as linhas vermelhas indicam o valor analítico de informação mútua.

Realizamos testes estatísticos para confirmar a hipótese de que a distribuição para cada classe e método seja significativamente diferente (Apêndice B). A escolha por testes estatísticos não-paramétricos se deu por não conhecermos as distribuições dos dados. Para avaliar as diferenças nas medianas dos valores de  $\tau$  com ambos os métodos dentro de uma única classe, foi utilizado o *Wilcoxon signed-rank test*. Os resultados obtidos são apresentados na Tabela 5.1.

Tabela 5.1: Resultados dos testes estatísticos utilizando *Wilcoxon signed-rank test*

$\tau$		
Classes	$p$	$h$
SDL (Naive) $\times$ SDL (KSG)	0.0042	1
EDM (Naive) $\times$ EDM (KSG)	0.0001	1
NDL (Naive) $\times$ NDL (KSG)	0.0005	1
PRL (Naive) $\times$ PRL (KSG)	$1.1767 \times 10^{-06}$	1

Por outro lado, para avaliar as diferenças entre as classes utilizando o método KSG, foi utilizado o *Wilcoxon ranksum test*. Os resultados obtidos são apresentados na Tabela 5.2.

Dessa forma, podemos observar que existem diferenças significativas entre o  $\tau$  estimado com o estimador ingênuo e o  $\tau$  estimado com o estimador KSG para todas as classes

Tabela 5.2: Resultados dos testes estatísticos utilizando *Wilcoxon ranksum test*

$\tau$		
Classes	$p$	$h$
SDL (KSG) $\times$ EDM (KSG)	0.0003	1
SDL (KSG) $\times$ NDL (KSG)	$7.6711 \times 10^{-05}$	1
SDL (KSG) $\times$ PRL (KSG)	0.0041	1
SDL (KSG) $\times$ PTL (KSG)	$7.3734 \times 10^{-05}$	1
EDM (KSG) $\times$ NDL (KSG)	0.1017	0
PRL (KSG) $\times$ EDM (KSG)	0.6084	0
PRL (KSG) $\times$ NDL (KSG)	0.0314	1

consideradas. Quando consideramos apenas o  $\tau$  estimado com o KSG, observamos que as comparações de SDL  $\times$  EDM, SDL  $\times$  NDL, SDL  $\times$  PRL, SDL  $\times$  PTL e PRL  $\times$  NDL apresentaram diferenças estatísticas significativas entre as estimativas. Apenas para os casos EDM  $\times$  NDL e PRL  $\times$  EDM, observamos que não houve diferença significativa.

### 5.1.2 Entropia de Shannon

Nas Figuras 5.5 e 5.6, estão ilustradas a distribuição dos valores médios da entropia de Shannon para sinais de vozes saudáveis (SDL) e patológicas com paralisia (PRL), edema de Reinke (EDM) e nódulos (NDL), utilizando-se o estimador ingênuo e o estimador KSG/KOLE respectivamente.

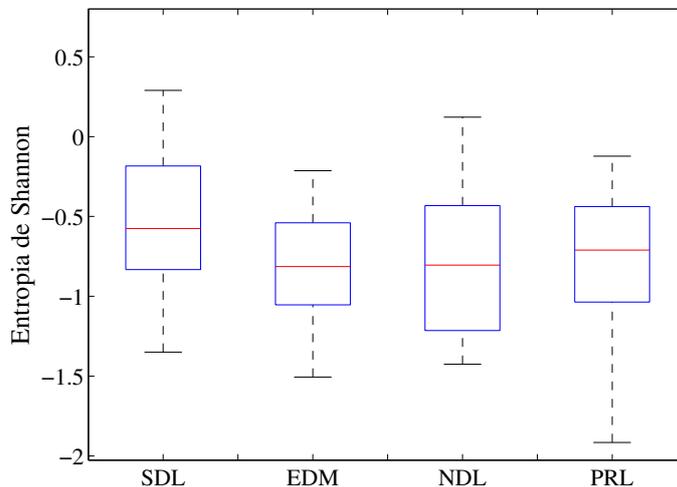


Figura 5.5: Distribuição média dos valores de entropia de Shannon para os sinais de vozes saudáveis (SDL) e patológicos com as patologias edema de Reinke (EDM), nódulo (NDL) e paralisia das pregas vocais (PRL), estimadas com o estimador ingênuo.

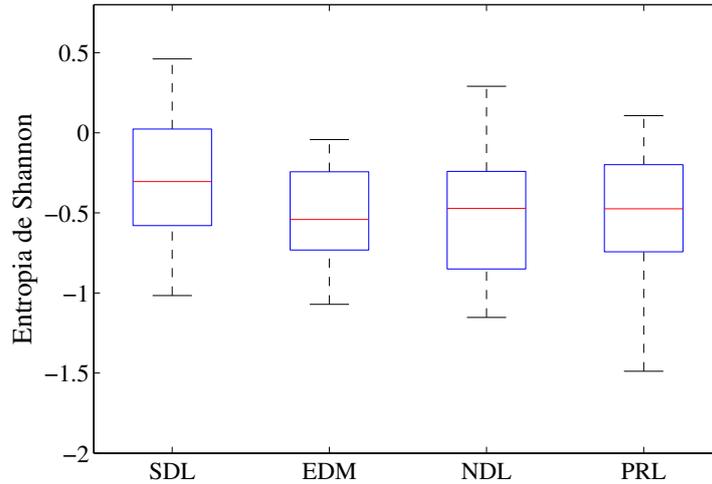


Figura 5.6: Distribuição média dos valores de entropia de Shannon para os sinais de vozes saudáveis (SDL) e patológicos com as patologias edema de Reinke (EDM), nódulo (NDL) e paralisia das pregas vocais (PRL), estimadas com o estimador KOLE.

### Teste dos Estimadores Ingênuo e KOLE

Para testar o desempenho dos estimadores, assim como foi realizado para as medidas anteriores, foram feitas simulações com conjuntos de dados gaussianos bivariados, com média zero e variância unitária. Para esses casos, existe um valor analítico bem estabelecido para a entropia, que pode ser usado para comparar as estimativas. A Figura 5.7 ilustra a simulação com os estimadores ingênuo e KOLE.

$$X \sim \phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (5.2)$$

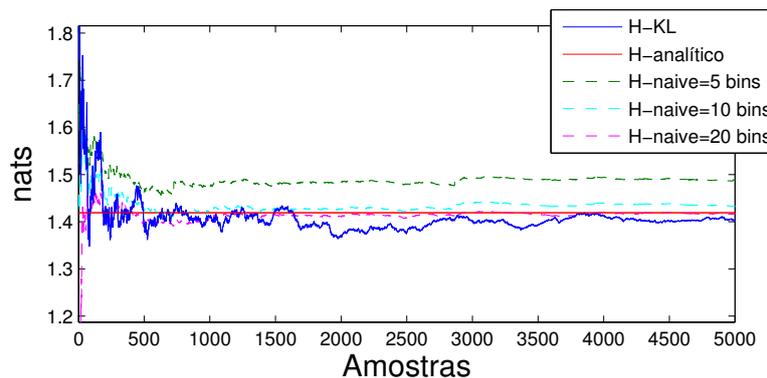


Figura 5.7: Estimativas de entropia de Shannon usando os estimadores KOLE e ingênuo. Os dados possuíam distribuições gaussianas com média zero e variância unitária. A linha vermelha indica o valor analítico para entropia diferencial. A linha em azul sólida indica valores médios estimados utilizando o estimador KOLE. As linhas tracejadas em verde, azul claro e lilás, indicam respectivamente, os valores estimados com o estimador ingênuo, para 5, 10 e 20 bins.

Podemos observar que o estimador ingênuo para a entropia de Shannon converge muito bem para o valor analítico.

Similarmente à medida anterior, realizamos testes estatísticos para confirmar a hipótese de que a distribuição para cada classe e método seja significativamente diferente (Apêndice B). Para avaliar as diferenças nas medianas dos valores de entropia de Shannon com ambos os métodos dentro de uma única classe, foi utilizado o *Wilcoxon signed-rank test*. Os resultados obtidos são apresentados na Tabela 5.3.

Tabela 5.3: Resultados dos testes estatísticos utilizando *Wilcoxon signed-rank test*

<b>Entropia de Shannon</b>		
<b>Classes</b>	<i>p</i>	<i>h</i>
SDL (Naive) × SDL (KOLE)	$3.7149 \times 10^{-10}$	1
EDM (Naive) × EDM (KOLE)	$7.6159 \times 10^{-09}$	1
NDL (Naive) × NDL (KOLE)	$1.9644 \times 10^{-04}$	1
PRL (Naive) × PRL (KOLE)	$3.5039 \times 10^{-06}$	1

Por outro lado, para avaliar as diferenças entre as classes utilizando o método KSG/KOLE, foi utilizado o *Wilcoxon rank-sum test*. Os resultados obtidos são apresentados na Tabela 5.4.

Tabela 5.4: Resultados dos testes estatísticos utilizando *Wilcoxon ranksum test*

<b>Entropia de Shannon</b>		
<b>Classes</b>	<i>p</i>	<i>h</i>
SDL (KOLE) × EDM (KOLE)	0.0037	1
SDL (KOLE) × NDL (KOLE)	$0.0425 \times 10^{-05}$	1
SDL (KOLE) × PRL (KOLE)	0.0216	1
SDL (KOLE) × PTL (KOLE)	0.0017	1
EDM (KOLE) × NDL (KOLE)	0.9505	0
PRL (KOLE) × EDM (KOLE)	0.4207	0
PRL (KOLE) × NDL (KOLE)	0.5409	1

Quando consideramos apenas os valores de entropia de Shannon estimados com o KOLE, a análise dos resultados apresentados indica diferenças estatísticas significativas dos valores da entropia de Shannon entre o grupo de vozes saudáveis e cada um dos três grupos de vozes patológicas. Entretanto, os testes revelaram não haver evidências estatísticas para o uso da entropia de Shannon no sentido de diferenciar entre os grupos de vozes patológicas.

A partir do resultado dos testes estatísticos, podemos concluir que, a entropia de Shannon pode ser usada, em potencial, como um discriminador entre vozes saudáveis e as patologias da laringe consideradas neste trabalho, mas não para diferenciar entre os diferentes tipos de patologia.

## 5.2 Classificação de Sinais de Vozes Saudáveis *Versus* Patológicas Usando Medidas da Análise Não-Linear

Nesta seção são apresentados os resultados obtidos do processo de estimação e classificação de sinais de vozes saudáveis e patológicas realizado a fim de investigar o potencial discriminativo das medidas da análise dinâmica não-linear  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ . O objetivo é comparar o desempenho dos estimadores KSG/KOLE e ingênuo na classificação entre vozes saudáveis e vozes afetadas pelas patologias edema de Reinke, paralisia das pregas vocais e nódulos nas pregas vocais, a fim de observar se, a utilização de um método de estimação mais robusto, eleva os valores de acurácia nos casos de classificação considerados nesta tese.

Para cada caso de classificação considerado: SDL  $\times$  PTL, SDL  $\times$  PRL, SDL  $\times$  EDM, SDL  $\times$  NDL, as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$  foram avaliadas de maneira individual e combinada (todas as medidas agrupadas), para observarmos se a junção de medidas eleva os valores de acurácia. Como mencionado no capítulo 4, foram utilizados dois classificadores diferentes: LDA e SVM, e os resultados para cada classificador é apresentado separadamente em cada caso de classificação considerado.

### 5.2.1 Classificação entre Vozes Saudáveis *versus* Vozes Afetadas por Patologias Laríngeas

Os resultados do processo de classificação utilizando cada característica de forma individual estão apresentados nas Tabelas 5.5, 5.6, para o classificador LDA e nas Tabelas 5.7 e 5.8 para o classificador SVM. Os resultados do processo de classificação utilizando as características de forma combinada estão apresentados nas Tabelas 5.9, 5.10, para o classificador LDA e nas Tabelas 5.11 e 5.12 para o classificador SVM.

#### Características Avaliadas Individualmente

- LDA

Na Tabela 5.5, estão apresentados os valores obtidos de acurácia (Ac), sensibilidade (Sen) e especificidade (Esp), utilizando o classificador LDA, para os sinais de vozes saudáveis (SDL) e patológicas (PTL) utilizando o estimador ingênuo, para as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , e na Tabela 5.6 utilizando o estimador KSG/KOLE. Os sinais de vozes patológicas compreendem, neste caso, os sinais de todas as patologias (edema, paralisia e nódulo) agrupadas em uma mesma classe.

Tabela 5.5: Taxas de Classificação (%) Saudável x Patológico - Estimador Ingênuo.

Vozes Saudáveis x Vozes Patológicas				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	60.66 $\pm$ 4.80	43.71 $\pm$ 5.50	61.69 $\pm$ 3.38	41.95 $\pm$ 2.67
Sensibilidade	62.67 $\pm$ 6.54	46.67 $\pm$ 6.64	61.34 $\pm$ 9.14	60.67 $\pm$ 4.00
Especificidade	57.12 $\pm$ 5.36	35.68 $\pm$ 6.19	61.36 $\pm$ 3.79	33.48 $\pm$ 4.15

Tabela 5.6: Taxas de Classificação (%) Saudável x Patológico - Estimador KSG/KOLE.

<b>Vozes Saudáveis x Vozes Patológicas</b>				
<b>Medidas</b>	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	$68.86 \pm 3.64$	$54.52 \pm 4.36$	$59.44 \pm 2.82$	$73.75 \pm 3.65$
Sensibilidade	$64.67 \pm 8.84$	$52.67 \pm 7.66$	$58.34 \pm 7.23$	$90.66 \pm 4.40$
Especificidade	$70.38 \pm 4.92$	$55.69 \pm 5.49$	$59.55 \pm 5.27$	$65.91 \pm 4.38$

Como pode ser observado, os resultados obtidos utilizando-se as medidas  $\tau$ ,  $PMIM$  e  $K_2$  estimados com o estimador KSG/KOLE apresentaram maiores taxas de acurácia média do classificador quando comparadas com os resultados obtidos com o estimador ingênuo. Além disso, também ofereceu o melhor desempenho nos critérios de sensibilidade média e especificidade média, para ambas as medidas. Apenas para a medida  $H$  o estimador ingênuo e KSG/KOLE apresentaram resultados semelhantes de classificação dentro do desvio padrão.

- SVM

Na Tabela 5.7, estão apresentados os valores obtidos de acurácia, Utilizando o classificador SVM, para os sinais de vozes saudáveis e patológicas utilizando o estimador ingênuo, para as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , e na Tabela 5.8 utilizando o estimador KSG/KOLE.

Tabela 5.7: Taxa de Classificação (%) Saudável x Patológico - Estimador Ingênuo.

<b>Vozes Saudáveis x Vozes Patológicas</b>				
<b>Medidas</b>	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	55.8824	54.5455	51.5152	55.8824

Tabela 5.8: Taxa de Classificação (%) Saudável x Patológico - Estimador KSG/KOLE.

<b>Vozes Saudáveis x Vozes Patológicas</b>				
<b>Medidas</b>	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	69.6970	57.5758	60.6061	78.7879

### Características Avaliadas de Forma Combinada

A combinação das características visa aumentar a quantidade de informação fornecida ao classificador, no intuito de melhorar as taxas de classificação. Para isso, foram avaliadas as combinações entre as medidas da análise não-linear  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , comparando o desempenho dos estimadores KSG/KOLE/KOLE e ingênuo, tanto para o classificador LDA quanto para o SVM.

- LDA

Na Tabela 5.9, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, em sinais de vozes saudáveis e vozes afetadas por patologias, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.10 utilizando o estimador KSG/KOLE.

Tabela 5.9: Taxas de Classificação (%) Saudável x Patológico. Medidas Combinadas - Estimador Ingênuo.

<b>Vozes Saudáveis x Vozes Patológicas</b>		
<b>Medidas</b>	$\tau + PMIM$	$\tau + PMIM + H + K_2$
Acurácia	$63.89 \pm 4.15$	$70.66 \pm 3.15$
Sensibilidade	$69.00 \pm 5.63$	$75.67 \pm 6.27$
Especificidade	$51.32 \pm 5.37$	$68.34 \pm 4.96$

Tabela 5.10: Taxas de Classificação (%) Saudável x Patológico. Medidas Combinadas - Estimador KSG/KOLE.

<b>Vozes Saudáveis x Vozes Patológicas</b>		
<b>Medidas</b>	$\tau + PMIM$	$\tau + PMIM + H + K_2$
Acurácia	$72.93 \pm 4.62$	$76.62 \pm 2.09$
Sensibilidade	$69.00 \pm 5.63$	$88.34 \pm 4.37$
Especificidade	$69.55 \pm 4.92$	$70.84 \pm 2.84$

- SVM

Na Tabela 5.11, estão apresentados os valores obtidos de acurácia, utilizando o classificador SVM, em sinais de vozes saudáveis e vozes afetadas por patologias, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.12 utilizando o estimador KSG/KOLE.

Tabela 5.11: Taxa de Classificação (%) Saudável x Patológico. Medidas Combinadas - Estimador Ingênuo.

<b>Vozes Saudáveis x Vozes Patológicas</b>		
<b>Medidas</b>	$\tau + PMIM$	$\tau + PMIM + H + K_2$
Acurácia	63.6250	56.2500

Os resultados obtidos mostram que, como era esperado, a combinação das características aumenta a quantidade de informação fornecida ao classificador, e dessa forma melhora as taxas de classificação, para ambos os classificadores. É possível ainda observar que o uso de uma maior quantidade de medidas (todas juntas), para este caso de classificação não melhora as taxas de classificação, ou ainda fazem essas taxas caírem.

Tabela 5.12: Taxa de Classificação (%) Saudável x Patológico. Medidas Combinadas - Estimador KSG/KOLE.

Vozes Saudáveis x Vozes Patológicas		
Medidas	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	69.6970	69.6970

## 5.2.2 Classificação entre Vozes Saudáveis *versus* Vozes Afetadas pela Patologia Edema de Reinke

Nesta seção são apresentados os resultados da classificação entre sinais de vozes saudáveis e vozes afetadas por edema nas pregas vocais.

### Características Avaliadas Individualmente

- LDA

Na Tabela 5.13, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, para os sinais de vozes saudáveis e vozes afetadas por edema nas pregas vocais utilizando o estimador ingênuo, para as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , e na Tabela 5.14 utilizando o estimador KSG/KOLE.

Tabela 5.13: Taxas de Classificação (%) Saudável x Edema - Estimador Ingênuo.

Vozes Saudáveis x Vozes com Edema				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	$63.78 \pm 5.51$	$47.55 \pm 6.42$	$65.66 \pm 4.98$	$53.33 \pm 5.08$
Sensibilidade	$65.00 \pm 6.62$	$56.00 \pm 9.55$	$64.00 \pm 4.69$	$69.34 \pm 6.68$
Especificidade	$60.50 \pm 9.77$	$38.00 \pm 7.28$	$68.50 \pm 9.35$	$33.05 \pm 5.64$

Tabela 5.14: Taxas de Classificação (%) Saudável x Edema - Estimador KSG/KOLE.

Vozes Saudáveis x Vozes com Edema				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	$69.22 \pm 5.51$	$52.33 \pm 5.26$	$62.77 \pm 2.98$	$75.22 \pm 2.79$
Sensibilidade	$66.00 \pm 7.18$	$49.34 \pm 7.06$	$62.00 \pm 5.82$	$88.67 \pm 4.31$
Especificidade	$72.50 \pm 5.69$	$56.50 \pm 8.66$	$63.00 \pm 7.93$	$58.50 \pm 7.93$

também para este caso de classificação, os resultados obtidos utilizando-se as medidas  $\tau$ ,  $PMIM$  e  $K_2$  estimados com o estimador KSG/KOLE apresentaram maiores taxas de acurácia média do classificador quando comparadas com os resultados obtidos com o estimador ingênuo. Apenas para a medida  $H$  o estimador ingênuo apresentou melhor desempenho.

- SVM

Na Tabela 5.15, estão apresentados os valores obtidos de acurácia, Utilizando o classificador SVM, para os sinais de vozes saudáveis e vozes afetadas por edema nas pregas vocais utilizando o estimador ingênuo, para as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , e na Tabela 5.16 utilizando o estimador KSG/KOLE.

Tabela 5.15: Taxa de Classificação (%) Saudável x Edema - Estimador Ingênuo.

Vozes Saudáveis x Vozes com Edema				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	71.6418	44.6154	57.3529	50.7463

Tabela 5.16: Taxa de Classificação (%) Saudável x Edema - Estimador KSG/KOLE.

Vozes Saudáveis x Vozes com edema				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	74.6268	50.7463	61.194	77.6119

Para a classificação utilizando o classificador SVM, os resultados obtidos, utilizando-se todas as medidas individualmente, estimados com o estimador KSG/KOLE apresentaram maiores taxas de acurácia média do classificador quando comparadas com os resultados obtidos com o estimador ingênuo.

### Características Avaliadas de Forma Combinada

- LDA

Na Tabela 5.17, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, em sinais de vozes saudáveis e vozes afetadas por edema nas pregas vocais, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.18 utilizando o estimador KSG/KOLE.

Tabela 5.17: Taxas de Classificação (%) Saudável x Edema. Medidas Combinadas - Estimador Ingênuo.

Vozes Saudáveis x Vozes com Edema		
Medidas	$\tau + PMIM$	$\tau + PMIM + H + K_2$
Acurácia	$65.11 \pm 5.72$	$72.00 \pm 4.88$
Sensibilidade	$68.00 \pm 6.37$	$75.33 \pm 5.95$
Especificidade	$60.50 \pm 10.21$	$66.50 \pm 10.62$

Tabela 5.18: Taxas de Classificação (%) Saudável x Edema. Medidas Combinadas - Estimador KSG/KOLE.

Vozes Saudáveis x Vozes com Edema		
Medidas	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	$72.11 \pm 5.10$	$76.11 \pm 4.66$
Sensibilidade	$67.67 \pm 8.13$	$81.67 \pm 6.80$
Especificidade	$66.00 \pm 5.95$	$70.00 \pm 7.15$

- SVM

Na Tabela 5.19, estão apresentados os valores obtidos de acurácia (Ac), , utilizando o classificador SVM, em sinais de vozes saudáveis e vozes afetadas por edema nas pregas vocais, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.20 utilizando o estimador KSG/KOLE.

Tabela 5.19: Taxa de Classificação (%) Saudável x Edema. Medidas Combinadas - Estimador Ingênuo.

Vozes Saudáveis x Vozes com Edema		
Medidas	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	61.7647	62.6866

Tabela 5.20: Taxa de Classificação (%) Saudável x Edema. Medidas Combinadas - Estimador KSG/KOLE.

Vozes Saudáveis x Vozes com Edema		
Medidas	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	68.6567	76.4706

Também para este caso de classificação, os resultados obtidos mostram que, a combinação das características melhora as taxas de classificação, para ambos os classificadores, e os resultados obtidos com o estimador KSG/KOLE foi superior ao estimador ingênuo.

### 5.2.3 Classificação entre Vozes Saudáveis *versus* Vozes Afetadas pela Patologia Nódulo nas Pregas Vocais

Nesta seção são apresentados os resultados da classificação entre sinais de vozes saudáveis e vozes afetadas pela patologia nódulo nas pregas vocais.

## Características Avaliadas Individualmente

- LDA

Na Tabela 5.21, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, para os sinais de vozes saudáveis e vozes afetadas por nódulo nas pregas vocais utilizando o estimador ingênuo, para as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , e na Tabela 5.22 utilizando o estimador KSG/KOLE.

Tabela 5.21: Taxas de Classificação (%) Saudável x Nódulo - Estimador Ingênuo.

Vozes Saudáveis x Vozes com Nódulo				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	73.32 ± 5.34	65.00 ± 3.44	63.21 ± 5.38	41.96 ± 7.35
Sensibilidade	81.67 ± 6.05	70.00 ± 4.13	62.00 ± 7.19	45.34 ± 8.76
Especificidade	70.00 ± 13.34	50.00 ± 12.91	65.00 ± 10.68	30.00 ± 11.06

Tabela 5.22: Taxas de Classificação (%) Saudável x Nódulo - Estimador KSG/KOLE.

Vozes Saudáveis x Vozes com Nódulo				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	80.00 ± 7.13	57.85 ± 6.55	61.96 ± 5.22	84.64 ± 4.39
Sensibilidade	81.67 ± 7.19	55.00 ± 6.84	60.34 ± 6.68	90.34 ± 3.24
Especificidade	75.00 ± 11.18	60.00 ± 12.48	70.00 ± 11.06	70.00 ± 13.34

Como pode ser observado, os resultados obtidos utilizando-se as medidas  $\tau$ ,  $PMIM$  e  $K_2$  estimados com o estimador KSG/KOLE apresentaram maiores taxas de acurácia média do classificador quando comparadas com os resultados obtidos com o estimador ingênuo. Além disso, também ofereceu o melhor desempenho nos critérios de sensibilidade média e especificidade média, para ambas as medidas. Apenas para a medida  $H$  o estimador ingênuo mostrou melhor desempenho.

- SVM

Na Tabela 5.23, estão apresentados os valores obtidos de acurácia ( $Ac$ ), Utilizando o classificador SVM, para os sinais de vozes saudáveis e afetadas por nódulo nas pregas vocais utilizando o estimador ingênuo, para as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , e na Tabela 5.24 utilizando o estimador KSG/KOLE.

Tabela 5.23: Taxa de Classificação (%) Saudável x Nódulo - Estimador Ingênuo.

Vozes Saudáveis x Vozes com Nódulo				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	61.5385	50.00	57.6923	42.3077

Tabela 5.24: Taxa de Classificação (%) Saudável x Nódulo - Estimador KSG/KOLE.

<b>Vozes Saudáveis x Vozes com Nódulo</b>				
<b>Medidas</b>	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	65.3846	53.8462	61.5385	76.923

Como pode ser observado, para a classificação utilizando o classificador SVM, os resultados obtidos utilizando-se as medidas  $\tau$  e  $K_2$  estimados com o estimador KSG/KOLE apresentaram maiores taxas de acurácia média do classificador quando comparadas com os resultados obtidos com o estimador ingênuo. Para as medidas  $PMIM$  e  $H$  os resultados de classificação utilizando o estimador ingênuo apresentou desempenho superior.

### Características Avaliadas de Forma Combinada

- LDA

Na Tabela 5.25, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, em sinais de vozes saudáveis e vozes afetadas por nódulo nas pregas vocais, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.26 utilizando o estimador KSG/KOLE.

Tabela 5.25: Taxas de Classificação (%) Saudável x Nódulo. Medidas Combinadas - Estimador Ingênuo.

<b>Vozes Saudáveis x Vozes com Nódulo</b>		
<b>Medidas</b>	$\tau + PMIM$	$\tau + PMIM + H + K_2$
Acurácia	$77.14 \pm 4.86$	$84.14 \pm 3.95$
Sensibilidade	$81.00 \pm 4.92$	$85.33 \pm 3.19$
Especificidade	$60.00 \pm 14.54$	$80.00 \pm 7.64$

Tabela 5.26: Taxas de Classificação (%) Saudável x Nódulo. Medidas Combinadas - Estimador KSG/KOLE.

<b>Vozes Saudáveis x Vozes com Nódulo</b>		
<b>Medidas</b>	$\tau + PMIM$	$\tau + PMIM + H + K_2$
Acurácia	$79.82 \pm 6.18$	$90.00 \pm 3.05$
Sensibilidade	$75.67 \pm 7.66$	$94.34 \pm 4.16$
Especificidade	$70.00 \pm 11.06$	$80.00 \pm 8.17$

Também para este caso de classificação, os resultados obtidos mostram que, a combinação das características melhora as taxas de classificação, e os resultados obtidos com o estimador KSG/KOLE foi superior ao estimador ingênuo.

- SVM

Na Tabela 5.27, estão apresentados os valores obtidos de acurácia ( $A_c$ ), utilizando o classificador SVM, em sinais de vozes saudáveis e vozes afetadas por nódulo nas pregas vocais, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.28 utilizando o estimador KSG/KOLE.

Tabela 5.27: Taxa de Classificação (%) Saudável x Nódulo. Medidas Combinadas - Estimador Ingênuo.

Vozes Saudáveis x Vozes com Nódulo		
Medidas	$\tau + PMIM$	$\tau + PMIM + H + K_2$
Acurácia	69.2308	65.3846

Tabela 5.28: Taxa de Classificação (%) Saudável x Nódulo. Medidas Combinadas - Estimador KSG/KOLE.

Vozes Saudáveis x Vozes com Nódulo		
Medidas	$\tau + PMIM$	$\tau + PMIM + H + K_2$
Acurácia	76.923	76.9231

Como pode ser observado, para este caso de classificação com o classificador SVM, a incorporação de mais duas medidas (caso onde todas as medidas foram agrupadas) não aumenta as taxas de acurácia do classificador. Ainda assim, todos os resultados obtidos com o uso do estimador KSG/KOLE apresentaram maiores taxas de acurácia quando comparados com os resultados obtidos com o estimador ingênuo.

#### 5.2.4 Classificação entre Vozes Saudáveis *versus* Vozes Afetadas pela Patologia Paralisia das Pregas Vocais

Nesta seção são apresentados os resultados da classificação entre sinais de vozes saudáveis e vozes afetadas pela patologia paralisia das pregas vocais.

##### Características Avaliadas Individualmente

- LDA

Na Tabela 5.29, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, para os sinais de vozes saudáveis e vozes afetadas por paralisia das pregas vocais utilizando o estimador ingênuo, para as medidas  $\tau$ , PMIM, H e  $K_2$ , e na Tabela 5.30 utilizando o estimador KSG/KOLE.

Tabela 5.29: Taxas de Classificação (%) Saudável x Paralisia - Estimador Ingênuo.

<b>Vozes Saudáveis x Vozes com Paralisia</b>				
<b>Medidas</b>	$\tau$	<i>PMIM</i>	<i>H</i>	$K_2$
Acurácia	68.64 ± 5.52	53.27 ± 4.79	57.00 ± 3.61	59.09 ± 5.93
Sensibilidade	60.67 ± 4.92	34.00 ± 3.89	59.34 ± 6.61	45.34 ± 7.96
Especificidade	66.00 ± 9.11	53.67 ± 8.08	54.00 ± 4.00	73.67 ± 7.51

Tabela 5.30: Taxas de Classificação (%) Saudável x Paralisia - Estimador KSG/KOLE.

<b>Vozes Saudáveis x Vozes com Paralisia</b>				
<b>Medidas</b>	$\tau$	<i>PMIM</i>	<i>H</i>	$K_2$
Acurácia	70.18 ± 5.76	63.91 ± 4.12	54.90 ± 5.88	80.90 ± 3.10
Sensibilidade	67.34 ± 7.21	61.00 ± 6.71	55.00 ± 8.62	90.67 ± 4.00
Especificidade	64.00 ± 6.14	67.67 ± 4.66	54.67 ± 7.00	71.00 ± 6.40

- SVM

Na Tabela 5.31, estão apresentados os valores obtidos de acurácia, Utilizando o classificador SVM, para os sinais de vozes saudáveis e afetadas por paralisia das pregas vocais utilizando o estimador ingênuo, para as medidas  $\tau$ , PMIM, H e  $K_2$ , e na Tabela 5.32 utilizando o estimador KSG/KOLE.

Tabela 5.31: Taxa de Classificação (%) Saudável x Paralisia - Estimador Ingênuo.

<b>Vozes Saudáveis x Vozes com Paralisia</b>				
<b>Medidas</b>	$\tau$	<i>PMIM</i>	<i>H</i>	$K_2$
Acurácia	63.0137	56.1644	54.7945	63.0137

Tabela 5.32: Taxa de Classificação (%) Saudável x Paralisia - Estimador KSG/KOLE.

<b>Vozes Saudáveis x Vozes com Paralisia</b>				
<b>Medidas</b>	$\tau$	<i>PMIM</i>	<i>H</i>	$K_2$
Acurácia	69.863	63.0137	50.6849	76.7123

Como pode ser observado, os resultados obtidos utilizando-se as medidas  $\tau$ , PMIM e  $K_2$  estimados com o estimador KSG/KOLE apresentaram maiores taxas de acurácia média do classificador quando comparadas com os resultados obtidos com o estimador ingênuo. Apenas para a medida *H* o estimador ingênuo mostrou melhor desempenho, para ambos os classificadores.

### Características Avaliadas de Forma Combinada

- LDA

Na Tabela 5.33, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, em sinais de vozes saudáveis e vozes afetadas por paralisia das pregas vocais, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.34 utilizando o estimador KSG/KOLE.

Tabela 5.33: Taxas de Classificação (%) Saudável x Paralisia. Medidas Combinadas - Estimador Ingênuo.

<b>Vozes Saudáveis x Vozes com Paralisia</b>		
<b>Medidas</b>	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	$65.18 \pm 5.39$	$68.45 \pm 4.09$
Sensibilidade	$76.67 \pm 6.21$	$73.00 \pm 6.06$
Especificidade	$53.34 \pm 7.03$	$63.67 \pm 5.97$

Tabela 5.34: Taxas de Classificação (%) Saudável x Paralisia. Medidas Combinadas - Estimador KSG/KOLE.

<b>Vozes Saudáveis x Vozes com Paralisia</b>		
<b>Medidas</b>	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	$75.54 \pm 4.81$	$79.64 \pm 3.79$
Sensibilidade	$74.34 \pm 7.86$	$88.67 \pm 3.11$
Especificidade	$73.00 \pm 5.27$	$70.67 \pm 6.96$

- SVM

Na Tabela 5.35, estão apresentados os valores obtidos de acurácia, utilizando o classificador SVM, em sinais de vozes saudáveis e vozes afetadas por paralisia das pregas vocais, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.36 utilizando o estimador KSG/KOLE.

Tabela 5.35: Taxa de Classificação (%) Saudável x Paralisia. Medidas Combinadas - Estimador Ingênuo.

<b>Vozes Saudáveis x Vozes com Paralisia</b>		
<b>Medidas</b>	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	68.4932	61.6438

Tabela 5.36: Taxa de Classificação (%) Saudável x Paralisia. Medidas Combinadas - Estimador KSG/KOLE.

Vozes Saudáveis x Vozes com Paralisia		
Medidas	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	71.2328	73.9726

Também para este caso de classificação, os resultados obtidos mostram que, a combinação das características melhora as taxas de classificação, para ambos os classificadores, e os resultados obtidos com o estimador KSG/KOLE foi superior ao estimador ingênuo.

## Discussão dos Resultados

Neste capítulo foram apresentados os resultados da classificação entre vozes saudáveis e patológicas. Como apresentado nas tabelas, podemos observar que, para quase todos os casos de classificação considerados, as medidas  $\tau$ ,  $PMIM$  e  $K_2$ , avaliadas individualmente, utilizando o estimador KSG/KOLE, apresentaram maiores valores de acurácia média quando comparados com o estimador ingênuo para ambos os classificadores empregados. Em alguns casos as taxas de classificação foram superiores a 80% como pode ser observado nas Figuras 5.8 e 5.9.

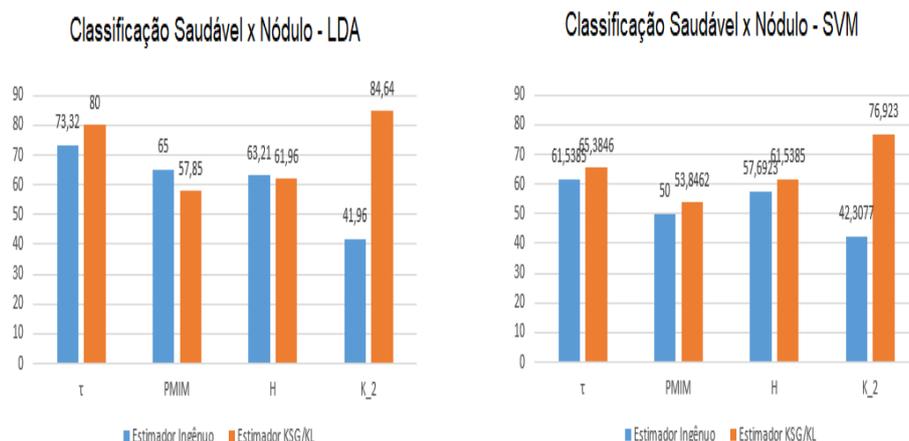


Figura 5.8: Classificação Saudável x Nódulo - Medidas Individuais

Apenas para a medida  $H$  os resultados obtidos utilizando o estimador ingênuo foi superior ao KSG/KOLE, o que pode ser explicado uma vez que, devido a dimensionalidade dos dados, a estimação da entropia de Shannon com o estimador ingênuo converge muito bem para o valor analítico, como foi apresentado na seção 5.1.2 deste capítulo.

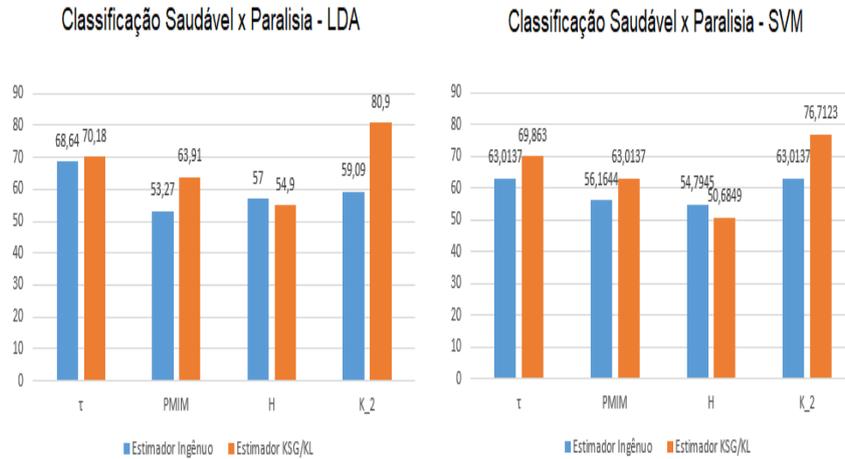


Figura 5.9: Classificação Saudável x Paralisia - Medidas Individuais

Podemos observar também que a combinação das medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$  utilizando o estimador KSG/KOLE apresentaram maiores taxas de classificação, quando comparadas ao estimador ingênuo, com taxas acima de 70%, o que é bastante interessante, pois o uso de apenas quatro medidas foram capazes de fazer uma boa classificação. O melhor caso de classificação foi para classe SDL x NDL, onde a combinação de todas as medidas Utilizando o estimador KSG e classificador LDA produziu uma acurácia de 90% como pode ser observado na Figura 5.10.

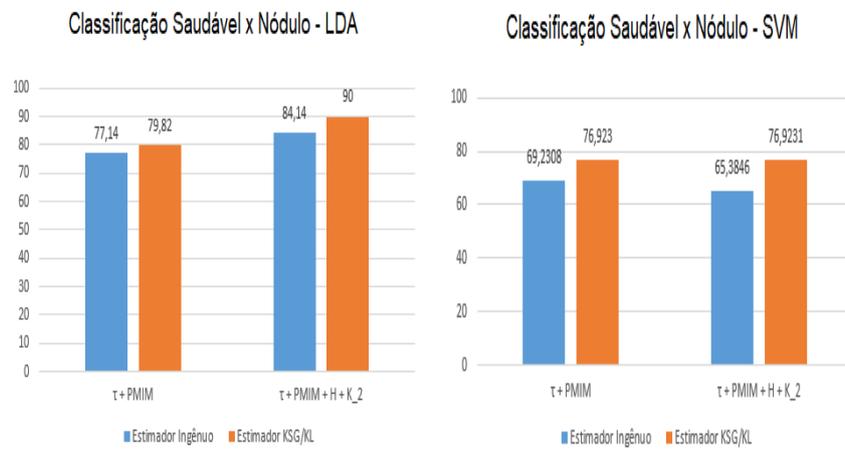


Figura 5.10: Classificação Saudável x Nódulo - Medidas Combinadas

### 5.3 Classificação de Sinais de Vozes Patológicas Usando Medidas da Análise Não-Linear

Nesta seção são apresentados os resultados obtidos do processo de estimação e classificação de sinais de vozes patológicas realizado a fim de investigar o potencial discriminativo das medidas da análise dinâmica não-linear  $\tau, PMIM, H$  e  $K_2$ . Este tipo de classificação

tem sido menos explorado na literatura, uma vez que as mediadas comumente testadas apresentam pouco caráter discriminatório.

O objetivo é comparar o desempenho dos estimadores KSG/KOLE e ingênuo na classificação entre vozes afetadas pelas patologias edema de Reinke, paralisia das pregas vocais e nódulos nas pregas vocais, a fim de observar se, a utilização de um método de estimação mais robusto, eleva os valores de acurácia na diferenciação entre tipos de patologias.

Para cada caso de classificação considerado: PRL  $\times$  EDM, PRL  $\times$  NDL, e EDM  $\times$  NDL, as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$  foram avaliadas de maneira individual e combinada (todas as medidas agrupadas), para observarmos se a junção de medidas eleva os valores de acurácia. Similarmente a seção anterior, foram utilizados dois classificadores diferentes: LDA e SVM, e os resultados para cada classificador é apresentado separadamente em cada caso de classificação considerado.

### 5.3.1 Classificação entre Vozes Afetadas por Paralisia das Pregas Vocais *versus* Vozes Afetadas por Nódulo nas Pregas Vocais

Os resultados do processo de classificação utilizando cada característica de forma individual estão apresentados nas Tabelas 5.37, 5.38, para o classificador LDA e nas Tabelas 5.39 e 5.40 para o classificador SVM. Os resultados do processo de classificação utilizando as características de forma combinada estão apresentados nas Tabelas 5.41, 5.42, para o classificador LDA e nas Tabelas 5.43 e 5.44 para o classificador SVM.

#### Características Avaliadas Individualmente

- LDA

Na Tabela 5.37, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, com o estimador ingênuo, para as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , e na Tabela 5.38 utilizando o estimador KSG/KOLE.

Tabela 5.37: Taxas de Classificação (%) Paralisia x Nódulo - Estimador Ingênuo.

Vozes com Paralisia x Vozes com Nódulo				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	57.14 $\pm$ 6.39	74.28 $\pm$ 4.16	50.00 $\pm$ 6.48	54.14 $\pm$ 5.21
Sensibilidade	61.33 $\pm$ 7.77	81.34 $\pm$ 5.16	51.00 $\pm$ 7.74	67.00 $\pm$ 5.27
Especificidade	50.00 $\pm$ 12.92	60.00 $\pm$ 12.48	40.00 $\pm$ 12.48	30.00 $\pm$ 13.33

Tabela 5.38: Taxas de Classificação (%) Paralisia x Nódulo - Estimador KSG/KOLE.

<b>Voices com Paralisia x Voices com Nódulo</b>				
<b>Medidas</b>	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	64.29 ± 4.39	71.49 ± 7.37	44.29 ± 6.19	67.14 ± 5.24
Sensibilidade	66.67 ± 8.02	69.67 ± 7.24	47.67 ± 8.26	74.00 ± 5.81
Especificidade	65.00 ± 13.03	75.00 ± 11.18	35.00 ± 13.02	50.00 ± 14.91

- SVM

Na Tabela 5.39, estão apresentados os valores obtidos de acurácia, Utilizando o classificador SVM, para os sinais de vozes saudáveis e patológicas utilizando o estimador ingênuo, para as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , e na Tabela 5.40 utilizando o estimador KSG/KOLE.

Tabela 5.39: Taxas de Classificação (%) Paralisia x Nódulo - Estimador Ingênuo.

<b>Voices com Paralisia x Voices com Nódulo</b>				
<b>Medidas</b>	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	53.8462	61.5385	42.3077	46.1538

Tabela 5.40: Taxa de Classificação (%) Paralisia x Nódulo - Estimador KSG/KOLE.

<b>Voices com Paralisia x Voices com Nódulo</b>				
<b>Medidas</b>	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	61.5385	65.3846	46.1538	65.3846

Para a classificação utilizando o classificador SVM, os resultados obtidos utilizando-se as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$  estimados com o estimador KSG/KOLE apresentaram maiores taxas de acurácia média do classificador quando comparadas com os resultados obtidos com o estimador ingênuo.

### Características Avaliadas de Forma Combinada

- LDA

Na Tabela 5.41, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, em sinais de vozes afetadas por paralisia das pregas vocais e vozes afetadas por edema nas pregas vocais, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.42 utilizando o estimador KSG/KOLE.

Tabela 5.41: Taxas de Classificação (%) Paralisia x Nódulo. Medidas Combinadas - Estimador Ingênuo.

<b>Vozes com Paralisia x Vozes com Nódulo</b>		
<b>Medidas</b>	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	$61.42 \pm 7.39$	$57.14 \pm 6.39$
Sensibilidade	$70.67 \pm 8.97$	$60.34 \pm 6.17$
Especificidade	$40.00 \pm 12.48$	$50.00 \pm 12.91$

Tabela 5.42: Taxas de Classificação (%) Paralisia x Nódulo. Medidas Combinadas - Estimador KSG/KOLE.

<b>Vozes com Paralisia x Vozes com Nódulo</b>		
<b>Medidas</b>	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	$69.71 \pm 6.46$	$67.14 \pm 4.29$
Sensibilidade	$65.67 \pm 7.10$	$65.34 \pm 6.51$
Especificidade	$69.00 \pm 12.48$	$75.00 \pm 8.34$

- SVM

Na Tabela 5.43, estão apresentados os valores obtidos de acurácia, utilizando o classificador SVM, em sinais de vozes afetadas por paralisia das pregas vocais e vozes afetadas por edema nas pregas vocais, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.44 utilizando o estimador KSG/KOLE.

Tabela 5.43: Taxa de Classificação (%) Paralisia x Nódulo. Medidas Combinadas - Estimador Ingênuo.

<b>Vozes com Paralisia x Vozes com Nódulo</b>		
<b>Medidas</b>	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	50.0000	53.8462

Tabela 5.44: Taxa de Classificação (%) Paralisia x Nódulo. Medidas Combinadas - Estimador KSG/KOLE.

<b>Vozes com Paralisia x Vozes com Nódulo</b>		
<b>Medidas</b>	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	61.5384	65.3846

Os resultados obtidos mostram que, como era esperado, a combinação das características aumenta a quantidade de informação fornecida ao classificador, e dessa forma melhora as taxas de classificação, para ambos os classificadores. É possível ainda observar que o uso do estimador KSG/KOLE, para este caso de classificação melhora as taxas de acurácia para ambos os classificadores.

### 5.3.2 Classificação entre Vozes Afetadas por Paralisia das Pregas Vocais *versus* Vozes Afetadas pela Patologia Edema de Reike

Nesta seção são apresentados os resultados da classificação entre sinais de vozes saudáveis e vozes afetadas por edema nas pregas vocais.

#### Características Avaliadas Individualmente

- LDA

Na Tabela 5.45, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, para os sinais de vozes afetadas por paralisia das pregas vocais e vozes afetadas por edema nas pregas vocais utilizando o estimador ingênuo, para as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , e na Tabela 5.46 utilizando o estimador KSG/KOLE.

Tabela 5.45: Taxas de Classificação (%) Paralisia x Edema - Estimador Ingênuo.

Vozes com Paralisia x Vozes com Edema				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	55.56 ± 5.46	63.22 ± 3.20	43.77 ± 6.01	57.33 ± 4.05
Sensibilidade	52.00 ± 6.43	60.00 ± 5.91	44.67 ± 7.91	55.67 ± 8.58
Especificidade	58.50 ± 8.76	48.50 ± 7.20	43.00 ± 8.28	35.50 ± 6.60

Tabela 5.46: Taxas de Classificação (%) Paralisia x Edema - Estimador KSG/KOLE.

Vozes com Paralisia x Vozes com Edema				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	43.44 ± 3,99	65.77 ± 2.91	58.33 ± 4.67	55.33 ± 6.65
Sensibilidade	48.34 ± 5.43	63.34 ± 6.78	56.34 ± 9.36	60.67 ± 8.42
Especificidade	38.00 ± 5.01	68.50 ± 6.83	62.00 ± 7.86	47.50 ± 10.76

Para este caso de classificação, os resultados obtidos utilizando-se as medidas  $PMIM$  e  $H$ , estimados com o estimador KSG/KOLE apresentaram maiores taxas de acurácia média do classificador quando comparadas com os resultados obtidos com o estimador ingênuo. Para as medidas  $\tau$  e  $K_2$  o estimador ingênuo apresentou melhor desempenho.

- SVM

Na Tabela 5.47, estão apresentados os valores obtidos de acurácia, Utilizando o classificador SVM, para os sinais de vozes afetadas por paralisia das pregas vocais e vozes afetadas por edema nas pregas vocais utilizando o estimador ingênuo, para as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , e na Tabela 5.48 utilizando o estimador KSG/KOLE.

Tabela 5.47: Taxa de Classificação (%) Paralisia x Edema - Estimador Ingênuo.

<b>Vozes com Paralisia x Vozes com Edema</b>				
<b>Medidas</b>	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	52.9412	55.8224	51.4706	55.8823

Tabela 5.48: Taxa de Classificação (%) Paralisia x Edema - Estimador KSG/KOLE.

<b>Vozes com Paralisia x Vozes com edema</b>				
<b>Medidas</b>	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	59.7014	58.8235	54.4118	58.8235

Para a classificação utilizando o classificador SVM, os resultados obtidos, utilizando-se todas as medidas individualmente, estimados com o estimador KSG/KOLE apresentaram maiores taxas de acurácia média do classificador quando comparadas com os resultados obtidos com o estimador ingênuo.

### Características Avaliadas de Forma Combinada

- LDA

Na Tabela 5.49, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, em sinais de vozes afetadas por paralisia das pregas vocais e vozes afetadas por edema nas pregas vocais, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.50 utilizando o estimador KSG/KOLE.

Tabela 5.49: Taxas de Classificação (%) Paralisia x Edema. Medidas Combinadas - Estimador Ingênuo.

<b>Vozes com Paralisia x Vozes com Edema</b>		
<b>Medidas</b>	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	$53.44 \pm 3.33$	$54.33 \pm 4.12$
Sensibilidade	$50.00 \pm 5.73$	$59.34 \pm 3.77$
Especificidade	$57.50 \pm 9.81$	$49.00 \pm 9.57$

Tabela 5.50: Taxas de Classificação (%) Paralisia x Edema. Medidas Combinadas - Estimador KSG/KOLE.

Vozes com Paralisia x Vozes com Edema		
Medidas	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	67,77 ± 4,53	61,44 ± 5,42
Sensibilidade	63,34 ± 7,98	59,67 ± 6,78
Especificidade	64,00 ± 8,87	63,00 ± 5,58

- SVM

Na Tabela 5.51, estão apresentados os valores obtidos de acurácia (Ac), utilizando o classificador SVM, em sinais de vozes afetadas por paralisia das pregas vocais e vozes afetadas por edema nas pregas vocais, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.52 utilizando o estimador KSG/KOLE.

Tabela 5.51: Taxa de Classificação (%) Paralisia x Edema. Medidas Combinadas - Estimador Ingênuo.

Vozes com Paralisia x Vozes com Edema		
Medidas	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	50.0000	58.8235

Tabela 5.52: Taxa de Classificação (%) Paralisia x Edema. Medidas Combinadas - Estimador KSG/KOLE.

Vozes com Paralisia x Vozes com Edema		
Medidas	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	58.8235	64.7058

Também para este caso de classificação, os resultados obtidos mostram que, a combinação das características melhora as taxas de classificação, para ambos os classificadores, e os resultados obtidos com o estimador KSG/KOLE foi superior ao estimador ingênuo.

### 5.3.3 Classificação entre Vozes Afetadas por Edema de Reike nas Pregas Vocais *versus* Vozes Afetadas por Nódulo nas Pregas Vocais

Nesta seção são apresentados os resultados da classificação entre sinais de vozes afetadas pela patologia edema de reike e vozes afetadas pela patologia nódulo nas pregas vocais.

## Características Avaliadas Individualmente

- LDA

Na Tabela 5.53, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, para os sinais de vozes afetadas por edema de Reike nas pregas vocais e vozes afetadas por nódulo nas pregas vocais utilizando o estimador ingênuo, para as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , e na Tabela 5.54 utilizando o estimador KSG/KOLE.

Tabela 5.53: Taxas de Classificação (%) Edema x Nódulo - Estimador Ingênuo.

Vozes com Edema x Vozes com Nódulo				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	59.05 ± 7.31	52.85 ± 6.69	38.81 ± 3.66	43.81 ± 5.69
Sensibilidade	61.05 ± 9.84	53.50 ± 8.27	48.50 ± 6.55	36.50 ± 8.67
Especificidade	55.00 ± 8.98	50.00 ± 12.92	15.00 ± 7.64	60.00 ± 12.48

Tabela 5.54: Taxas de Classificação (%) Edema x Nódulo - Estimador KSG/KOLE.

Vozes com Edema x Vozes com Nódulo				
Medidas	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	59.52 ± 6.83	66.19 ± 7.11	35.23 ± 4.97	54.76 ± 5.97
Sensibilidade	63.50 ± 10.06	69.00 ± 7.18	44.00 ± 7.99	52.00 ± 10.06
Especificidade	45.00 ± 15.73	60.00 ± 7.28	20.00 ± 11.06	60.00 ± 10.00

Como pode ser observado, os resultados obtidos utilizando-se as medidas  $PMIM$  e  $K_2$  estimados com o estimador KSG/KOLE apresentaram maiores taxas de acurácia média do classificador quando comparadas com os resultados obtidos com o estimador ingênuo. Além disso, também ofereceu o melhor desempenho nos critérios de sensibilidade média e especificidade média, para ambas as medidas. Apenas para a medida  $H$  o estimador ingênuo mostrou melhor desempenho. Para a medida  $\tau$  ambos os estimadores apresentaram desempenho semelhante.

- SVM

Na Tabela 5.55, estão apresentados os valores obtidos de acurácia ( $Ac$ ), Utilizando o classificador SVM, para os sinais de vozes afetadas por edema de Reike nas pregas vocais e afetadas por nódulo nas pregas vocais utilizando o estimador ingênuo, para as medidas  $\tau$ ,  $PMIM$ ,  $H$  e  $K_2$ , e na Tabela 5.56 utilizando o estimador KSG/KOLE.

Tabela 5.55: Taxa de Classificação (%) Edema x Nódulo - Estimador Ingênuo.

<b>Vozes com Edema x Vozes com Nódulo</b>				
<b>Medidas</b>	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	53.8462	42.3077	46.1538	46.1538

Tabela 5.56: Taxa de Classificação (%) Edema x Nódulo - Estimador KSG/KOLE.

<b>Vozes com Edema x Vozes com Nódulo</b>				
<b>Medidas</b>	$\tau$	$PMIM$	$H$	$K_2$
Acurácia	61.5385	50.0000	50.0000	53.8462

Como pode ser observado, para a classificação utilizando o classificador SVM, os resultados obtidos utilizando-se todas as medidas, individualmente, estimados com o estimador KSG/KOLE apresentaram maiores taxas de acurácia média do classificador quando comparadas com os resultados obtidos com o estimador ingênuo.

### Características Avaliadas de Forma Combinada

- LDA

Na Tabela 5.57, estão apresentados os valores obtidos de acurácia, sensibilidade e especificidade, utilizando o classificador LDA, em sinais de vozes afetadas por edema de Reike nas pregas vocais e vozes afetadas por nódulo nas pregas vocais, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.58 utilizando o estimador KSG/KOLE.

Tabela 5.57: Taxas de Classificação (%) Edema x Nódulo. Medidas Combinadas - Estimador Ingênuo.

<b>Vozes com Edema x Vozes com Nódulo</b>		
<b>Medidas</b>	$\tau + PMIM$	$\tau + PMIM + H + K_2$
Acurácia	$54.76 \pm 9.20$	$52.86 \pm 5.43$
Sensibilidade	$59.50 \pm 11.02$	$42.50 \pm 5.43$
Especificidade	$45.00 \pm 13.85$	$45.00 \pm 15.73$

Tabela 5.58: Taxas de Classificação (%) Edema x Nódulo. Medidas Combinadas - Estimador KSG/KOLE.

<b>Vozes com Edema x Vozes com Nódulo</b>		
<b>Medidas</b>	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	$69.52 \pm 5.89$	$60.95 \pm 3.71$
Sensibilidade	$66.50 \pm 6.37$	$63.50 \pm 6.46$
Especificidade	$60.00 \pm 12.48$	$50.00 \pm 12.92$

Também para este caso de classificação, os resultados obtidos mostram que, a combinação das características melhora as taxas de classificação, e os resultados obtidos com o estimador KSG/KOLE foi superior ao estimador ingênuo.

- SVM

Na Tabela 5.59, estão apresentados os valores obtidos de acurácia ( $A_c$ ), utilizando o classificador SVM, em sinais de vozes afetadas por edema de Reike nas pregas vocais e vozes afetadas por nódulo nas pregas vocais, utilizando o estimador ingênuo para as seguintes medidas combinadas:  $\tau + PMIM$  e todas as medidas juntas ( $\tau + PMIM + H + K_2$ ) e na Tabela 5.60 utilizando o estimador KSG/KOLE.

Tabela 5.59: Taxa de Classificação (%) Edema x Nódulo. Medidas Combinadas - Estimador Ingênuo.

<b>Vozes com Edema x Vozes com Nódulo</b>		
<b>Medidas</b>	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	53.8461	53.8462

Tabela 5.60: Taxa de Classificação (%) Edema x Nódulo. Medidas Combinadas - Estimador KSG/KOLE.

<b>Vozes com Edema x Vozes com Nódulo</b>		
<b>Medidas</b>	$\tau+PMIM$	$\tau+PMIM+H+K_2$
Acurácia	80.7692	61.5384

Como pode ser observado, para este caso de classificação com o classificador SVM, a incorporação de mais duas medidas (caso onde todas as medidas foram agrupadas) não aumenta as taxas de acurácia do classificador. Com a combinação de apenas duas medidas,  $\tau$  e  $PMIM$ , foi conseguida uma taxa de mais de 80% de acurácia. Ainda assim, todos os resultados obtidos com o uso do estimador KSG/KOLE apresentaram maiores taxas de acurácia quando comparados com os resultados obtidos com o estimador ingênuo.

## Discussão dos Resultados

Nesta sessão foram apresentados os resultados da classificação entre vozes afetadas por diferentes patologias nas pregas vocais. Quando observamos as medidas utilizadas de maneira individual, podemos concluir que, quase todos os casos de classificação considerados

utilizando o estimador KSG/KOLE e o classificador LDA, apresentaram maiores valores de acurácia média quando comparados com o estimador ingênuo. Para o classificador SVM, o uso do estimador KSG/KOLE apresentou melhores taxas de acurácia para todos os casos de classificação considerados.

Quando observamos as medidas utilizadas de maneira combinada, as taxas de acurácia para ambos os classificadores, utilizando o estimador KSG/KOLE, foram bem mais elevadas quando comparadas ao estimador ingênuo, a exemplo da classificação EDM x NDL que, com o estimador KSG/KOLE e classificador SVM, apresentou taxa de acurácia de 80.7692%, contra 53.8461% com o estimador ingênuo como pode ser observado na Figura 5.11.

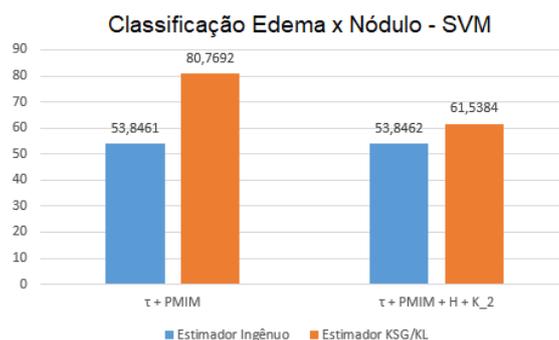


Figura 5.11: Classificação Edema x Nódulo - Medidas Combinadas

# Capítulo 6

## Conclusões e Trabalhos Futuros

Ao longo desta tese foram investigados métodos de estimação de medidas de informação de maneira a avaliar se o uso de um modelo diferente de estimação produzia melhores resultados na classificação entre sinais de vozes saudáveis e patológicas e entre sinais de vozes patológicas. A análise baseada na teoria dos sistemas dinâmicos não-lineares e na teoria do caos tem oferecido caminhos alternativos para a extração de informações do sistema de produção da fala a partir da saída desse sistema representado por meio de uma série temporal (sinal de voz).

Neste trabalho foram realizados experimentos computacionais (simulações) reunindo um conjunto de quatro medidas decorrentes da análise não-linear e da teoria do caos que, via de regra, são abordadas de forma fragmentada na literatura. O potencial discriminativo de cada uma dessas medidas é avaliado, primeiro, por meio de testes estatísticos, e posteriormente, como características (de forma individual e combinada) nos processos de classificação, comparando os resultados obtidos com dois métodos de estimação diferentes o KSG e o ingênuo, além de empregar o uso de dois classificadores distintos, o classificador baseado em validação cruzada e o SVM.

Neste capítulo, resumimos as principais conclusões desta tese, além de apontar direções para novas pesquisas.

### 6.1 Resultados dos testes Experimentais

Os resultados apresentados utilizando sinais gaussianos para avaliar o desempenho do estimador KSG em relação ao estimador ingênuo revelaram diferenças significativas nos valores de  $\tau$  estimados com o estimador ingênuo e com o estimador KSG para todas as classes consideradas. Em alguns casos os valores de  $\tau$  caem mais que a metade, como por exemplo: para a classe edema de Reinke, com o estimador ingênuo estimamos um valor de  $\tau = 16$  e com o estimador KSG, para o mesmo sinal de voz, estimamos um valor de  $\tau = 8$ , para a classe nódulo, com o estimador ingênuo estimamos um valor de  $\tau = 15$  e com o estimador KSG, para o mesmo sinal de voz, estimamos um valor de  $\tau = 7$ , para a classe paralisia, com o estimador ingênuo estimamos um valor de  $\tau = 20$  e com o estimador KSG, para o mesmo sinal de voz, estimamos um valor de  $\tau = 9$ .

Além disso o resultado dos testes estatísticos mostraram que, existem diferenças significativas entre os valores de  $\tau$  estimados com o estimador KSG e o valor de  $\tau$  estimado com o estimador ingênuo para todas as classes consideradas. Ao utilizar o método KSG para estimar  $\tau$ , observamos que as comparações de SDL x EDM, SDL x NDL, SDL x PRL, SDL x PTL e PRL x NDL apresentaram diferenças estatísticas significativas.

Observamos que os únicos casos em que não houve diferença significativa entre o KSG estimado  $\tau$  foram as comparações EDM x NDL e PRL x EDM.

Quando consideramos apenas os valores de entropia de Shannon estimados com o estimador KSG, a análise dos resultados apresentados indica diferenças estatísticas significativas dos valores da entropia de Shannon entre o grupo de vozes saudáveis e cada um dos três grupos de vozes patológicas. Entretanto, os testes revelaram não haver evidências estatísticas para o uso da entropia de Shannon no sentido de diferenciar entre os grupos de vozes patológicas.

A partir do resultado dos testes estatísticos, podemos concluir que, a entropia de Shannon pode ser usada, em potencial, como um discriminativo entre vozes saudáveis e as patologias da laringe consideradas neste trabalho, mas não para diferenciar entre os diferentes tipos de patologias.

## 6.2 Classificação de Sinais de Vozes Saudáveis *Ver-sus* Patológicas Usando Medidas da Análise Não-Linear

Os resultados apresentados para este grupo de classificação mostraram que, para quase todos os casos de classificação considerados, as medidas  $\tau$ , *PMIM* e  $K_2$  utilizando o estimador KSG, apresentaram maiores valores de acurácia média quando comparados com o estimador ingênuo para ambos os classificadores empregados.

Apenas para a medida  $H$  os resultados obtidos utilizando o estimador ingênuo e classificador baseado em validação cruzada, foi superior ao KSG, o que pode ser explicado uma vez que, devido a dimensionalidade dos dados, a estimação da entropia de Shannon com o estimador ingênuo converge muito bem para o valor analítico, como foi apresentado na seção 5.1.2 do capítulo que antecede a este.

Para o classificador SVM, apenas para o caso de classificação saudável  $\times$  paralisia, o estimador ingênuo apresentou melhor desempenho em termos de acurácia. Para todos os outros casos considerados o estimador KSG mostrou sua superioridade.

Podemos observar também que a combinação das medidas  $\tau$ , *PMIM*  $H$  e  $K_2$  utilizando o estimador KSG, para ambos os classificadores, apresentaram maiores taxas de classificação, quando comparadas ao estimador ingênuo, com taxas acima de 70%, o que é bastante interessante, pois o uso de apenas quatro medidas foram capazes de fazer uma boa classificação.

## 6.3 Classificação de Sinais de Vozes Patológicas Usando Medidas da Análise Não-Linear

Quando observamos as medidas utilizadas de maneira individual, podemos concluir que, para quase todos os casos de classificação considerados utilizando o estimador KSG e o classificador validação cruzada, apresentaram maiores valores de acurácia média quando comparados com o estimador ingênuo. Para o classificador SVM, o uso do estimador KSG apresentou melhores taxas de acurácia para todos os casos de classificação considerados.

Quando observamos as medidas utilizadas de maneira combinada, as taxas de acurácia para ambos os classificadores, utilizando o estimador KSG, foram bem mais elevadas

quando comparadas ao estimador ingênuo, a exemplo da classificação EDM x NDL que, com o estimador KSG e classificador SVM, apresentou taxa de acurácia de 80.7692%, contra 53.8461% com o estimador ingênuo, que é bastante interessante, pois o uso de apenas quatro medidas foram capazes de fazer uma boa classificação entre classes patológicas.

Por fim podemos concluir que a escolha de um estimador mais preciso torna o classificador mais eficiente, uma vez que a estimativa do KSG se aproxima de um valor que melhor se adapta às características de um verdadeiro atraso de reconstrução além de apresentar melhores resultados para as demais medidas consideradas. É um fato bem conhecido que um atraso de reconstrução muito curto não capturará a dinâmica dos dados, enquanto um atraso de reconstrução muito grande o tornará completamente independente em um sentido estatístico [85]. O atraso de reconstrução estimado com o método KSG captura a dinâmica dos sinais de vozes de forma mais confiável, o que, por sua vez, permite um melhor desempenho de classificação.

Ressaltamos que os resultados obtidos nesta tese são interessantes para o caso de detecção entre sinais de vozes saudáveis e patológicas, como também para o caso de classificação entre patologias. A classificação entre patologias ainda é um assunto pouco explorado na literatura.

## 6.4 Perspectiva de Trabalhos Futuros

Embora os resultados obtidos nesta pesquisa possam ser, de uma forma geral, considerados relevantes, o presente trabalho abre espaço para novas pesquisas.

- Extensão do trabalho realizado nesta tese para outras aplicações, a exemplo dos sinais de EEG;
- Avaliação do desempenho das medidas no processo de classificação, após a separação por gênero dos locutores e grau de severidade das patologias;
- Ampliação da análise dos espaços de fases reconstruídos a partir da representação gráfica dos gráficos de recorrência.

# Referências Bibliográficas

- [1] Mara Behlau. *Voz: O livro do Especialista*, volume 1 of *Optics and Photonics*. Revinter, Rio de Janeiro, Brasil., 2001.
- [2] Silvana L N C Costa. *Análise Acústica, Baseada no Modelo Linear de Produção da Fala, para Discriminação de Vozes Patológicas*. PhD thesis, Campina Grande, Brasil, 2008.
- [3] Zuleica Camargo. Avaliação objetiva da voz. *Carrara-Angelis E, Furia LB, Mourão LF, Kowalski LP. A atuação da Fonoaudiologia no câncer de cabeça e pescoço. São Paulo: Lovise*, pages 175–92, 2000.
- [4] Lilian Elisa Minikel Brod and Izabel Christine Seara. As vogais orais do português brasileiro na fala infantil e adulta: uma análise comparativa. *Revista Linguagem & Ensino*, 16(1):111–130, 2013.
- [5] Philippe H Dejonckere, Patrick Bradley, Pais Clemente, Guy Cornut, Lise Crevier-Buchman, Gerhard Friedrich, Paul Van De Heyning, Marc Remacle, and Virginie Woisard. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. *European Archives of Oto-rhino-laryngology*, 258(2):77–82, 2001.
- [6] Jagadish Nayak, P Subbanna Bhat, R Acharya, and UV Aithal. Classification and analysis of speech abnormalities. *ITBM-RBM*, 26(5):319–327, 2005.
- [7] Jenö Hirschberg, PH Dejonckere, Minoru Hirano, Kazunori Mori, H-J Schultz-Coulon, and Karel Vrtička. Voice disorders in children. *International journal of pediatric otorhinolaryngology*, 32:S109–S125, 1995.
- [8] Janina K Casper and Rebecca Leonard. *Understanding voice problems: A physiological perspective for diagnosis and treatment*. Lippincott Williams & Wilkins, 2006.
- [9] Iuberi Carson Zwetsch. Processamento do sinal de voz para o diagnóstico diferencial de alterações laríngeas. Master’s thesis, Pontifícia Universidade Católica do Rio Grande do Sul, 2006.
- [10] Mikaelle Oliveira Santos. Análise acústica de desvios vocais infantis utilizando a transformada wavelet. Master’s thesis, Instituto Federal de Educação Ciência e Tecnologia da Paraíba, João Pessoa, Brasil.

- [11] W. C. A Costa. *Análise dinâmica não linear de sinais de voz para detecção de patologias laríngeas*. PhD thesis, Universidade Federal de Campina Grande. Tese de Doutorado, 2012.
- [12] Patricia Henríquez, Jesús B Alonso, Miguel A Ferrer, Carlos M Travieso, Juan I Godino-Llorente, and Fernando Díaz-de María. Characterization of healthy and pathological voice through measures based on nonlinear dynamics. *IEEE Transactions on audio, speech, and language processing*, 17(6):1186–1195, 2009.
- [13] Steven B Davis. Acoustic characteristics of normal and pathological voices. *Speech and language: advances in basic research and practice*, 1(27):1–314, 1979.
- [14] Francis Quek, Mary Harper, Y Haciahtmetoglou, Lei Chen, and Lorraine O Ramig. Speech pauses and gestural holds in parkinson’s disease. In *Proc. 2002 Int. Conf. Spoken Language Processing*, pages 2485–2488, 2002.
- [15] Joseph C Stemple, Nelson Roy, and Bernice K Klaben. *Clinical voice pathology: Theory and management*. Plural Publishing, 2014.
- [16] Louise Cummings. *Clinical linguistics*. Edinburgh University Press, 2008.
- [17] CA Araújo. Cuidados com a voz. *AMB-Rev Assoc Med Bras*, 48:24, 2004.
- [18] Anadergh Barbosa-Branco and Marilene S Romariz. Doenças das cordas vocais e sua relação com o trabalho. *Comum Ciênc Saúde*, 17(1):37–45, 2006.
- [19] Adriane Parraga. Aplicação da transformada wavelet packet na análise e classificação de sinais de vozes patológicas. Master’s thesis, Universidade Federal do Rio Grande do Sul, 2002.
- [20] Juan Ignacio Godino-Llorente, Pedro Gomez-Vilda, and Manuel Blanco-Velasco. Dimensionality reduction of a pathological voice quality assessment system based on gaussian mixture models and short-term cepstral parameters. *Biomedical Engineering, IEEE Transactions on*, 53(10):1943–1953, 2006.
- [21] Yu Zhang and Jack J Jiang. Chaotic vibrations of a vocal fold model with a unilateral polyp. *The Journal of the Acoustical Society of America*, 115(3):1266–1269, 2004.
- [22] Jack J Jiang, Yu Zhang, and Jennifer Stern. Modeling of chaotic vibrations in symmetric vocal folds. *The Journal of the Acoustical Society of America*, 110(4):2120–2128, 2001.
- [23] Patrick Mergell, Hanspeter Herzel, and Ingo R Titze. Irregular vocal-fold vibration—high-speed observation and modeling. *The Journal of the Acoustical Society of America*, 108(6):2996–3002, 2000.
- [24] Hanspeter Herzel, David Berry, Ingo Titze, and Ina Steinecke. Nonlinear dynamics of the voice: signal analysis and biomechanical modeling. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1):30–34, 1995.

- [25] Jack J Jiang, Yu Zhang, and Charles N Ford. Nonlinear dynamics of phonations in excised larynx experiments. *The Journal of the Acoustical Society of America*, 114(4):2198–2205, 2003.
- [26] David A Berry, Hanspeter Herzel, Ingo R Titze, and Brad H Story. Bifurcations in excised larynx experiments. *Journal of Voice*, 10(2):129–138, 1996.
- [27] Pablo Henrique U Pinho, Mary Kallynna Couras, Evandson CS Dantas, Silvana Cunha Costa, and Suzete Élide Nóbrega Correia. Classificação de patologias laríngeas por meio de características do espaço de fase reconstruído. *XXXIV Simpósio Brasileiro de Telecomunicações– SBrT*, 2016.
- [28] Vinícius Jefferson Dias Vieira. Avaliação de distúrbios da voz por meio de análise de quantificação de recorrência. Master’s thesis, Instituto de Educação, Ciência e Tecnologia da Paraíba, 2014.
- [29] Yu Zhang and Jack J Jiang. Acoustic analyses of sustained and running voices from patients with laryngeal pathologies. *Journal of Voice*, 22(1):1–9, 2008.
- [30] Lotfi Salhi and Adnane Cherif. Selection of pertinent acoustic features for detection of pathological voices. In *Modeling, Simulation and Applied Optimization (ICMSAO), 2013 5th International Conference on*, pages 1–6. IEEE, 2013.
- [31] Washington César de Almeida Costa, Silvana Luciene do Nascimento Cunha Costa, Francisco Marcos de Assis, and Benedito Guimarães Aguiar Neto. Classificação de sinais de vozes saudáveis e patológicas por meio da combinação entre medidas da análise dinâmica não linear e codificação preditiva linear. *Revista Brasileira de Engenharia Biomédica*, 29(1), 2013.
- [32] Paulo Rogério Scalassara, Maria Eugenia Dajer, Carlos Dias Maciel, and Jose Carlos Pereira. Voice signals characterization through entropy measures. In *BIOSIGNALS (2)*, pages 163–170, 2008.
- [33] Jack J Jiang, Yu Zhang, and Clancy McGilligan. Chaos in voice, from modeling to measurement. *Journal of Voice*, 20(1):2–17, 2006.
- [34] Iasonas Kokkinos and Petros Maragos. Nonlinear speech analysis using models for chaotic systems. *IEEE Transactions on Speech and Audio Processing*, 13(6):1098–1109, 2005.
- [35] HH FALCÃO, S Correia, SC Costa, JM Fachine, and B Aguiar Neto. O uso da entropia na discriminação de vozes patológicas. In *Anais do 21o Congresso Brasileiro de Engenharia Biomédica. Salvador: Sociedade Brasileira de Engenharia Biomédica*, pages 1599–1602, 2008.
- [36] ME Torres, LG Gamero, HL Rufiner, C Martínez, DH Milone, and G Schlotthauer. Study of complexity in normal and pathological speech signals. In *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, volume 3, pages 2339–2342. IEEE, 2003.
- [37] Elon Lages Lima. *Variedades diferenciáveis*. Number 15. Instituto Matemática Puro e Aplicada, Conselho Nacional de Pesquisas, 1973.

- [38] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. Estimating mutual information. *Physical review E*, 69(6):066138, 2004.
- [39] John R Deller, John HL Hansen, and John G Proakis. *Discrete-time processing of speech signals*. Ieee New York, NY, USA:, 2000.
- [40] B P dos S Siqueira, P L Kalls, J A Faria, and D B Fernandes. Características dos sons das vogais do português falados no brasil. *Incitel*, 2013.
- [41] M. E. Dajer. Padrões visuais de sinais de voz através de técnica de análise de não-linear. Master’s thesis, São Paulo, Brasil., 2006.
- [42] D M Bless M Hirano. Exame videoestroboscópico da laringe. 1993.
- [43] Leonardo Wanderley Lopes, Ivonaldo Leidson Barbosa Lima, Larissa Nadjara Alves Almeida, Débora Pontes Cavalcante, and Anna Alice Figueirêdo de Almeida. Severity of voice disorders in children: correlations between perceptual and acoustic data. *Journal of Voice*, 26(6):819–e7, 2012.
- [44] Mara Behlau and Paulo Pontes. Avaliação global da voz. *São Paulo, EPPM*, 1990.
- [45] Lawrence R Rabiner and Ronald W Schafer. *Digital processing of speech signals*, volume 100. Prentice-hall Englewood Cliffs, 1978.
- [46] Radish Kumar Balasubramanium, Jayashree S Bhat, Sahikha Fahim, and Ria Raju. Cepstral analysis of voice in unilateral adductor vocal fold palsy. *Journal of Voice*, 25(3):326–329, 2011.
- [47] Xiang Wang, Jianping Zhang, and Yonghong Yan. Discrimination between pathological and normal voices using gmm-svm approach. *Journal of Voice*, 25(1):38–43, 2011.
- [48] B Radish Kumar, Jayashree S Bhat, and Neitica Prasad. Cepstral analysis of voice in persons with vocal nodules. *Journal of Voice*, 24(6):651–653, 2010.
- [49] Joseana Macêdo FECHINE. *Reconhecimento automático de identidade vocal utilizando modelagem híbrida: Paramétrica e Estatística. 2000. 212 f.* PhD thesis, Tese (Doutorado em Engenharia Elétrica)–Universidade Federal de Campina Grande, Campina Grande, 2000.
- [50] W Koenig, H K Dunn, and L Y Lacy. The sound spectrograph. *The Journal of the Acoustical Society of America*, 18(1):19–49, 1946.
- [51] Geraldo Pereira Jotz. *Configuração laríngea, análise perceptiva auditiva e computadorizada da voz de crianças institucionalizadas do sexo masculino*. PhD thesis, Universidade Federal de Sao Paulo. Escola Paulista de Medicina, 1997.
- [52] R.D. Kent and C. Read. The acoustic analysis of speech. 1992.
- [53] Gunnar Fant. *Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations*, volume 2. Walter de Gruyter, 1971.
- [54] HM Teager and SM Teager. Evidence for nonlinear sound production mechanisms in the vocal tract. *Speech production and speech modelling*, 55:241–261, 1990.

- [55] Petros Maragos, Alexander G Dimakis, and Iasonas Kokkinos. Some advances in nonlinear speech modeling using modulations, fractals, and chaos. In *Digital Signal Processing, 2002. DSP 2002. 2002 14th International Conference on*, volume 1, pages 325–332. IEEE, 2002.
- [56] Marcelo de Oliveira Rosa. *Laringe digital*. PhD thesis, Universidade de São Paulo, 2002.
- [57] Adriano Petry. Reconhecimento automático de locutor utilizando medidas de invariantes dinâmicas não-lineares. 2002.
- [58] Lingying Chai, Alicia J Sprecher, Yi Zhang, Yufang Liang, Huijun Chen, and Jack J Jiang. Perturbation and nonlinear dynamic analysis of adult male smokers. *Journal of voice*, 25(3):342–347, 2011.
- [59] Ghazaleh Vaziri, Farshad Almasganj, and Roozbeh Behroozmand. Pathological assessment of patients’ speech signals using nonlinear dynamical analysis. *Computers in biology and medicine*, 40(1):54–63, 2010.
- [60] Shaheen N Awan, Nelson Roy, and Jack J Jiang. Nonlinear dynamic analysis of disordered voice: The relationship between the correlation dimension (d 2) and pre-/post-treatment change in perceived dysphonia severity. *Journal of Voice*, 24(3):285–293, 2010.
- [61] B. Benjamin. Cirurgia endolaríngea. *E. Revinter, Rio de Janeiro*, 2000.
- [62] Dario Marcotullio, Giuseppe Magliulo, and Tiziana Pezone. Reinke’s edema and risk factors: clinical and histopathologic aspects. *American journal of otolaryngology*, 23(2):81–84, 2002.
- [63] Frederik G Dikkers and Peter GJ Nikkels. Benign lesions of the vocal folds: histopathology and phonotrauma. *Annals of Otology, Rhinology & Laryngology*, 104(9):698–703, 1995.
- [64] A Giovanni, N Henin, JM Triglia, M Cannoni, and A Pech. Formation nodulaires et paranodulaires. *Rev Laryngol*, 108:393–398, 1987.
- [65] A Hantzakos, Marc Remacle, FG Dikkers, J-C Degols, Monique Delos, G Friedrich, A Giovanni, and N Rasmussen. Exudative lesions of reinke’s space: a terminology proposal. *European Archives of Oto-rhino-laryngology*, 266(6):869, 2009.
- [66] Frederik G Dikkers and Peter GJ Nikkels. Lamina propria of the mucosa of benign lesions of the vocal folds. *The Laryngoscope*, 109(10):1684–1689, 1999.
- [67] O Kleinsasser. Microlaryngoscopic and histologic appearances of polyps, nodules, cysts, reinke’s edema, and granulomas of the vocal cords. In *Vocal fold histopathology: a symposium*, pages 51–55. College Hill Press, San Diego, California, 1986.
- [68] Náthalee Cavalcanti de Almeida. Sistema inteligente para diagnóstico de patologias na laringe utilizando máquinas de vetor de suporte. 2010.

- [69] New York Eye and Ear Infirmary of Mount Sinai (NYEE). Vocal fold paralysis. <http://www.nyee.edu/patient-care/otolaryngology/voice-swallowing/neurological-disorders/vocal-cord-paralysis>. Agosto, 2017.
- [70] Abdel Latif Hamdan, Roger V Moukarbel, Firas Farhat, and Mounir Obeid. Vocal cord paralysis after open-heart surgery. *European Journal of Cardio-thoracic surgery*, 21(4):671–674, 2002.
- [71] Nelson Fiedler-Ferrara and CP Cintra do Prado. *Caos: uma introdução*. Edgar Blucher, 1995.
- [72] MA Savi. Chaos and order in biomedical rhythms. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 27(2):157–169, 2005.
- [73] Holger Kantz and Thomas Schreiber. *Nonlinear time series analysis*, volume 7. Cambridge university press, 2004.
- [74] Arun Kumar and SK Mullick. Nonlinear dynamical analysis of speech. *The Journal of the Acoustical Society of America*, 100(1):615–629, 1996.
- [75] Gernot Kubin. Synthesis and coding of continuous speech with the nonlinear oscillator model. In *icassp*, pages 267–270. IEEE, 1996.
- [76] Marcelo Amorin Savi. *Dinâmica não-linear e caos*. Editora E-papers, 2006.
- [77] Floris Takens. *Detecting strange attractors in turbulence*. Springer, 1981.
- [78] JJ Jiang and Yu Zhang. Nonlinear dynamic analysis of speech from pathological subjects. *Electronics Letters*, 38(6):294–295, 2002.
- [79] Michael T Johnson, Richard J Povinelli, Andrew C Lindgren, Jinjin Ye, Xiaolin Liu, and Kevin M Indrebo. Time-domain isolated phoneme classification using reconstructed phase spaces. *IEEE Transactions on Speech and Audio Processing*, 13(4):458–466, 2005.
- [80] Julián D Arias-Londoño, Juan I Godino-Llorente, Nicolás Sáenz-Lechón, Víctor Osma-Ruiz, and Germán Castellanos-Domínguez. Automatic detection of pathological voices using complexity measures, noise parameters, and mel-cepstral coefficients. *IEEE Transactions on Biomedical Engineering*, 58(2):370–379, 2011.
- [81] Ingo R Titze. *Workshop on acoustic voice analysis: Summary statement*. National Center for Voice and Speech, 1995.
- [82] Everton Granemann Souza. Caracterização de sistemas dinâmicos através de gráficos de recorrência. *Universidade Federal do Paraná. Dissertação de Mestrado*, 2008.
- [83] Luiz Henrique Alves Monteiro. *Sistemas dinâmicos*. Editora Livraria da Física, 2006.
- [84] Edward Ott. *Chaos in dynamical systems*. Cambridge university press, 2002.
- [85] Henry DI Abarbanel. *Analysis of observed chaotic data*. Springer, 1996.

- [86] Henri Poincaré. *Les méthodes nouvelles de la mécanique céleste: Méthodes de MM. Newcomb, Glydén, Lindstedt et Bohlin. 1893*, volume 2. Gauthier-Villars it fils, 1893.
- [87] Edward N Lorenz. Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130–141, 1963.
- [88] Andrew M Fraser and Harry L Swinney. Independent coordinates for strange attractors from mutual information. *Physical review A*, 33(2):1134, 1986.
- [89] Wentian Li. Mutual information functions versus correlation functions. *Journal of statistical physics*, 60(5-6):823–837, 1990.
- [90] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.
- [91] Mikaelle Oliveira Santos, Juliana Martins de Assis, Vinícius Jefferson Dias Vieira, and Francisco Marcos de Assis. Ksg estimation of reconstruction delay to detect vocal disorders in nonlinear dynamical analysis. *Research on Biomedical Engineering*, 34(3):217–225, 2018.
- [92] Juliana M de Assis, Mikaelle O Santos, and Francisco M de Assis. Auditory stimuli coding by postsynaptic potential and local field potential features. *PloS one*, 11(8):e0160089, 2016.
- [93] Georges A Darbellay and Igor Vajda. Estimation of the information by an adaptive partitioning of the observation space. *IEEE Transactions on Information Theory*, 45(4):1315–1321, 1999.
- [94] LF Kozachenko and Nikolai N Leonenko. Sample estimate of the entropy of a random vector. *Problemy Peredachi Informatsii*, 23(2):9–16, 1987.
- [95] Milton Abramowitz and Irene A Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation, 1964.
- [96] Claude Elwood Shannon. A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423, 1948.
- [97] Max A Little, Patrick E McSharry, Stephen J Roberts, Declan AE Costello, and Irene M Moroz. Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *Biomedical engineering online*, 6(1):23, 2007.
- [98] LF Kozachenko and Nikolai N Leonenko. Sample estimate of the entropy of a random vector. *Problemy Peredachi Informatsii*, 23(2):9–16, 1987.
- [99] Bernhard Riemann. On the number of prime numbers less than a given quantity (ueber die anzahl der primzahlen unter einer gegebenen grösse.). *Monatsberichte der Berliner Akademie*, 1859.
- [100] Guido Nolte, Andreas Ziehe, and Klaus-Robert Müller. Noise robust estimates of correlation dimension and k 2 entropy. *Physical review E*, 64(1):016112, 2001.

- [101] Peter Grassberger and Itamar Procaccia. Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, 9(1-2):189–208, 1983.
- [102] Diks Cees. *Nonlinear Time Series Analysis: Methods And Applications*, volume 4. World Scientific, 1999.
- [103] Kay Elemetrics. Kay elemetrics corp. disordered voice database, 1994.
- [104] Rainer Hegger, Holger Kantz, and Thomas Schreiber. Practical implementation of nonlinear time series methods: The tisean package. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 9(2):413–435, 1999.
- [105] Norbert Marwan and Jürgen Kurths. Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters A*, 302(5):299–307, 2002.
- [106] Gunnar Fant. *Speech acoustics and phonetics: Selected writings*, volume 24. Springer Science & Business Media, 2006.
- [107] Ian H Witten and Eibe Frank. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005.
- [108] Ron Kohavi et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145, 1995.
- [109] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- [110] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. A practical guide to support vector classification. 2003.
- [111] Christopher JC Burges. Geometry and invariance in kernel based methods. 1999.
- [112] Nello Cristianini, John Shawe-Taylor, et al. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.
- [113] Joseph T Lizier. Jidt: An information-theoretic toolkit for studying the dynamics of complex systems. *arXiv preprint arXiv:1408.3270*, 2014.
- [114] Athanasios Papoulis and S Unnikrishna Pillai. *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.
- [115] Paul L Meyer. Probabilidade: aplicações à estatística. In *Probabilidade: aplicações à estatística*. Livro Técnico, 1983.
- [116] Lorí Viali. Série estatística básica - estimação., 2019.
- [117] Ivo Grosse. *Applications of statistical physics and information theory to the analysis of DNA sequences*. PhD thesis, Boston University, 2000.
- [118] Sidney Siegel and N John Castellan Jr. *Estatística não-paramétrica para ciências do comportamento*. Artmed Editora, 2006.
- [119] Harold J Larson and Harold J Larson. *Introduction to probability theory and statistical inference*, volume 12. Wiley New York, 1982.

# Apêndice A

## Conceitos de Probabilidade e Estatística

Este apêndice tem por finalidade apresentar alguns conceitos de probabilidade e estatística que podem ser úteis no entendimento deste trabalho. Primeiramente, é interessante afirmar que enquanto a estatística lida com aplicações da teoria a problemas reais e suas conclusões são inferências baseadas em observações, a probabilidade é uma disciplina matemática desenvolvida como um modelo abstrato e suas deduções são baseadas em axiomas [114].

Em diversas aplicações, é mais viável retirar conclusões do todo (população inteira) a partir de uma amostra aleatória. A inferência estatística consiste em retirar conclusões acerca da natureza de algum sistema baseado em dados sujeitos a variação aleatória [115].

Um importante conceito de probabilidade é o de variável aleatória. Variável aleatória é uma função que associa a cada elemento do espaço amostral um número real. Uma amostra aleatória de uma variável aleatória  $X$  é um vetor de variáveis aleatórias  $(X_1, X_2, \dots, X_N)$ , cada uma com a mesma distribuição de  $X$  meyer1983. Na investigação estatística, há duas classes gerais de problemas. Na primeira classe, há um modelo probabilístico conhecido e deseja-se fazer previsões de futuras observações - portanto, parte-se do modelo para as observações.

Já na segunda classe, um ou mais parâmetros do modelo são desconhecidos e deseja-se estimá-los (estimação de parâmetro) ou decidir se estão em um determinado subconjunto de valores (teste de hipótese). Nesse caso, parte-se das observações para o modelo [114], classe esta considerada ao longo da presente tese.

O problema de estimação está bem definido para o caso paramétrico, ou seja, aquele em que se conhece a distribuição de probabilidade, porém, um ou mais dos seus parâmetros  $\theta$  são desconhecidos fazendo-se necessária sua estimação. Neste caso, o estimador é uma função de uma amostra aleatória, i.e.,  $\hat{\theta} = g(X_1, X_2, \dots, X_N)$ , e dessa forma constitui outra variável aleatória. Seu valor específico para uma determinada realização amostral é chamada de estimativa  $\hat{\theta} = g(x_1, x_2, \dots, x_N)$ . Qualquer função de uma amostra aleatória é chamada de estatística, logo, um estimador é uma estatística [114].

Um estimador  $\hat{\theta}$  é dito não enviesado do parâmetro  $\theta$  se  $\mathbb{E}\hat{\theta} = \theta$ . Caso contrário, é dito enviesado com viés  $b = \mathbb{E}\hat{\theta} - \theta$ . A média amostral  $\bar{X}$  de uma variável aleatória  $X$  é um exemplo de estimador não enviesado da esperança de  $X$ ,  $\mathbb{E}X$ :

$$\begin{aligned}
\mathbb{E}\bar{X} &= \mathbb{E} \frac{1}{N} \sum_{n=1}^N X_n \\
&= \frac{1}{N} \sum_{n=1}^N \mathbb{E}X_n \\
&= \frac{1}{N} \sum_{n=1}^N \mathbb{E}X \\
&= \mathbb{E}X.
\end{aligned}$$

Caso a função  $g(X_1^N)$  seja selecionada adequadamente, o erro de estimação  $\hat{\theta} - \theta$  diminui à medida que  $N$  aumenta [114]. Existem vários critérios de convergência para uma sequência de variável aleatória, por exemplo:

- $\hat{\theta}_n$  converge em probabilidade se  $\lim_{n \rightarrow \infty} P[|\hat{\theta}_n - \theta| > \epsilon] = 0$  [64];
- $\hat{\theta}_n$  converge com probabilidade 1 (ou quase certamente) para  $\theta$  se  $P[\lim_{n \rightarrow \infty} \hat{\theta}_n = \theta] = 1$  [13];

Um estimador é dito consistente se seu erro de estimação converge para zero em probabilidade. Além disso, diz-se que um estimador  $\hat{\theta}_1$  de um parâmetro  $\theta$  é mais acurado que um estimador  $\hat{\theta}_2$  se [116]

$$\mathbb{E}(\hat{\theta}_1 - \theta)^2 < \mathbb{E}(\hat{\theta}_2 - \theta)^2$$

Quanto menor o erro médio quadrático de um estimador em relação ao parâmetro que se deseja estimar, maior sua acurácia. Seja  $b$  o viés do estimador  $\hat{\theta}$  do parâmetro  $\theta$  de uma distribuição. É possível desenvolver a seguinte relação entre acurácia ( $\mathbb{E}(\hat{\theta} - \theta)^2$ ), viés e variância de um estimador:

$$\begin{aligned}
b &= \mathbb{E}(\hat{\theta} - \theta), \text{ viés do estimador,} \\
\mathbb{E}(\hat{\theta} - \theta)^2 &= \mathbb{E}(\hat{\theta}^2 - 2\theta\hat{\theta} + \theta^2) \\
&= \mathbb{E}\hat{\theta}^2 - 2\theta\mathbb{E}\hat{\theta} + \theta^2 \\
&= \mathbb{E}\hat{\theta}^2 - (\mathbb{E}\hat{\theta})^2 + (\mathbb{E}\hat{\theta})^2 - 2\theta\mathbb{E}\hat{\theta} + \theta^2 \\
&= \text{var}(\hat{\theta}) + (b + \theta)^2 - 2\theta(b + \theta) + \theta^2 \\
&= \text{var}(\hat{\theta}) + b^2.
\end{aligned}$$

Dessa forma, quanto menores a variância e o viés do estimador  $\hat{\theta}$ , maior será sua acurácia. É esperado que o valor verdadeiro de um parâmetro da população esteja relacionado à amostra aleatória encontrada na realização de experimentos. Ou seja, é esperado que a função densidade de probabilidade  $f(x_1, x_2, \dots, x_N; \theta)$  dependa do valor de  $\theta$ . Por exemplo, uma amostra aleatória de uma variável aleatória gaussiana apresenta valores que dependem de sua média. Dessa forma, em geral procura-se achar a melhor função

dos dados para estimar um parâmetro, o que pode ser feito no sentido da máxima verossimilhança. A estimativa de máxima verossimilhança (ML - *maximum likelihood*) do parâmetro  $\theta$  associada à distribuição de probabilidade conjunta de  $f(x_1, x_2, \dots, x_n; 0)$  de uma variável aleatória pode ser determinada a partir da equação de verossimilhança como:

$$\hat{\theta}_{ML} = \sup f(x_1, x_2, \dots, x_n; 0). \quad (\text{A.1})$$

ou através da função de log-verossimilhança

$$L(x_1, x_2, \dots, x_n; 0) = \ln f(x_1, x_2, \dots, x_n; 0). \quad (\text{A.2})$$

Se  $L(x_1, x_2, \dots, x_n; 0)$  for diferenciável e um supremo  $\hat{\theta}_{ML}$  existir, então ele deve satisfazer à equação [114]:

$$\left. \frac{\partial \ln f(x_1, x_2, \dots, x_n; 0)}{\partial \theta} \right|_{\theta=\hat{\theta}_{ML}} \quad (\text{A.3})$$

Podemos observar que a conexão entre a probabilidade é baseada na frequência relativa de eventos, isto é, a probabilidade de um evento  $A$  costuma ser estimada pela contagem de sua ocorrência dividida pelo tamanho amostral [114]:

$$p = \frac{N_A}{N} \quad (\text{A.4})$$

Já é estabelecido que a frequência relativa é o estimador de máxima verossimilhança para um vetor de probabilidades de eventos discretos [117]. Apesar disso, o uso deste estimador para probabilidades que serão usadas no funcional de informação mútua, por exemplo, resulta em estimativas de informação mútua enviesadas. Diferentemente da estimação de entropia, em que o uso da frequência relativa para estimá-la resulta na média em estimativas menores que a verdadeira, não há uma direção para o viés da informação mútua quando a frequência relativa é utilizada [117]. O viés da estimação de informação mútua quando a frequência relativa é utilizada pode depender das distribuições conjuntas de probabilidade  $p(X, Y)$  e o tamanho amostral  $N$ .

# Apêndice B

## Testes Estatísticos

Este apêndice tem por finalidade trazer um breve esclarecimento acerca do uso dos testes estatísticos nesta tese.

Além de inferir parâmetros, muitas vezes em estatística deseja-se retirar alguma conclusão acerca dos dados observados - verificar se há de fato algum padrão neles ou se o que foi obtido deve-se somente ao acaso. O uso de testes estatísticos em uma pesquisa tem por objetivo fornecer um caráter científico às conclusões auferidas a partir dos resultados obtidos nos experimentos realizados. Para chegar a uma decisão objetiva sobre se uma determinada hipótese é confirmada, ou não, por um conjunto de dados, é preciso ter um procedimento objetivo para aceitar ou rejeitar uma hipótese estabelecida [118].

De uma forma geral, os testes de hipótese são classificados em dois tipos: paramétricos e não paramétricos.

- Testes paramétricos:
  - teste t pareado (compara a diferença entre as médias de variáveis dependentes);
  - teste t não pareado (compara a diferença entre as médias de variáveis independentes);
- Testes não paramétricos:
  - *Wilcoxon rank-sum* (leva em conta magnitude e direção da diferença entre duas variáveis independentes);
  - *Wilcoxon sign-rank* (leva em conta magnitude e direção da diferença entre duas variáveis dependentes);

Os testes paramétricos utilizam o pressuposto que os dados disponíveis formam uma amostra aleatória de uma (ou de mais de uma) população cuja distribuição é a normal. Por outro lado, os testes não paramétricos não necessitam de suposições sobre o tipo de distribuição da população a qual pertencem os dados da amostra utilizada [11].

Muitos testes paramétricos são robustos, apontando corretamente a evidência de uma hipótese ser verdadeira ou falsa ainda que a distribuição subjacente não seja exatamente aquela considerada para o modelo (normal, exponencial, binomial, etc.). Ao se testar uma mesma hipótese sob ambos tipos de teste - paramétrico e não paramétrico - há um resultado mais confiável no primeiro caso se a distribuição suposta pelo teste corresponder à verdadeira. Contudo, quando não se pode afirmar com certeza a lei de probabilidade envolvida, o teste não paramétrico é preferível [119].

Para compreender como os testes de hipóteses paramétricos funcionam, considera-se a questão a seguir. Dada uma variável  $X$ , que possua um modelo conhecido de distribuição, com parâmetro desconhecido  $\theta$ , testa-se a suposição de que  $\theta \neq \theta_0$  (hipótese nula,  $H_0$ ) contra a suposição de que  $\theta = \theta_0$  (hipótese alternativa,  $H_1$ ). A hipótese alternativa também pode ser unilateral, isto é, testa-se se  $\theta > \theta_0$  ou se  $\theta < \theta_0$ . As hipóteses são testadas baseadas em evidências experimentais, isto é, na amostra aleatória  $X_1^N$  observada. O propósito do teste não é determinar a veracidade de  $H_0$  ou  $H_1$ , mas sim estabelecer se há evidências que apoiam a rejeição de  $H_0$  [114].

Existem dois tipos de erros que podem ocorrer em testes de hipótese e devem ser considerados:

- Erro do tipo I: rejeitar  $H_0$  muito embora  $H_0$  seja verdadeiro. A probabilidade de tal erro é denotada por  $\alpha$  e é chamada de nível de significância do teste;
- Erro do tipo II: aceitar  $H_0$  muito embora  $H_0$  seja falso. A probabilidade de tal erro é denotada por  $\beta$  e a diferença  $1 - \beta$  é chamada de potência do teste.

# Apêndice C

## Lista de Publicações do Doutorado

- Mikaelle O. Santos, Juliana M. de Assis, Vinícius J. D. Vieira, Francisco M. de Assis. “KSG estimation of reconstruction delay to detect vocal disorders in nonlinear dynamical analysis”. Res. Biomed. Eng. vol.34 no.3 Rio de Janeiro July./Sept. 2018 Epub Aug 23, 2018;
- Juliana M. de Assis, Mikaelle O. Santos, Francisco M. de Assis. “Auditory Stimuli Coding by Postsynaptic Potential and Local Field Potential Features”. Plos One, v. 11, p. e0160089, 2016.