

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO
PROGRAMA DE FÍSICA APLICADA À MEDICINA E BIOLOGIA

**Inferência Estatística em Métodos de Análise de
Ressonância Magnética Funcional**

Brenno Caetano Troca Cabella

Ribeirão Preto
2008

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO
PROGRAMA DE FÍSICA APLICADA À MEDICINA E BIOLOGIA

Inferência Estatística em Métodos de Análise de Ressonância Magnética Funcional

Brenno Caetano Troca Cabella

Dissertação submetida ao Programa de Pós-Graduação em Física Aplicada à Medicina e Biologia da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo como parte dos requisitos para a obtenção do título de Mestre em Física Aplicada à Medicina e Biologia.

Orientador: Prof. Dr. Ubiraci Pereira da Costa Neves

Ribeirão Preto, Março de 2008.

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE DOCUMENTO, POR MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Cabella, B. C. T.

Inferência Estatística em Métodos de Análise de Ressonância Magnética Funcional / Brenno Caetano Troca Cabella; orientador Prof. Dr. Uiraci Pereira da Costa Neves.

– Ribeirão Preto/SP, 2008.

95 p.

Dissertação (Mestrado – Programa de Pós-Graduação em Física Aplicada à Medicina e Biologia) – Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo.

ressonância magnética funcional, inferência estatística, simulações numéricas, testes de hipóteses, distância de Kullback-Leibler, métodos bayesianos.

Banca Examinadora:

Resumo

CABELLA, B. C. T. **Inferência Estatística em Métodos de Análise de Ressonância Magnética Funcional**. Dissertação (Mestrado) – Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2008.

No presente trabalho, conceitos de inferência estatística são utilizados para aplicação e comparação de diferentes métodos de análise de sinais de ressonância magnética funcional. A idéia central baseia-se na obtenção da distribuição de probabilidade da variável aleatória de interesse, para cada método estudado e sob diferentes valores da relação sinal-ruído (SNR). Este objetivo é atingido através de simulações numéricas da função resposta hemodinâmica (HRF) acrescida de ruído gaussiano. Tal procedimento nos permite avaliar a sensibilidade e a especificidade dos métodos empregados através da construção das curvas ROC (*receiver operating characteristic*) para diferentes valores de SNR.

Sob específicas condições experimentais, aplicamos métodos clássicos de análise (teste t de Student e correlação), medidas de informação (distância de Kullback-Leibler e sua forma generalizada) e um método Bayesiano (método do pixel independente). Em especial, mostramos que a distância de Kullback-Leibler (D) (ou entropia relativa) e sua forma generalizada são medidas úteis para análise de sinais dentro do cenário de teoria da informação. Estas entropias são usadas como medidas da "distância" entre as funções de probabilidade p_1 e p_2 dos níveis do sinal relacionados a estímulo e repouso. Para prevenir a ocorrência de valores divergentes de D , introduzimos um pequeno parâmetro δ nas definições de p_1 e p_2 . Estendemos a análise, apresentando um estudo original da distância de Kullback-Leibler generalizada D_q (q é o parâmetro de Tsallis). Neste caso, a escolha apropriada do intervalo $0 < q < 1$ permite assegurar que D_q seja finito. Obtemos as densidades de probabilidade $f(\bar{D})$ e $f(\bar{D}_q)$ das médias amostrais das variáveis D e D_q , respectivamente, calculadas ao longo das N épocas de todo o experimento. Para pequenos valores de N ($N < 30$), mostramos que $f(\bar{D})$ e $f(\bar{D}_q)$ são muito bem aproximadas por distribuições Gamma ($\chi^2 < 0,0009$). Em seguida, estudamos o método (Bayesiano) do pixel independente, considerando a probabilidade *a posteriori* como variável aleatória e obtendo sua distribuição para várias SNR's e probabilidades *a priori*. Os resultados das simulações apontam para o fato de que a correlação e o método do pixel independente apresentam melhor desempenho do que os demais métodos empregados (para SNR > -20 dB). Contudo, deve-se ponderar que o teste t e os métodos entrópicos compartilham da vantagem de não se utilizarem de um modelo para HRF na análise de dados reais. Finalmente, para os diferentes métodos, obtemos os mapas funcionais correspondentes a séries de dados reais de um voluntário assintomático submetido a estímulo motor de evento relacionado, os quais demonstram ativação nas áreas cerebrais motoras primária e secundária. Enfatizamos que o procedimento adotado no presente estudo pode, em princípio, ser utilizado em outros métodos e sob diferentes condições experimentais.

Palavras-chave: ressonância magnética funcional, inferência estatística, simulações numéricas, testes de hipóteses, distância de Kullback-Leibler, métodos bayesianos.

Abstract

CABELLA, B. C. T.. **Statistical Inference in Methods of Analysis of Functional Magnetic Resonance**. Dissertation (Master) – Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2008.

In the present work, concepts of statistical inference are used for application and comparison of different methods of signal analysis in functional magnetic resonance imaging. The central idea is based on obtaining the probability distribution of the random variable of interest, for each method studied under different values of signal-to-noise ratio (SNR). This purpose is achieved by means of numerical simulations of the hemodynamic response function (HRF) with gaussian noise. This procedure allows us to assess the sensitivity and specificity of the methods employed by the construction of the ROC curves (*receiver operating characteristic*) for different values of SNR.

Under specific experimental conditions, we apply classical methods of analysis (Student's t test and correlation), information measures (distance of Kullback-Leibler and its generalized form) and a Bayesian method (independent pixel method). In particular, we show that the distance of Kullback-Leibler D (or relative entropy) and its generalized form are useful measures for analysis of signals within the information theory scenario. These entropies are used as measures of the "distance" between the probability functions p_1 and p_2 of the signal levels related to stimulus and non-stimulus. In order to avoid undesirable divergences of D , we introduced a small parameter δ in the definitions of p_1 and p_2 . We extend such analysis, by presenting an original study of the generalized Kullback-Leibler distance D_q (q is Tsallis parameter). In this case, the appropriate choice of range $0 < q < 1$ ensures that D_q is finite. We obtain the probability densities $f(\bar{D})$ and $f(\bar{D}_q)$ of the sample averages of the variables D and D_q , respectively, calculated over the N epochs of the entire experiment. For small values of N ($N < 30$), we show that $f(\bar{D})$ and $f(\bar{D}_q)$ are well approximated by Gamma distributions ($\chi^2 < 0.0009$). Afterward, we studied the independent pixel bayesian method, considering the probability *a posteriori* as a random variable, and obtaining its distribution for various SNR's and probabilities *a priori*. The results of simulations point to the fact that the correlation and the independent pixel method have better performance than the other methods used (for SNR > -20 dB). However, one should consider that the Student's t test and the entropic methods share the advantage of not using a model for HRF in real data analysis. Finally, we obtain the maps corresponding to real data series from an asymptomatic volunteer submitted to an event-related motor stimulus, which shows brain activation in the primary and secondary motor brain areas. We emphasize that the procedure adopted in this study may, in principle, be used in other methods and under different experimental conditions.

Keywords: functional magnetic resonance imaging, statistical inference, numerical simulations, hypothesis testing, Kullback-Leibler distance, bayesian methods.

DEDICATÓRIA

À minha família.

Agradecimentos

Ao meu orientador, Prof. Dr. Ubiraci Pereira da Costa Neves, pela orientação e amizade. Sua seriedade, ética e conhecimento foram essenciais para o desenvolvimento deste trabalho.

Ao Prof. Dr. Dráulio Barros de Araújo pelas discussões teóricas pertinentes e por incentivar este estudo.

Ao Prof. Dr. Alexandre Souto Martinez pela contribuição na minha formação acadêmica.

Ao amigo Márcio Junior Sturzbecher por mostrar o caminho das pedras no desenvolvimento deste trabalho.

Aos amigos Aquino de Lauri Espíndola , César Augusto Sangalleti Terçariol, Felipe de Moura Kiipper, Juliana Berbert, Leandro Rizzi, Marcelo Alves Pereira, Rodrigo da Silva Gonzalez, Tiago José Arruda e Wilnice Tavares Reis Oliveira que ajudaram a tornar o dia-a-dia de trabalho mais produtivo e agradável.

Aos funcionários Fábio José Moretti e Adriano de Jesus Holanda por garantirem o pleno funcionamento dos equipamentos necessários para o desenvolvimento deste trabalho e agradeço também às funcionárias Nilza Marina Leone Marino, Rita de Cassia Ribeiro e Sonia Aparecida Nali de Paula pela presteza.

À Fabíola Lima Pereira pelo carinho e apoio nos momentos difíceis.

Ao Departamento de Física e Matemática pelo suporte ao trabalho.

Ao CNPq pelo apoio financeiro.

Sumário

Resumo	p. v
Abstract	p. vi
Lista de Siglas	p. iv
Lista de Figuras	p. v
Lista de Tabelas	p. xii
1 Introdução	p. 1
2 Imagem Funcional por Ressonância Magnética	p. 5
2.1 O contraste BOLD	p. 6
2.2 Tipos de paradigma	p. 8
2.3 Modelo de função resposta hemodinâmica (HRF)	p. 9
2.4 Aquisição das imagens	p. 10
2.5 Pré-processamento	p. 12
2.5.1 Correção temporal das fatias	p. 12
2.5.2 Movimento da cabeça	p. 13
2.5.3 Correção da distorção	p. 14
2.5.4 Filtros	p. 15
2.6 Processamento	p. 15
3 Simulações das Séries Temporais, Testes de Hipóteses e Mapas Estatísticos	p. 16

3.1	Simulações das séries temporais	p. 16
3.1.1	Ruído	p. 16
3.1.2	SNR	p. 19
3.2	Testes de hipóteses	p. 21
3.3	Curvas ROC	p. 24
3.4	Protocolo de aquisição de dados reais	p. 25
3.5	Apresentação dos mapas estatísticos	p. 26
3.5.1	Correção de Šidák - Bonferonni	p. 26
3.6	Equipamentos e aplicativos utilizados	p. 27
3.6.1	Aquisição das imagens	p. 27
3.6.2	Pré-processamento	p. 27
3.6.3	Processamento	p. 28
4	Métodos Clássicos em fMRI	p. 29
4.1	O Teste t	p. 29
4.1.1	Teste t - Simulações	p. 32
4.1.2	Teste t - Dados reais	p. 33
4.2	Correlação	p. 35
4.2.1	Correlação - Simulações	p. 37
4.2.2	Correlação - Dados reais	p. 39
5	Distâncias de Kullback-Leibler em fMRI	p. 42
5.1	Entropia de Shannon	p. 42
5.2	Entropia de Tsallis	p. 44
5.3	Distância de Kullback-Leibler	p. 45
5.3.1	Distância de Kullback-Leibler - Simulações	p. 47
5.3.1.1	Otimização do número de níveis (L) e do parâmetro δ	p. 50

<i>Sumário</i>	iii
5.3.2 Distância de Kullback-Leibler - Dados reais	p. 53
5.4 Distância de Kullback-Leibler generalizada	p. 56
5.4.1 Distância de Kullback-Leibler Generalizada - Simulações	p. 58
5.4.1.1 Otimização do número de níveis (L) e do parâmetro q	p. 61
5.4.2 Distância de Kullback-Leibler generalizada - Dados reais	p. 64
6 Estatística Bayesiana em fMRI	p. 66
6.1 O método do pixel independente	p. 66
6.1.1 Estimativa de parâmetros	p. 69
6.2 Estatística Bayesiana - Simulações	p. 70
6.3 Estatística Bayesiana - Dados reais	p. 75
7 Comparação entre Métodos	p. 77
8 Conclusões e Perspectivas	p. 79
Apêndice A – Imagem por Ressonância Magnética Nuclear	p. 81
A.1 Princípios físicos	p. 81
A.2 Formação da imagem	p. 86
A.3 Intensidade do sinal	p. 87
A.4 Agentes de contraste	p. 90
A.5 Gradientes de campo magnético	p. 91
Referências	p. 93

Lista de Siglas

BOLD	Blood-Oxygenation-Level Dependent
dHb	deoxi-hemoglobina
EPI	Echo Planar Imaging
FID	Free induction decay
fMRI	functional Magnetic Resonance Imaging
FOV	Field Of View
FVN	Fração de Verdadeiros Negativos
FVP	Fração de Verdadeiros Positivos
GLM	General Linear Model
Hb	Hemoglobina
HRF	Hemodynamic Response Function
ICA	Independent Component Analysis
IRM	Imagem por Ressonância Magnética
pdf	probability density functions
PET	Positron Emission Tomography
RMN	Ressonância Magnética Nuclear
ROC	Receiver Operating Characteristic
SNR	Signal-to-Noise Ratio
SPECT	Single Photon Emission Computed Tomography
TE	Time to Echo
TR	Time to Repeat

Lista de Figuras

- 2.1 (a) Representação esquemática de um neurônio; (b) Processo de transmissão de impulso na sinapse. p. 7
- 2.2 Representação esquemática da resposta hemodinâmica. (a) Resposta a um estímulo de curta duração; (b) Resposta a estímulos múltiplos consecutivos. p. 8
- 2.3 Representação dos tipos de paradigma. (a) Bloco; (b) Evento relacionado e semi-aleatório; (c) Misto. p. 9
- 2.4 Modelo de resposta hemodinâmica para um paradigma de evento relacionado, utilizando o modelo da eq. (2.2) com os parâmetros: $a = 6$, $a^0 = 12$, $b = b^0 = 0.9$, $c = 0.35$, $d = ab$ e $d^0 = a^0 b^0$ p. 10
- 2.5 Tipos de cortes anatômicos que podem ser visualizados em imagens de ressonância magnética funcional. p. 11
- 2.6 Sequência de imagens obtidas ao longo do tempo para duas fatias axiais. Ao lado, série temporal correspondente a um *voxel* da fatia 2. p. 11
- 2.7 Ilustração mostrando a interpolação linear para correção temporal das fatias. p. 13
- 2.8 Ilustração mostrando a transformação de corpo rígido para alinhamento das imagens. p. 14
- 3.1 Série temporal de um *voxel*. Variações são devidas principalmente ao ruído térmico. p. 18
- 3.2 Ruído gaussiano simulado para diferentes valores de média e desvio-padrão. p. 18
- 3.3 Distribuição do ruído simulado com diferentes médias e desvios-padrão. Em vermelho, a *pdf* Gaussiana. p. 19
- 3.4 Séries temporais simuladas para diferentes valores de SNR. Para valores altos de SNR, a HRF é mais evidente. p. 20
- 3.5 Densidade de probabilidade da estatística x . O valor da área sombreada (A) é a probabilidade de x estar entre -2 e 1. p. 22
- 3.6 Valor do nível descritivo p , obtido a partir da função densidade de probabilidade de x , supondo (H_0) verdadeira. p. 23

3.7	Valores de sensibilidade, especificidade, nível de significância e β dados um valor crítico x^* e uma SNR.	p. 23
3.8	Tipos de erros, sensibilidade e especificidade, associados aos testes de hipóteses. . .	p. 24
3.9	Curvas ROC e as áreas associadas a cada uma delas.	p. 25
3.10	Escala de cores para representar ativação.	p. 26
4.1	Distribuições da variável t para diferentes valores de graus de liberdade (ν) e distribuição normal padrão.	p. 31
4.2	Distribuição da estatística t^{θ} para diferentes valores de SNR e para ruído simulado. .	p. 33
4.3	Distribuição t com número de graus de liberdade ($\nu = 18$) sobreposta a distribuição da variável t' calculada para ruído simulado com os parâmetros ($n_x = n_y = 10$). . . .	p. 33
4.4	Curvas ROC para a estatística t^{θ} para diferentes valores de SNR do sinal simulado. Note que os valores de SNR maiores que j 8dB não estão representados uma vez que as curvas para tais valores tocam as extremidades dos eixos.	p. 34
4.5	Mapa de ativação obtido através do teste t . Nível de significância: $\alpha_f = 0,05$; $\alpha_t = 3,13 \times 10^{-6}$, valor t de corte: $t_c = 4,74$	p. 35
4.6	Cinco épocas de uma série temporal de um <i>voxel</i> em comparação com o modelo de resposta hemodinâmica. A correlação quantifica a "semelhança" entre as duas.	p. 37
4.7	Distribuição da estatística r para diferentes valores de SNR e para o ruído.	p. 38
4.8	Curvas ROC para o método da correlação para alguns valores de SNR.	p. 38
4.9	Mapas de ativação obtidos a partir do método da correlação. O valor de corte ($r_c = 0,242$) foi obtido a partir da distribuição de r correspondente a ruído somente, dado $\alpha_f = 0,05$	p. 39
4.10	Distribuição r obtida a partir de ruído simulado.	p. 40
4.11	Distribuição t com ($\nu = 22$) graus de liberdade, sobreposta a distribuição t obtida a partir do coeficiente de correlação r	p. 40
5.1	Entropia de Shannon $H(X)$ para diferentes valores de $p(x_1)$. A entropia é máxima para $p(x_1) = p(x_2) = 0,5$ e mínima para os casos extremos [$p(x_1) = 0; p(x_2) = 1$ e $p(x_1) = 1; p(x_2) = 0$].	p. 43

5.2 Entropia de Tsallis $S_q(X)$ em função de $p(x_1)$ para diferentes valores de q . Para valores negativos de q , a entropia diverge para os casos extremos [$p(x_1) = 0; p(x_2) = 1$ e $p(x_1) = 1; p(x_2) = 0$] e é mínima para [$p(x_1) = p(x_2) = 0,5$]. p. 44

5.3 Entropia de Tsallis $S_q(X)$ em função de $p(x_1)$ para diferentes valores de q . Para q positivo, a entropia é máxima para [$p(x_1) = p(x_2) = 0,5$] e mínima para [$p(x_1) = 0; p(x_2) = 1$ e $p(x_1) = 1; p(x_2) = 0$]. Quando $q = 0$ a entropia é constante independente do valor de $p(x_1)$, e para $q = 1$ obtemos a entropia de Shannon. p. 45

5.4 Curso temporal da HRF modelo. O gráfico ilustra os conjuntos de pontos associados ao aumento do sinal (W_1) e o conjunto correspondente a valores de linha de base (W_2). Os intervalos I_1, I_2, \dots, I_L definem os estados (níveis) acessíveis da série temporal. p. 46

5.5 Distribuição da variável D calculada para ruído; parâmetros $\delta = 0,1$ e $L = 2$ p. 48

5.6 Distribuição da variável D calculada para $SNR = j$ 4dB; parâmetros $\delta = 0,1$ e $L = 2$. p. 48

5.7 Distribuição da variável D calculada para $SNR = 0$ dB; parâmetros: $\delta = 0,1$ e $L = 2$. . p. 49

5.8 Distribuição da variável D calculada para $SNR = 4$ dB; parâmetros: $\delta = 0,1$ e $L = 2$. . p. 49

5.9 Densidade de probabilidade $f(\bar{D})$ para simulações de ruído, e alguns valores finitos de SNR; usando os parâmetros $N = 24, n_1 = n_2 = 7, L = 2$ e $\delta = 0,1$. As curvas contínuas correspondem à aproximação pela função Gamma ($\chi^2 < 0,0009$) dos valores simulados. p. 50

5.10 Otimização do parâmetro δ para os números de níveis ($L = 2$). O valor de δ que maximiza as áreas ROC é $\delta = 0,1$ p. 51

5.11 Otimização do parâmetro δ para os números de níveis ($L = 3$). O valor de δ que maximiza as áreas ROC é $\delta = 0,01$ p. 51

5.12 Otimização do parâmetro δ para os números de níveis ($L = 4$). O valor de δ que maximiza as áreas ROC é $\delta = 0,001$ p. 52

5.13 Otimização do parâmetro δ para os números de níveis ($L = 5$). O valor de δ que maximiza as áreas ROC é $\delta = 0,0001$ p. 52

5.14 Otimização do número de níveis L . A área abaixo da curva ROC é maior ao usar os parâmetros $L = 2$ e $\delta = 0,1$ para uma ampla faixa de valores de SNR. p. 53

5.15 Mapa de ativação obtido a partir da variável \bar{D} calculada para ruído simulado. Foram utilizados os parâmetros otimizados $L = 2$ e $\delta = 0,1$ p. 54

5.16 Mapa de ativação obtido a partir da variável \bar{D} calculada para ruído simulado. Foram utilizados os parâmetros não otimizados $L = 5$ e $\delta = 0,1$ p. 54

5.17 Gráficos ilustrando o número de falsos negativos usando parâmetros otimizados ($L = 2$; $\delta = 0,1$) e não otimizados ($L = 5$; $\delta = 0,1$) para dois valores de SNR ($2dB$ e $4dB$). Parâmetros não otimizados, geram uma quantidade de erros do tipo II maior em relação ao uso dos parâmetros otimizados. Como o nível de significância é o mesmo para ambos os casos ($\alpha_f = 0,05$), o número de falsos positivos é o mesmo. . . p. 55

5.18 Curvas ROC para os parâmetro otimizados ($L = 2$; $\delta = 0,1$) e para os parâmetros não-otimizados ($L = 5$; $\delta = 0,1$). p. 56

5.19 Valores de D_q com $L = 2$, para diferentes valores de q .(a) $q=0,1$; (b) $q=0,3$; (c) $q=0,6$; (d) $q=0,9$. Os valores mais altos de D_q ocorrem quando ($p_{2,1} = 1$ e $p_{1,1} = 0$) ou ($p_{2,1} = 0$ e $p_{1,1} = 1$). p. 58

5.20 Distribuição da variável D_q calculada para ruído; parâmetros: $q = 0,8$ e $L = 3$ p. 59

5.21 Distribuição da variável D_q calculada para $SNR = j 4dB$; parâmetros: $q = 0,8$ e $L = 3$. p. 59

5.22 Distribuição da variável D_q calculada para $SNR = 0dB$; parâmetros: $q = 0,8$ e $L = 3$. p. 60

5.23 Distribuição da variável D_q calculada para $SNR = 4dB$; parâmetros: $q = 0,8$ e $L = 3$. p. 60

5.24 Densidade de probabilidade $f(\bar{D}_q)$ para simulações de ruído, e alguns valores finitos de SNR; usando os parâmetros $N = 24$, $n_1 = n_2 = 7$, $L = 3$ e $q = 0,8$. As curvas contínuas correspondem à aproximação pela função Gamma ($\chi^2 < 0,0001$) dos valores simulados. p. 61

5.25 Otimização do parâmetro q para o números de níveis ($L = 2$). O valor de q que maximiza as áreas ROC é $q = 0,70$ p. 62

5.26 Otimização do parâmetro q para o números de níveis ($L = 3$). O valor de q que maximiza as áreas ROC é $q = 0,80$ p. 62

5.27 Otimização do parâmetro q para o números de níveis ($L = 4$). O valor de q que maximiza as áreas ROC é $q = 0,90$ p. 63

5.28 Otimização do parâmetro q para o números de níveis ($L = 5$). O valor de q que maximiza as áreas ROC é $q = 0,95$ p. 63

5.29 Otimização do número de níveis L . A área abaixo da curva ROC é maior ao usar os parâmetros $L = 3$ e $q = 0,80$ para a faixa de valores de $SNR > j 8dB$ p. 64

5.30 Curvas ROC para os parâmetro otimizados ($L = 3; q = 0,8$). p. 65

5.31 Mapa de ativação obtido a partir do método da distância de Kullback-Leibler Generalizada. Para o cálculo de \bar{D}_q , foram utilizados os parâmetros $L = 3$ e $q = 0,8$. O valor de corte $\bar{D}_q^\alpha \approx 1,73$ foi obtido considerando o nível de significância por família de testes ($\alpha_f = 0,05$). p. 65

6.1 Gráficos mostrando as distribuições densidade de probabilidade da estatística $p(H_1|s_v)$, para SNR=-4dB e para o ruído. p. 70

6.2 Gráficos mostrando as distribuições densidade de probabilidade da estatística $p(H_1|s_v)$, para SNR=-12dB e para o ruído. p. 71

6.3 Gráficos mostrando as distribuições densidade de probabilidade da estatística $p(H_1|s_v)$, para SNR=-16dB e para o ruído. p. 71

6.4 Gráficos mostrando as distribuições densidade de probabilidade da estatística $p(H_1|s_v)$, para SNR=-20dB e para o ruído. p. 72

6.5 Curvas ROC para o método Bayesiano com SNR=-12, -16 e -20dB. p. 72

6.6 Distribuição densidade de probabilidade do *posteriori* $P(H_1|s_v)$ para o valor de *priori* $P(H_1) = 0,3$ p. 73

6.7 Distribuição densidade de probabilidade do *posteriori* $P(H_1|s_v)$ para o valor de *priori* $P(H_1) = 0,5$ p. 74

6.8 Distribuição densidade de probabilidade do *posteriori* $P(H_1|s_v)$ para o valor de *priori* $H_1) = 0,7$ p. 74

6.9 Distribuição densidade de probabilidade do *posteriori* $P(H_1|s_v)$ para o valor de *priori* $P(H_1) = 0,9$ p. 75

6.10 Comportamento da área ROC para (SNR= -8, -12,-16 e -20 dB) em função do *priori*. p. 75

6.11 Mapa de ativação obtido a partir do método Bayesiano usando um *priori* $P(H_1) = 0,3$. O valor de corte $P(H_1|s_v)^\alpha \approx 0,999966$ foi obtido considerando o nível de significância por família de testes ($\alpha_f = 0,05$). p. 76

6.12 Mapa de ativação obtido a partir do método Bayesiano usando um *priori* $P(H_1) = 0,8$. O valor de corte $P(H_1|s_v)^\alpha \approx 0,999995$ foi obtido considerando o nível de significância por família de testes ($\alpha_f = 0,05$). p. 76

7.1 Gráfico comparando os métodos através do comportamento da área ROC. p. 78

A.1	Momento magnético que gira em torno de um eixo. Aproximação de um <i>spin</i> a um ímã(adapt. R.B. Lufkin, 1990).[1]	p. 82
A.2	Spins na ausência de campo magnético externo(adapt. R.B. Lufkin, 1990).	p. 82
A.3	<i>Spins</i> num meio onde se estabeleceu um campo magnético B_0 . M_0 tem o significado de magnetização total do meio(adapt. R.B. Lufkin, 1990).	p. 82
A.4	Representação de <i>spins</i> a precessionarem em torno de um campo magnético externo (B_0) e magnetização total do meio (M_0)(adapt. R.B. Lufkin, 1990).	p. 83
A.5	Conseqüências da aplicação de um campo de rádio-freqüência na magnetização total. Exemplo de um pulso de 90° (adapt. R.B. Lufkin, 1990).	p. 84
A.6	1.8 – Mecanismo de desfasagem dos spins, com conseqüente decaimento do sinal (adapt. R.B. Lufkin, 1990).	p. 84
A.7	Esquema do decaimento provocado pela desfasagem dos spins (adapt. R.B. Lufkin, 1990).	p. 85
A.8	Mecanismo de recuperação da magnetização longitudinal, devido à reorganização das populações de spin entre os estados energéticos, com conseqüente libertação de energia para o meio (adapt. R.B. Lufkin, 1990).	p. 85
A.9	Esquema da evolução da magnetização transversal com o comportamento dos <i>spins</i> , em resposta a um pulso de 180° (adapt. R.B. Lufkin, 1990).	p. 86
A.10	Esquema explicativo sobre como, através da técnica de ecos de <i>spin</i> , é possível obter um sinal que é dependente apenas das interações entre os spins e não considera as heterogeneidades do campo magnético estático (adapt. R.B. Lufkin, 1990).	p. 87
A.11	Esquema explicativo do comportamento da magnetização devido a T_2 e devido a T_1 (adapt. R.B. Lufkin, 1990).	p. 88
A.12	Esquema simplificado do comportamento da magnetização no caso em que $T_1 \ll T_2$	p. 88
A.13	Gráfico da intensidade do sinal em função do tempo para dois tecidos caracterizados por tempos de relaxação <i>spin</i> /rede diferentes (adapt. R.B. Lufkin, 1990).	p. 89
A.14	Gráfico da intensidade do sinal em função do tempo para dois tecidos caracterizados por tempos de relaxação <i>spin/spin</i> diferentes (adapt. R.B. Lufkin, 1990).	p. 89

A.15 Comparação entre as imagens de RMN obtidas através de contraste em: a) densidade de prótons, b) T_1 e c) T_2 . Repare-se que o osso aparece em todas as imagens a escuro (baixa densidade de prótons); o líquido céfalo-raquidiano é escuro na imagem a T_1 é branco na imagem em T_2 ; a mielina é branca nas imagens em T_1 e escura nas imagens em T_2 p. 91

Lista de Tabelas

3.1	Valores da área abaixo da curva <i>ROC</i> associadas às respectivas qualidades.	p. 24
5.1	Parâmetros do ajuste da distribuição de \bar{D} , pela distribuição Gamma. Simulações realizadas utilizando os parâmetro ($N = 24, n_1 = n_2 = 7; L = 2; \delta = 0, 1$).	p. 50
5.2	Parâmetros do ajuste da distribuição de \bar{D}_q simulado, pela distribuição Gamma. Simulações realizadas utilizando os parâmetro ($N = 24, n_1 = n_2 = 7; L = 3; q = 0, 8$).	p. 61
A.1	Valores de T_1 e T_2 para alguns tecidos biológicos.	p. 90

1 *Introdução*

O cérebro é o órgão mais fascinante e um dos menos entendidos do corpo humano. A relação do corpo físico com o comportamento tem sido discutida ao longo dos séculos por cientistas e filósofos. No séc. XVII, Thomas Willis propôs que funções específicas são atribuídas a diferentes áreas do córtex cerebral. Já no séc. XIX, Gall introduziu a "ciência" da frenologia em que o formato do crânio é usado para determinar volumes de regiões cerebrais e com isso inferir o comportamento do indivíduo. Muitas críticas surgiram aos frenologistas, entre elas o fato de fazerem correlações a partir de depoimentos subjetivos com características físicas proeminentes. Além disso, não havia experimentos e muito menos validação estatística dos mapas cranianos propostos. A frenologia, no entanto, foi muito popular, gerando vários livros sobre o assunto. Um deles, por exemplo, fornecia aos empregadores dicas sobre quais eram as características desejadas para um bom funcionário. No auge da frenologia, os mapas possuíam mais de 150 áreas distintas, algumas correspondentes a funções obscuras como "comicalidade" e "velocidade". Por essas e outras razões, a frenologia foi demolida pela comunidade científica no final do séc. XIX. Atualmente, a premissa básica de que uma personalidade poder ser determinada em grande parte pelo formato do crânio é considerada como falsa.

Nas décadas seguintes, a idéia de examinar o crânio foi abandonada e os estudos se voltaram para estímulos do córtex cerebral de animais usando corrente elétrica. Isto levou ao mapeamento da função motora em animais e mais tarde em humanos. O trabalho *Epilepsy and Cerebral Localization: A Study of the Mechanism, Treatment and Prevention of Epileptic Seizures* publicado em 1941 pelo neurocirurgião Wilder Graves Penfield [2], mapeou o córtex motor e o somatosensorial usando estímulos corticais em pacientes submetidos a neurocirurgias. A natureza invasiva dessas medidas impediu um estudo sistemático e muito, no que diz respeito à cognição, permanecia inacessível. Muitos dos estudos da segunda metade do séc. XX são provenientes de pacientes com distúrbios neurológicos ou de medidas com eletrodos em animais. Apenas nas últimas décadas, as técnicas de imagem cerebral possibilitaram o estudo em voluntários humanos saudáveis.

Atualmente, além das imagens, também são utilizadas técnicas não invasivas que medem

atividade elétrica (EEG - *electroencephalography*) e magnética (MEG - *magnetoencephalography*) no cérebro. Esses métodos conseguem detectar mudanças rápidas nos potenciais elétricos e no fluxo magnético, possibilitando o estudo do tempo de processos cerebrais. Porém, a impossibilidade de localizar espacialmente as fontes dessas atividades impedem as técnicas de serem empregadas no mapeamento de ativação neural. Para o mapeamento são utilizadas técnicas de neuroimagens que fornecem a informação a respeito da localização espacial da atividade. Inicialmente, as técnicas utilizadas foram o PET (*Positron Emission Tomography*) e o SPECT (*Single Photon Emission Computer Tomography*), que usam marcadores radioativos para medir fluxo sanguíneo e/ou metabolismo da glucose, o que possibilitou a identificação de regiões cerebrais associadas às funções motoras, perceptivas e cognitivas. Porém, o uso de radiação ionizante, o alto custo de produção de radioisótopos e a baixa velocidade com que as imagens são adquiridas fizeram com que essas técnicas tivessem um crescimento menor nas últimas décadas.

A imagem funcional por ressonância magnética (fMRI) surgiu no cenário de neuroimagens com o diferencial de não submeter o voluntário a efeitos nocivos. A técnica consiste em obter imagens funcionais do córtex cerebral utilizando um agente de contraste endógeno, a deoxi-hemoglobina. O fato de não necessitar de um contraste exógeno aliado a imagens com alta resolução espacial (mm) e temporal (s) tornaram a imagem por ressonância magnética a principal escolha em estudos com muitos voluntários e para diferentes tipos de estímulos [3, 4].

A atividade neural devido a um estímulo causa um aumento do fluxo sanguíneo e oxigenação nos vasos locais. O contraste BOLD (*Blood Oxygenation Level Dependent*) da imagem é induzido por mudanças nas concentrações relativas de oxi e deoxi-hemoglobina. O sinal BOLD é a resposta da aplicação de um protocolo experimental (conhecido como paradigma) que é composto de mudanças periódicas entre estímulo e não-estímulo de diferentes áreas cerebrais. O resultado de tal experimento é uma série temporal correspondente ao sinal para cada voxel.

As imagens obtidas em um experimento de fMRI geralmente apresentam uma baixa relação sinal ruído (SNR - *signal to noise ratio*). Por esta razão, além do pré-processamento, métodos de processamento adquirem uma importância crucial na análise dos dados. Os métodos de processamento consistem em atribuir a cada voxel da imagem a condição de ativo ou inativo. Entretanto, existe um problema quanto à falta de consenso entre os pesquisadores em fMRI sobre a escolha do método ideal. Além disso, ao atribuir atividade a um voxel, deve se estabelecer um nível descritivo (*valor p*) que reflete a confiabilidade daquele voxel estar realmente ativo.

No processamento de séries temporais de fMRI, têm sido empregados diversos métodos para separar os sinais fisiologicamente induzidos (correspondentes a voxéis ativos) de ruído

(voxeis inativos). Métodos determinísticos (como a correlação) quantificam a similaridade entre a série temporal de cada voxel com uma função resposta hemodinâmica (HRF - *Hemodynamic Response Function*). Por outro lado, métodos estatísticos, como o teste t de Student, podem inferir quão significativa é a diferença entre sinais correspondentes a períodos de estímulo e repouso, com a vantagem de não necessitarem do conhecimento da forma da HRF para análise de dados reais. Nos últimos anos, métodos baseados em medidas de informação, tais como as entropias de Shannon e Tsallis e a informação mútua generalizada, têm sido empregados como alternativas à análise convencional de dados de fMRI [5, 6, 7, 8, 9]. Estas medidas de informação também não se utilizam de um modelo de HRF para análise. Finalmente, dentre muitos outros, caberia ainda aqui citarmos os métodos baseados em inferência Bayesiana, os quais têm sido aplicados com sucesso na análise de fMRI. Recentemente, Amaral et al. [10, 11] mostraram como os métodos Bayesianos (método do pixel independente e o de multiescala) podem ser aplicados nesta análise. Estes, por sua vez, requerem um modelo para a HRF.

No presente trabalho, conceitos de inferência estatística são utilizados para aplicação e comparação de diferentes métodos de análise de sinais de ressonância magnética funcional. Nosso estudo é projetado para paradigmas do tipo evento-relacionado (ER), em que um breve estímulo é seguido de um longo período de repouso. A idéia central baseia-se na obtenção da distribuição de probabilidade da variável aleatória de interesse, para cada método estudado e sob diferentes valores da relação sinal-ruído (SNR). Este objetivo é atingido através de simulações numéricas da função resposta hemodinâmica (HRF) acrescida de ruído gaussiano. Tal procedimento nos permite avaliar a sensibilidade e a especificidade dos métodos empregados através da construção das curvas ROC (*receiver operating characteristic*) para diferentes valores de SNR.

Sob condições experimentais específicas, aplicamos métodos clássicos de análise (teste t de Student e correlação), medidas de informação (distância de Kullback-Leibler e sua forma generalizada) e um método Bayesiano (método do *pixel* independente).

No capítulo 2, apresentamos os conceitos básicos envolvidos no estudo de imagens por ressonância magnética funcional, descrevendo os mecanismos que levam ao contraste endógeno conhecido como BOLD. Mostramos os tipos de paradigmas usados para evocar este contraste e a função resposta hemodinâmica (HRF) que o modela. Também abordamos as principais técnicas de pré-processamento e processamento das imagens.

No capítulo 3, mostramos como as séries temporais são simuladas com a finalidade de obter a distribuição de probabilidade da variável aleatória de interesse. Descrevemos os testes de hipóteses e como construir as curvas ROC. Apresentamos, também, o protocolo de aquisição de dados reais e como esses, após analisados, são mostrados em mapas estatísticos.

No capítulo 4, apresentamos os resultados das simulações para métodos clássicos de análise (teste t e correlação). Obtemos, através desses métodos, as curvas ROC e os mapas de ativação neural para dados reais.

No capítulo 5, mostramos que a distância de Kullback-Leibler D (ou entropia relativa) e sua forma generalizada são medidas úteis para análise de sinais dentro do cenário da teoria de informação. Obtemos as densidades de probabilidade das médias amostrais destas entropias, mostrando que estas podem ser aproximadas por distribuições Gamma (para pequenas amostras). Obtemos as curvas ROC e os mapas de ativação neural para dados reais.

No capítulo 6, apresentamos o método (Bayesiano) do pixel independente. Estimamos a distribuição das probabilidades *posteriori* para experimentos do tipo evento-relacionado. Deste modo, obtemos as curvas ROC para vários valores de SNR. Investigamos a influência do *priori* quanto ao desempenho do método e, finalmente, obtemos os mapas de ativação para um conjunto de dados reais.

No capítulo 7, mostramos o comportamento da área sob a curva ROC como função da SNR, permitindo assim a comparação entre os métodos empregados.

As conclusões deste trabalho e perspectivas futuras são apresentadas e discutidas no último capítulo.

O apêndice A aborda mais detalhadamente a teoria de imagem por ressonância magnética nuclear.

2 *Imagem Funcional por Ressonância Magnética*

Imagem por ressonância magnética (IRM) é uma técnica de obtenção de imagens baseada no fenômeno físico da ressonância magnética nuclear (RMN). Ela é utilizada em procedimentos clínicos para obtenção de imagens internas do corpo humano. A primeira imagem do corpo humano, através do uso da RMN, foi obtida em 1977, mas a tecnologia só foi adotada para uso clínico por volta de 1988.

A ressonância magnética nuclear é o fenômeno que ocorre quando partículas sub-atômicas presentes no núcleo (prótons e nêutrons), são imersas em um campo magnético estático e, em seguida, são expostas a um campo magnético variável. Quando o número de massa do elemento é ímpar (prótons e nêutrons desemparelhados), existe um momento magnético não nulo. Este momento magnético interage com o campo magnético externo em que está imerso. Portanto, os fenômenos de ressonância magnética só ocorrem para estes tipos de elementos.

O Hidrogênio é o elemento mais abundante no corpo humano e apresenta a característica de momento magnético não nulo. Quando colocado em um campo magnético externo (B_0), o vetor do *spin* se alinha com esse campo magnético. Surgem então duas possíveis configurações de energia, *spins* alinhados paralelamente com o campo magnético (baixa energia) e anti-paralelamente (alta energia). A quantidade de prótons na configuração de baixa energia é ligeiramente maior que a quantidade em alta energia, gerando uma magnetização da rede na direção do campo magnético externo. Ao aplicar um segundo campo magnético (B_1), perpendicularmente ao (B_0), os prótons de Hidrogênio realizam um movimento de precessão. Esta precessão ocorre com uma frequência que depende do campo (B_0),

$$f = \gamma B_0, \quad (2.1)$$

em que B_0 é o campo magnético medido em Tesla e $\gamma = 42,58\text{MHz}/\text{T}$ é razão giromagnética do Hidrogênio. A frequência de precessão f é também conhecida como *frequência de Larmor*.

Se a frequência do campo magnético oscilante B_1 , aplicado por um breve período de tempo, é a mesma da frequência de Larmor, uma quantidade de energia é transferida para os prótons, e estes passam a precessionar com coerência de fase. O sinal de ressonância magnética (RM) criado após o pulso de excitação decai no tempo devido a um fenômeno chamado de relaxação nuclear. São dois os mecanismos que contribuem para a perda do sinal de RM: tempo de relaxação longitudinal T_1 também conhecido como *spin/rede* e tempo de relaxação transversal T_2 também chamado de *spin/spin*.

As imagens por ressonância magnética são formadas a partir de características dos *spins* nucleares que constituem os tecidos analisados. Imagens ponderadas por densidade de prótons, fornecem informação a respeito do número de prótons contidos em um *voxel*. A ponderação pelos tempos de relaxação T_1 e T_2 refletem características da magnetização do tecido após a aplicação de um pulso eletromagnético.

O contraste endógeno conhecido como BOLD é responsável pelas mudanças no sinal de ressonância magnética que caracterizam a ativação neural. Essa mudança pode ser modelada através de uma função resposta hemodinâmica (*Hemodynamic Response Function - HRF*). Este capítulo tem como objetivo apresentar os fenômenos fisiológicos que levam à resposta hemodinâmica bem como um modelo plausível para a análise de dados e simulações. Também serão abordadas as principais técnicas de pré-processamento.

2.1 O contraste BOLD

O neurônio é a unidade básica de processamento de informação do sistema nervoso central. O cérebro humano contém cerca de 100 bilhões de neurônios, dos quais cerca de 20 bilhões estão localizados no córtex. Os neurônios possuem três partes principais: dendritos, corpo celular e axônio, cf. fig. 2.1(a). Através dos dendritos, cada neurônio recebe impulsos elétricos (potenciais de ação) provenientes de vários outros neurônios e esses impulsos são integrados e transmitidos a outros neurônios pelos terminais de transmissão. Essa transferência de impulsos entre neurônios ocorre através da despolarização da membrana pré-sináptica abrindo canais iônicos de Ca^{2+} . Esses íons, ao adentrarem a célula, fazem com que vesículas contendo neurotransmissores liberem seu conteúdo na sinapse. Os neurotransmissores se ligam com receptores em canais iônicos pós-sinápticos abrindo-os. Assim, íons como o Na^+ entram na célula pós-sináptica mudando seu potencial, gerando um novo impulso, cf. fig. 2.1(b).

Após a transmissão do impulso, energia é necessária para restaurar a configuração inicial e tornar a sinapse apta a receber um novo potencial de ação. Essa energia é obtida através da gli-

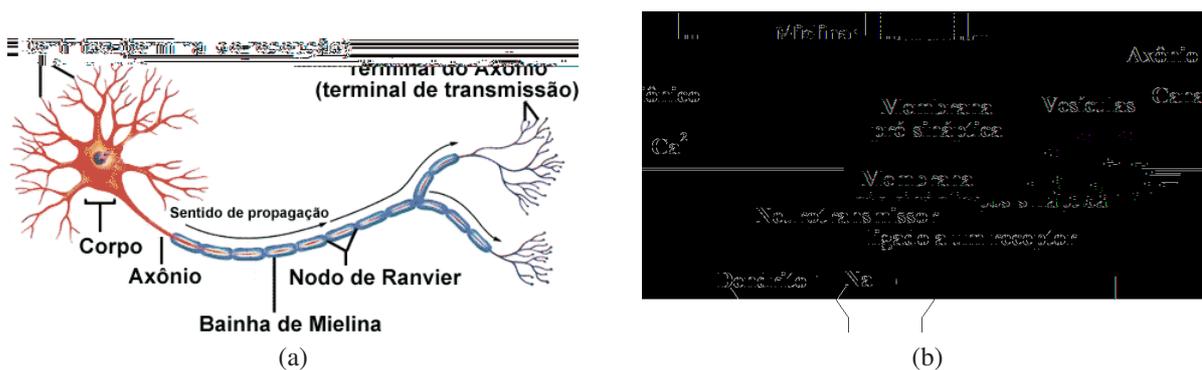


Figura 2.1: (a) Representação esquemática de um neurônio; (b) Processo de transmissão de impulso na sinapse.

cólise. Para a glicólise por processo aeróbico, é necessário o oxigênio trazido pela hemoglobina. A hemoglobina quando ligada ao oxigênio (*Hb*) é diamagnética, ou seja, não apresenta elétrons desemparelhados, resultando em um momento magnético nulo. A deoxi-hemoglobina (*dHb*) é paramagnética pois possui elétrons desemparelhados e conseqüentemente, momento magnético diferente de zero. É a concentração relativa de *dHb* que muda as propriedades magnéticas das moléculas de água próximas, gerando um contraste no sinal de ressonância magnética. O contraste endógeno resultante é conhecido como BOLD (*Blood-Oxygenation-Level Dependent*). É importante ressaltar que o contraste BOLD é uma conseqüência indireta da ativação neural e o seu mecanismo exato envolve o fluxo sanguíneo entre outros fatores ¹.

O sinal de ressonância magnética provocado pela mudança na concentração relativa de *dHb* é conhecida como resposta hemodinâmica. Como os exames de ressonância magnética funcional utilizam imagens ponderadas em T_2^* , quanto maior a concentração relativa de *dHb* (que é paramagnética), menor o sinal e vice-versa. Alguns estudos relatam que logo após o estímulo, é observado um pequeno decréscimo da resposta hemodinâmica, atribuído a um aumento inicial de *dHb*. Em seguida, após um curto período de latência, a demanda metabólica gera um aumento de sangue oxigenado. Assim, mais oxigênio é fornecido ao local de atividade do que extraído, e isto resulta na diminuição na quantidade de *dHb* em um elemento de volume (*voxel*), gerando um sinal acima de uma linha de base $2s$ após o estímulo, até alcançar um valor máximo (pico) $5s$ após o estímulo. Se a atividade neural for estendida, o pico é da mesma maneira estendido em um platô. Depois de atingir o pico, o sinal diminui até ficar abaixo da linha de base, este efeito, conhecido como *undershoot*, é gerado devido ao fato de o fluxo sanguíneo diminuir mais rápido do que o volume, aumentando a concentração relativa de *dHb*. A resposta hemodinâmica associada a um breve estímulo está representada na fig. 2.2 (a) enquanto que a

¹O modelo do balão proposto por Buxton [12] em 1998 é um modelo matemático e biomecânico que explica e prevê o comportamento da resposta BOLD.

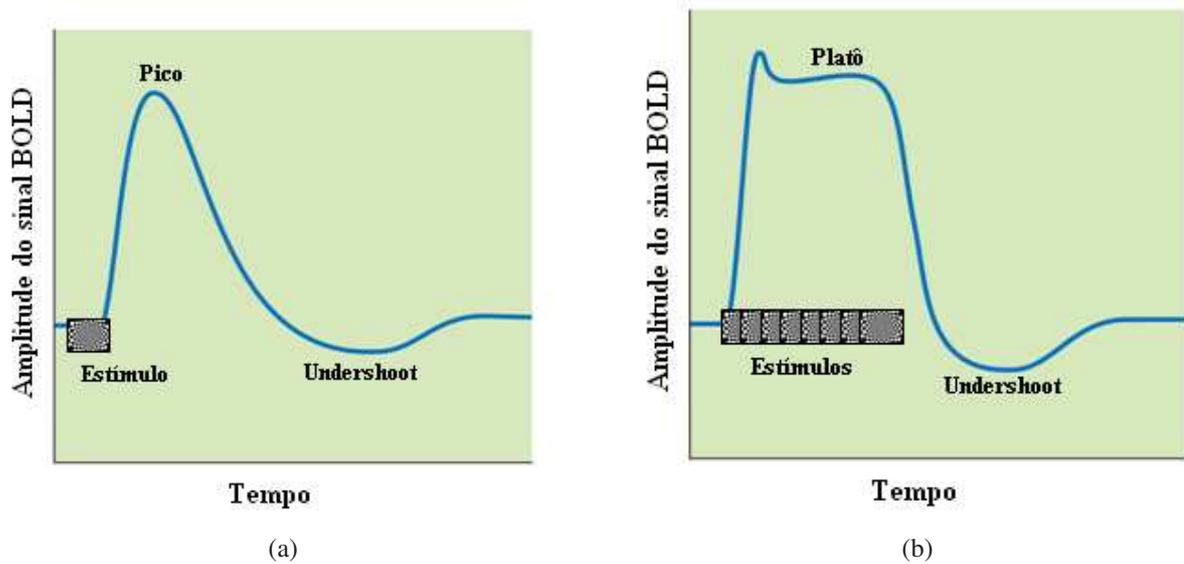


Figura 2.2: Representação esquemática da resposta hemodinâmica. (a) Resposta a um estímulo de curta duração; (b) Resposta a estímulos múltiplos consecutivos.

fig. 2.2 (b) mostra a resposta associada a vários estímulos consecutivos.

2.2 Tipos de paradigma

O tipo de resposta observada depende do tipo de estímulo e de como eles são apresentados. As diferentes formas de apresentar estímulos são chamadas de paradigmas e existem basicamente três tipos:

- ² bloco: as condições experimentais (por exemplo, estímulo e repouso) são separadas em blocos distintos e cada condição é apresentada por um longo período de tempo;
- ² evento relacionado: o estímulo é apresentado por um curto período de tempo e o repouso é longo. Um tipo particular de evento relacionado é o semi-aleatório em que a probabilidade de um estímulo ocorrer dentro de um intervalo de tempo muda sistematicamente ao longo do experimento;
- ² mistos: combinam o paradigma em bloco com o evento relacionado.

Nos estudos de ressonância magnética funcional não existe um paradigma ideal, a escolha depende dos objetivos do experimento. O paradigma em bloco possui uma análise simples e tem um excelente poder de detecção, porém possui baixo poder de estimação da resposta

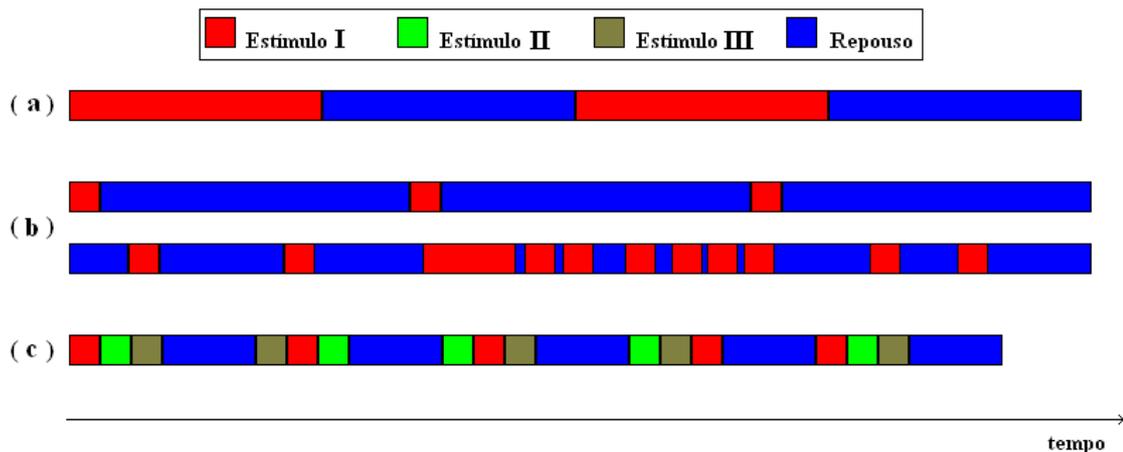


Figura 2.3: Representação dos tipos de paradigma. (a) Bloco; (b) Evento relacionado e semi-aleatório; (c) Misto.

hemodinâmica. Por sua vez, o evento relacionado tem um bom poder de estimação da resposta hemodinâmica, possui estratégias de análises flexíveis, mas por outro lado, tem o poder de detecção reduzido e é sensível ao modelo da resposta hemodinâmica escolhido. A estratégia mista pode ser uma boa opção por reunir as vantagens dos dois métodos anteriores, entretanto sua análise por ser mais complicada, pode comprometer o estudo.

2.3 Modelo de função resposta hemodinâmica (HRF)

A HRF pode ser modelada de diferentes formas, sendo encontrados na literatura diversos modelos, entre eles: distribuição de Poisson [13]; distribuição gamma [14]. Porém, um modelo mais preciso é a soma de duas distribuições gamma [15], que inclui a característica do *undershoot*. Neste trabalho utilizamos o modelo de função resposta hemodinâmica $h(t)$ conforme proposto por Glover em 1999 [16], em que:

$$h(t)$$

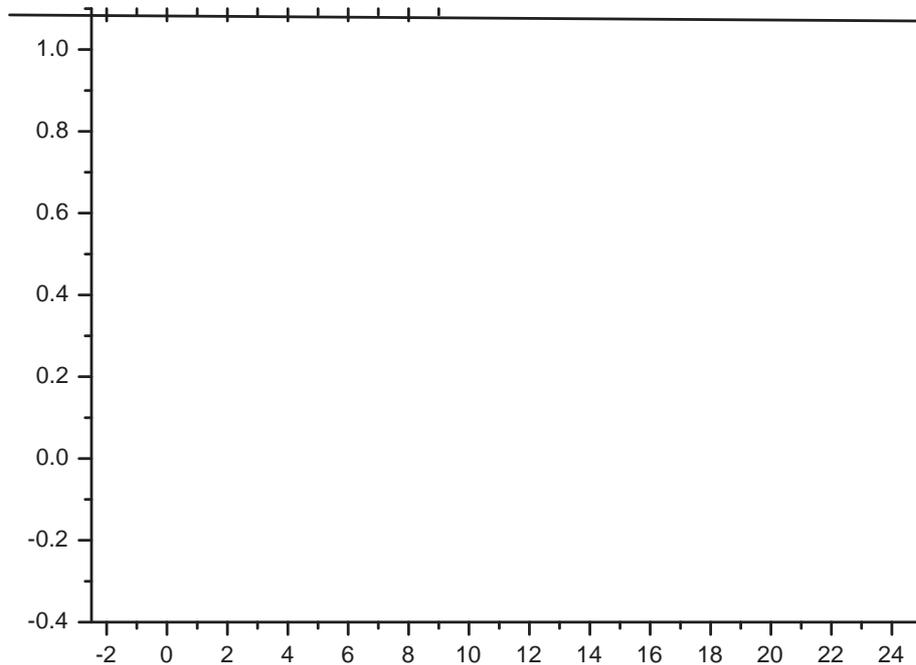


Figura 2.4: Modelo de resposta hemodinâmica para um paradigma de evento relacionado, utilizando o modelo da eq. (2.2) com os parâmetros: $a = 6$, $a^0 = 12$, $b = b^0 = 0.9$, $c = 0.35$, $d = ab$ e $d^0 = a^0 b^0$

2.4 Aquisição das imagens

Trataremos nesta seção de aspectos da técnica de aquisição de imagens e de como são obtidos os dados reais utilizados neste trabalho.

As imagens de fMRI podem ser adquiridas em diferentes cortes anatômicos. A fig. 2.5, mostra os tipos de cortes: axial, coronal e sagital. Para um mesmo tipo de corte em diferentes posições usamos o termo "fatia".

Em experimentos de fMRI, várias imagens são adquiridas ao longo do tempo para uma mesma fatia. Para cada *voxel*, existe um conjunto de pontos que representa a variação da intensidade desse *voxel* ao longo do tempo. Esse conjunto de pontos é chamado de *série temporal*, cf. fig. 2.6. Ao analisar a série temporal de determinado *voxel* podemos dizer se ele é ativo ou não através dos métodos de processamento das imagens.

Dentre as técnicas de obtenção de imagens por ressonância magnética, a técnica *Echo-Planar Imaging - EPI*, proposta por Peter Mansfield em 1977 [17], é uma das mais antigas. A

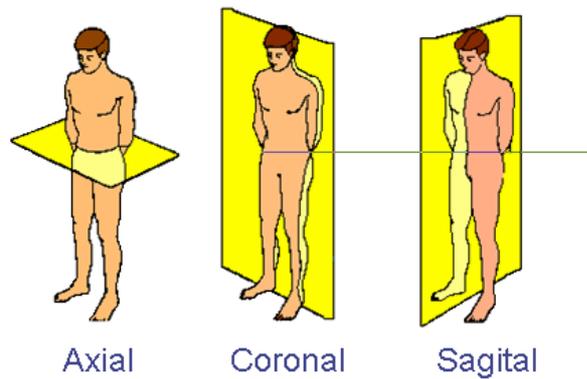


Figura 2.5: Tipos de cortes anatômicos que podem ser visualizados em imagens de ressonância magnética funcional.

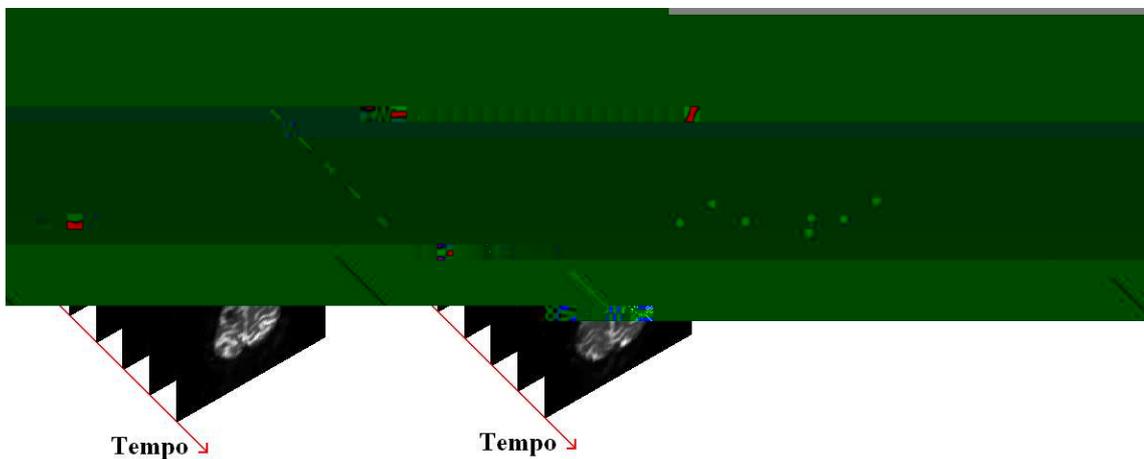


Figura 2.6: Sequência de imagens obtidas ao longo do tempo para duas fatias axiais. Ao lado, série temporal correspondente a um *voxel* da fatia 2.

imagem por EPI é obtida a partir de um único sinal FID (*Free Induction Decay*), entre 40 a 100 milissegundos.

Por ser uma técnica de obtenção de imagens muito rápida, a sequência de pulsos EPI pode fornecer uma base para estudos de processos dinâmicos, como por exemplo o efeito *BOLD*. Dessa forma, tal sequência é a principal técnica utilizada em fMRI.

Porém, a sequência de pulsos EPI exige muito do *hardware* pois grandes gradientes de campo têm que ser gerados. Outro grave inconveniente do método é a sua susceptibilidade a artefatos, que muitas vezes podem resultar em graves distorções nas imagens.

A importância dos estudos funcionais em neurociência é a principal razão para o rápido desenvolvimento da técnica EPI. Ao longo das últimas décadas, melhorias significativas foram feitas no *hardware*. Do mesmo modo, uma série de novas idéias relacionadas com experimentos e análise de dados para lidar com artefatos foram propostos. Atualmente, os aparelhos de

ressonância magnética chegam a operar com campos de 1,5 a 8 Teslas e são capazes de produzir imagens EPI com resolução milimétrica.

Por ser uma técnica susceptível à artefatos, as imagens obtidas possuem baixa relação sinal-ruído (*SNR - Signal to Noise Ratio*). Assim, para tornar possível a aplicação de métodos de análise estatística (processamento) é necessário antes um pré-processamento dos dados.

2.5 Pré-processamento

O propósito do pré-processamento é reduzir a variabilidade que não está associada ao estímulo e preparar os dados para o processamento.

Quando adquirimos as imagens, admitimos que cada *voxel* representa um local único e imutável no cérebro e que a amostragem ao longo do tempo ocorre a uma taxa regular. Porém, os dados de *fMRI* apresentam imprecisões temporais e espaciais causadas por vários fatores, entre eles:

- ² movimento da cabeça do indivíduo;
- ² oscilações fisiológicas (batimentos cardíacos, respiração);
- ² inhomogeneidades do campo estático;
- ² diferenças no tempo de aquisição das imagens.

Se não forem corrigidas, tais variabilidades podem diminuir ou até mesmo eliminar o poder de detecção de um experimento.

2.5.1 Correção temporal das fatias

Os dados de *fMRI*, são obtidos utilizando seqüências de pulsos, que adquirem uma fatia por vez. Assim, cada fatia estará defasada em relação a anterior, influenciando na correspondência entre a série temporal dos dados e a hipótese experimental. A solução usual para este tipo de problema é a interpolação de valores, colocando as fatias em fase. As interpolações podem ser realizadas por funções lineares, polinômios ou sinc². Na fig. 2.7 está ilustrada uma interpolação linear.

²a função sinc é definida como: $\text{sinc}(x) = \text{sen}(\pi x) / \pi x$.

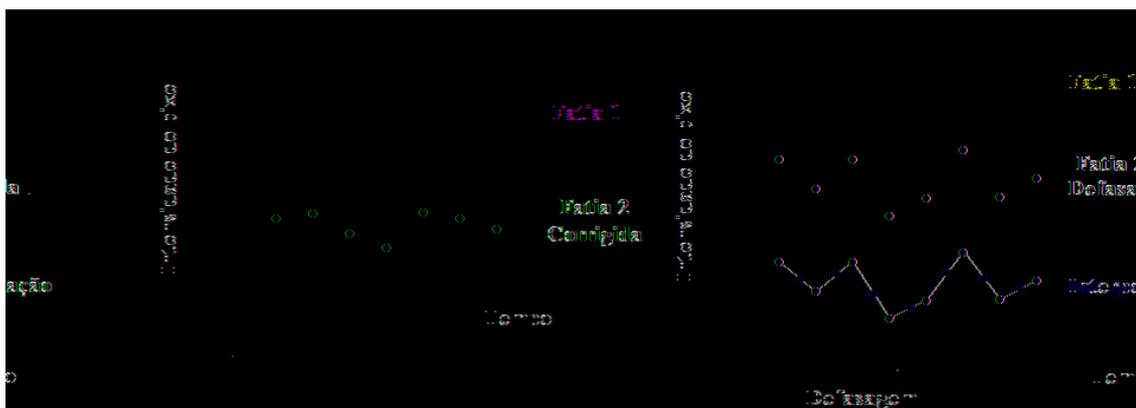


Figura 2.7: Ilustração mostrando a interpolação linear para correção temporal das fatias.

Apesar da correção temporal entre as fatias não recuperar totalmente a informação perdida entre as amostras, ela pode ser efetiva para taxas de amostragens maiores que a variabilidade dos dados.

2.5.2 Movimento da cabeça

O equipamento de ressonância magnética adquire as imagens da localização absoluta e não relativa da posição do cérebro. Assim, se o indivíduo mexer 5 mm a cabeça, as séries temporais da cada *voxel* conterão dados de duas partes diferentes do cérebro separadas por 5 mm. Mesmo um pequeno movimento de cabeça pode causar grandes mudanças no sinal ao longo do tempo.

O movimento de cabeça do indivíduo em geral é involuntário, porém o arranjo experimental pode fazer com o indivíduo tenha mais chance de mexer a cabeça. Algumas tarefas como o movimento de um *joystick* ou apertar um botão podem induzir o movimento. Como movimento da cabeça é mais facilmente prevenido do que corrigido, alguns recursos podem ser utilizados para restringir este tipo de movimento, entre eles:

- ² barras de morder, onde o indivíduo morde firmemente um molde dentário immobilizando a mandíbula, e conseqüentemente restringindo o movimento da cabeça;
- ² máscaras confeccionadas individualmente para cada sujeito que previnem o movimento ao fixar a posição da cabeça;
- ² sacos de vácuo colocadas como apoio para a cabeça. Os sacos são feitos de um material plástico flexível cheios de grânulos macios que ao ter o ar retirado, seu formato fica adequado à forma da cabeça do indivíduo, aumentando conforto e reduzindo mobilidade.

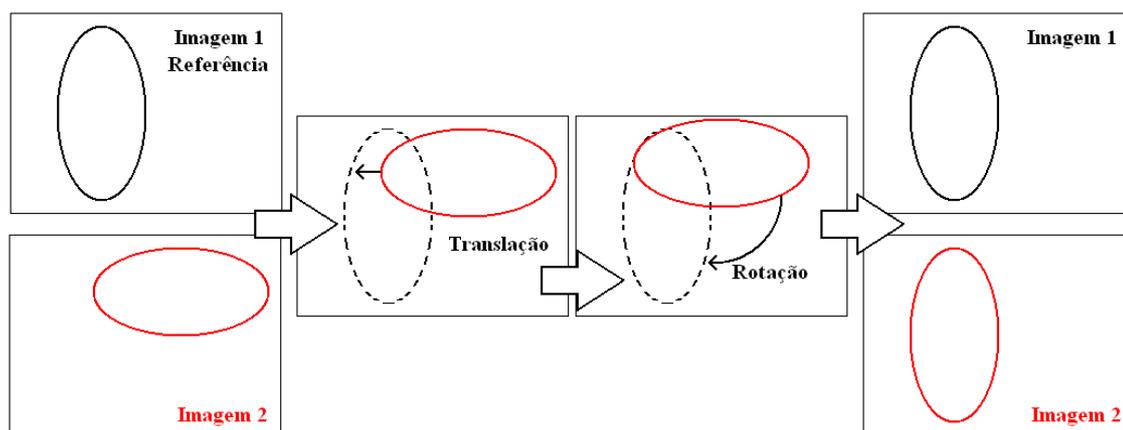


Figura 2.8: Ilustração mostrando a transformação de corpo rígido para alinhamento das imagens.

As técnicas de imobilização da cabeça são importantes, porém o fator mais crucial é a colaboração do sujeito. Por isso, é importante se certificar do conforto do indivíduo e interesse dele no estudo. Porém, nem sempre é possível garantir a imobilidade total da cabeça. Assim procedimentos que corrigem estes movimentos devem ser adotados.

O objetivo da correção de movimento de cabeça é ajustar as séries temporais das imagens de forma que o cérebro esteja sempre na mesma posição. O processo de alinhar essas imagens é chamado de *corregistro*. Para a correção de movimento, sucessivas imagens da série temporal são corregistradas em relação a uma imagem de referência. Através de transformação de *corpo rígido*, as imagens sofrem translações e rotações para se ajustarem a imagem de referência, cf. fig. 2.8. Este ajuste é dado por uma *função custo*, que quantifica a similaridade com a imagem de referência. Quanto menor o custo mais próximo da imagem de referência.

2.5.3 Correção da distorção

As imagens funcionais geralmente sofrem de distorções geométricas. A principal causa destas distorções é a inhomogeneidade do campo estático. Para prevenir a inhomogeneidade do campo, é usada uma bobina do tipo *shimming* que gera campos magnéticos compensatórios de forma a homogenizar espacialmente a intensidade do campo magnético. Outra forma, é mapear o campo magnético produzido e incorporar esses mapas na rotina de reconstrução da imagem para corrigir qualquer distorção geométrica.

2.5.4 Filtros

O uso de filtros temporais podem melhorar substancialmente a qualidade dos dados de fMRI, aumentando a SNR. Ao aplicar a *transformada de Fourier* em uma série temporal, podemos utilizar filtros para retirar frequências indesejáveis como a de batimentos cardíacos e da respiração. Outro tipo de artefato indesejável, conhecido como *drift*, consiste em uma variação linear lenta, que pode ser removida com o uso de um filtro passa alta. O *drift* ocorre devido a instabilidade instrumental, ruído instrumental ou movimentos do sujeito.

Filtros espaciais do tipo passa baixa são empregados para reduzir altas frequências espaciais, "borrando" as imagens. A técnica mais comum de borramento é a utilização de um filtro gaussiano. Quando este é aplicado, ele espalha a intensidade do *voxel* para *voxels* vizinhos e o efeito de borramento reduz a variabilidade dos dados, aumentando a SNR.

2.6 Processamento

Após a etapa de pré-processamento, os dados estão prontos para serem analisados. Na etapa de processamento são aplicados testes para identificar áreas de ativação neural. Estes testes utilizam as séries temporais dos *voxels* da imagem.

Métodos como a correlação, verificam a semelhança da série temporal com o que é esperado pelo modelo da HRF. Para uma dada série temporal, quanto maior o valor do coeficiente de correlação linear r , mais provável que aquele *voxel* esteja ativo.

Outros métodos não utilizam um modelo de HRF para inferir se há ou não ativação. O teste estatístico t de Student por exemplo, compara as médias dos dados relativos à ativação com a média dos dados relativos ao repouso. Se houver uma diferença significativa, o *voxel* é considerado ativo. Dentre as várias técnicas propostas para analisar dados de fMRI, podemos citar: Teste-t de Student [18]; Teste de Kolmogorov-Smirnov [19]; Correlação [20]; Análise de Fourier [21]; Transformada de ondaletas (*wavelets*) [22]; Modelo linear geral [23]; Métodos Bayesianos [11]; Métodos entrópicos [7, 8];

3 *Simulações das Séries Temporais, Testes de Hipóteses e Mapas Estatísticos*

Os dados obtidos em experimentos de fMRI devem ser processados a fim de revelar o mapa de ativação relativo ao estímulo aplicado. De modo geral, utilizamos a inferência estatística para testar a hipótese nula (H_0) de que um *voxel* está inativo (ruído) *versus* a hipótese alternativa (H_1) de que o mesmo está ativo. Para a realização deste teste de hipóteses, construímos a distribuição de probabilidade da estatística de interesse, através de simulações numéricas das séries temporais correspondentes ao sinal acrescido de ruído. Neste capítulo, apresentamos os tipos de séries temporais simuladas e como a sensibilidade e especificidade do teste são calculadas. Outrossim, mostramos como o método pode ser empregado para análise de dados reais, com determinação de seu nível descritivo (*valor-p*).

3.1 Simulações das séries temporais

As séries temporais obtidas são corrompidas com ruído e quanto maior o ruído, mais difícil é detectar ativação neural, caso ela exista. Alguns métodos de análise podem ser mais susceptíveis ao ruído que outros. As simulações são utilizadas para analisar quanto um dado método é susceptível ao ruído. Métodos pouco susceptíveis apresentam alta sensibilidade e especificidade, mesmo para valores baixos de relação sinal-ruído (SNR).

3.1.1 Ruído

O termo "ruído" é em geral usado para definir qualquer componente indesejável na observação de determinada característica (sinal). Em fMRI, o ruído é uma componente com variabilidade irrelevante, presente no sinal associado à resposta BOLD. As principais fontes de ruído associado à fMRI são:

1. **Ruído térmico intrínseco ao sujeito e à eletrônica do scanner.** O ruído térmico é proveniente do movimento dos elétrons no sujeito e também dos componentes eletrônicos do scanner. O sinal de fMRI é processado por uma série de componentes eletrônicos onde elétrons livres colidem com átomos, resultando em trocas de energia. Quanto maior a temperatura do sistema, maior o número de colisões e a distorção do sinal. Pelo fato dos *voxels* dentro do cérebro terem um sinal mais intenso, o ruído térmico causa um ligeiro aumento ou diminuição na intensidade dos valores medidos, resultando em valores com uma distribuição *gaussiana* das intensidades. Fora do cérebro, como só existe ar, não há sinal, assim, o ruído térmico só pode gerar pequenos valores positivos. As intensidades do sinal fora do cérebro seguem uma distribuição de *Rayleigh* [18].
2. **Ruídos do sistema associados a imperfeições do hardware.** São causados por inhomogeneidades do campo, não-linearidades e instabilidade nos gradientes de campo. O efeito de *drift* por exemplo, pode ser causado por ligeiras mudanças no campo magnético, alterando a frequência de ressonância dos prótons.
3. **Movimento e ruído fisiológico.** Artefatos resultantes do movimento da cabeça, ombro, pernas etc. e processos fisiológicos como batimentos cardíacos e respiração.
4. **Variabilidade na ativação neural não associadas às tarefas (estímulos).** O estímulo que se quer estudar pode sofrer interferência de outros estímulos não relacionados à tarefa de interesse, por exemplo, estímulos visuais, sonoros etc.
5. **Variabilidade no comportamento e na cognição do sujeito.** Tarefas que exigem execução de um movimento, por exemplo, podem sofrer variações no tempo de reação para diferentes sujeitos ou até mesmo para um mesmo sujeito ao longo do experimento.

A maioria dos tipos de ruído citados podem ser diminuídos através de técnicas de pré-processamento ou pela construção de um procedimento experimental adequado. Porém, o sinal nunca estará completamente livre de ruído, principalmente do ruído térmico pois este é intrínseco. A fig. 3.1 apresenta uma série temporal obtida a partir de dados reais que já foram submetidas ao pré-processamento.

Pelo fato de os dados reais por nós utilizados já terem passado por um pré-processamento, admitimos que a única componente ou a componente predominante do ruído, é o de origem térmica. Assim, neste trabalho, o termo "ruído térmico" será chamado apenas de "ruído".

Ao tentar identificar ativação neural, estamos examinando o córtex, portanto o modelo de ruído usado segue uma distribuição gaussiana. Em nossas simulações, utilizamos um algoritmo

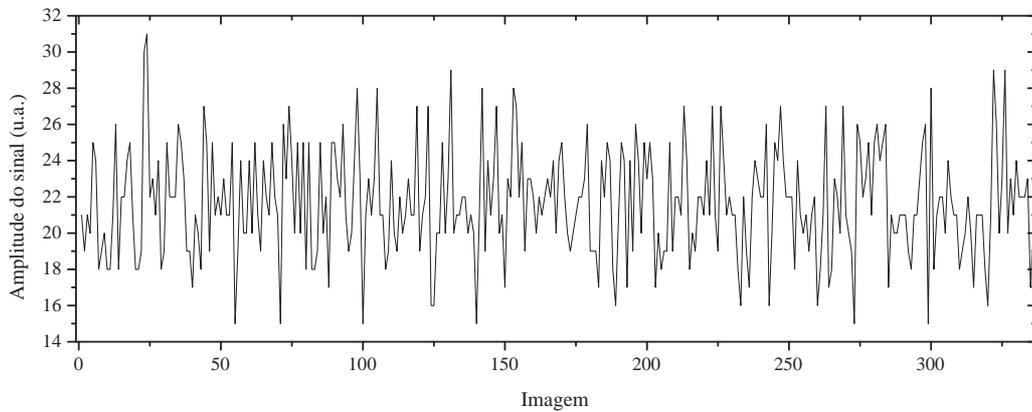


Figura 3.1: Série temporal de um *voxel*. Variações são devidas principalmente ao ruído térmico.

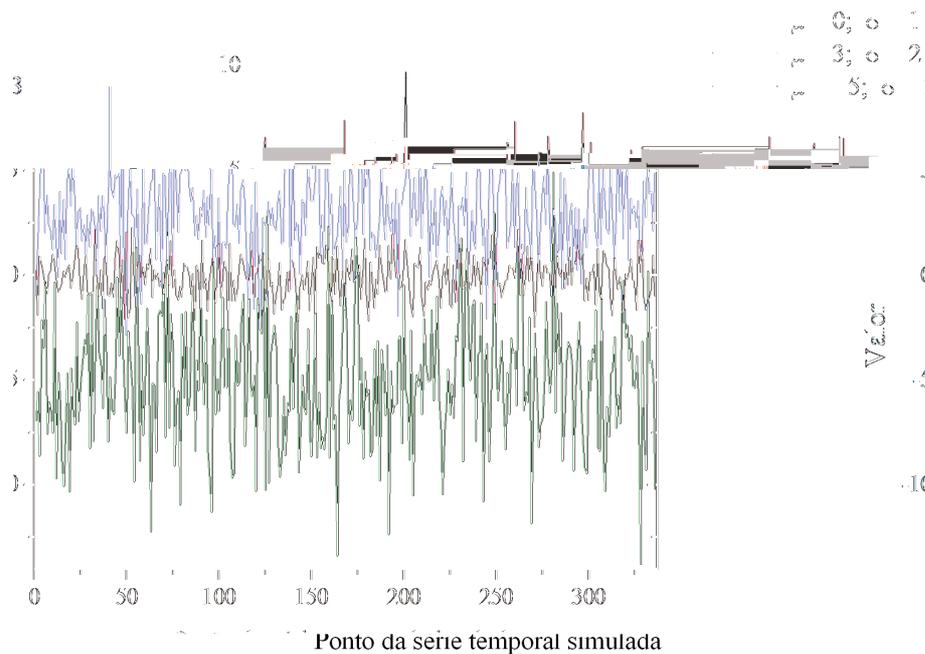


Figura 3.2: Ruído gaussiano simulado para diferentes valores de média e desvio-padrão.

que gera números pseudo-aleatórios com distribuição gaussiana. A fig. 3.2 mostra ruído simulado para diferentes valores de média e desvio-padrão e a fig. 3.3 as respectivas distribuições densidade de probabilidade. Os valores de desvio-padrão do ruído determinarão as características da relação sinal ruído (SNR - *Signal to Noise Ratio*), como será mostrado a seguir.

A função densidade de probabilidade (*pdf* - *probability density function*) da gaussiana é dada por

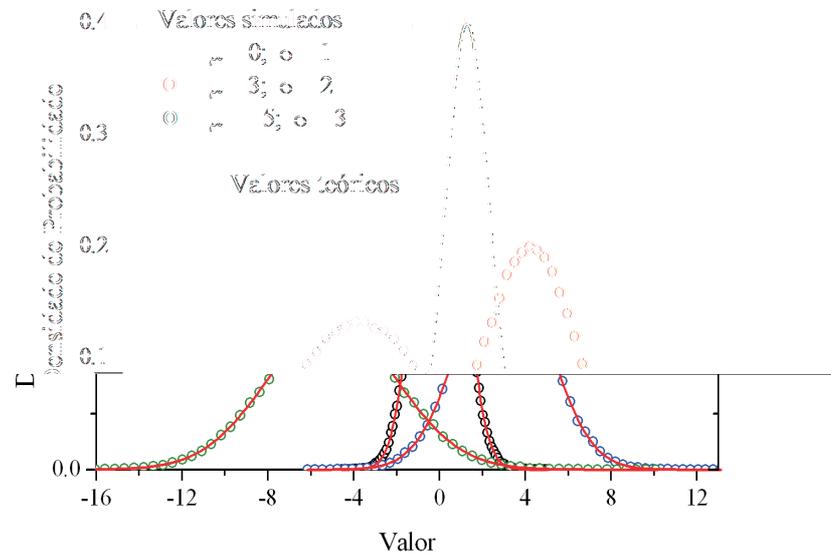


Figura 3.3: Distribuição do ruído simulado com diferentes médias e desvios-padrão. Em vermelho, a *pdf* Gaussiana.

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (3.1)$$

cuja integração de porções de $f(x)$ nos darão probabilidades.

3.1.2 SNR

A relação sinal-ruído (SNR - *Signal to Noise Ratio*), medida em *decibel* (dB) quantifica a relação entre a característica que se quer observar (sinal) e as variabilidades provenientes de outras fontes (ruído). Quanto maior o valor dessa relação, menor é a interferência associada ao ruído. Uma das formas de definir tal relação é através da razão entre as variâncias do sinal e do ruído em escala logarítmica, isto é,

$$SNR(dB) = 10 \log \left(\frac{\sigma_S^2}{\sigma_R^2} \right) = 20 \log \left(\frac{\sigma_S}{\sigma_R} \right), \quad (3.2)$$

em que σ_S é o desvio padrão do sinal e σ_R o desvio padrão do ruído.

Nas simulações realizadas, chamamos de "sinal" o modelo de resposta hemodinâmica descrito na seção 2.3. Para simular uma série temporal (sinal + ruído) para uma dada *SNR*, usamos apenas o desvio padrão do ruído, uma vez que o desvio padrão do sinal é fixo. Obtemos o valor do desvio padrão do ruído pela equação

$$\sigma_R = \frac{\sigma_S}{10^{\frac{SNR}{20}}}. \quad (3.3)$$

Os valores de ruído (ξ) são portanto, extraídos de uma distribuição gaussiana com média zero e desvio padrão σ_R . A série temporal $s(t)$ é composta de uma função resposta hemodinâmica $h(t)$ e do ruído gaussiano $\xi(t)$,

$$s(t) = h(t) + \xi(t). \quad (3.4)$$

Uma época corresponde a um período de estímulo e repouso. Para experimentos com várias épocas, a função $h(t)$ é repetida várias vezes seqüencialmente. A fig.3.4 apresenta séries temporais com 10 épocas para alguns valores de SNR.

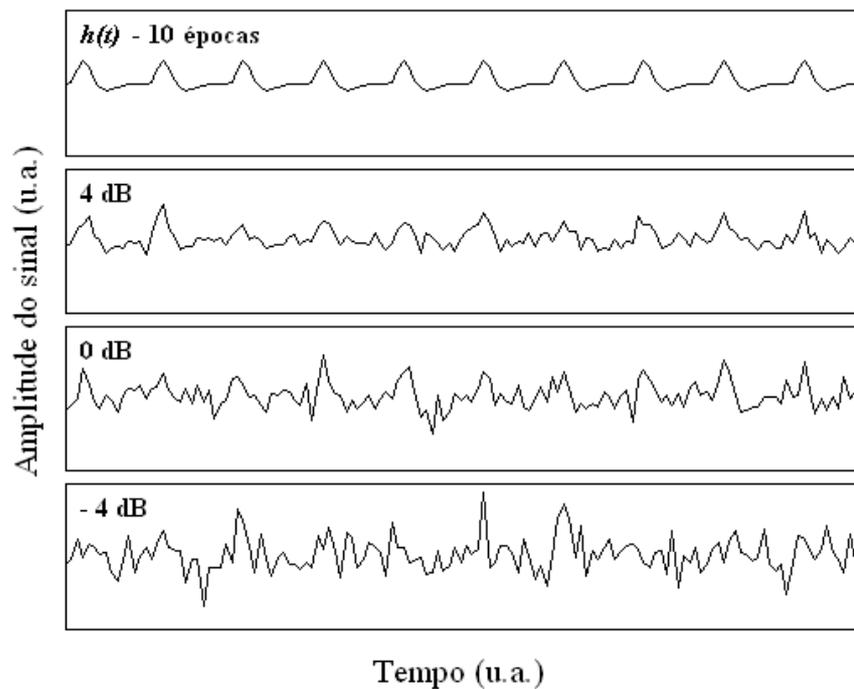


Figura 3.4: Séries temporais simuladas para diferentes valores de SNR. Para valores altos de SNR, a HRF é mais evidente.

As séries temporais são simuladas para analisar os métodos de análise em fMRI apresentados neste trabalho. A partir de uma série temporal (experimento), calculamos uma dada estatística x . Esta estatística pode ser a média amostral de uma quantidade, como, a entropia de Kullback-Leibler por exemplo(ver seção 5.3), calculada sobre várias épocas. Ao repetir muitos experimentos, podemos construir as distribuições de densidade de probabilidade de x . Estas distribuições são utilizadas nos testes de hipóteses para inferir a presença ou ausência de ativação neural. A seção seguinte mostra como os testes de hipóteses são construídos.

3.2 Testes de hipóteses

Em análises de sinais de fMRI, dada uma série temporal correspondente a um *voxel*, queremos identificar características desta série que sugiram a presença de ativação neural. Desta forma, podemos construir um teste de hipóteses com a hipótese nula (H_0) de que a série temporal apresenta apenas ruído e a hipótese alternativa (H_1) de que a série apresenta sinal acrescido de ruído.

H_0 : ruído (*voxel* inativo)

H_1 : sinal + ruído (*voxel* ativo)

A estatística x_o , observada para uma série temporal de um *voxel*, é então usada para testar a hipótese nula. Se o valor de x_o estiver na região de rejeição de H_0 , consideramos o *voxel* ativo (+). Caso contrário (x_o fora da região rejeição de H_0), consideramos o *voxel* inativo (j). Fixado um nível de significância α , delimita-se a região de rejeição ao longo de x , sob a curva densidade de probabilidade.

Para obter as distribuições de densidade de probabilidade, realizamos experimentos computacionais. Tais experimentos consistem em aplicar um determinado método de análise às séries temporais e, para cada série temporal simulada, associar um valor da estatística x .

Ao realizarmos muitos experimentos, podemos alocar os valores de x calculados em intervalos consecutivos, obtendo a distribuição de frequência desses valores. Ao dividir a frequência pelo número de experimentos e pela largura do intervalo, obtemos a distribuição de densidade de frequência. Ao fazermos a largura do intervalo suficientemente pequena, podemos associar à esta distribuição uma função contínua chamada densidade de probabilidade. Podemos calcular a probabilidade de ocorrer valores dentro de um intervalo através da área abaixo da curva da distribuição, limitada pelos valores desse intervalo. A fig.3.5, mostra uma distribuição densidade de probabilidade como função de determinada estatística (x). A área sombreada é a probabilidade de ocorrer um valor dentro do intervalo ($j_2 < x < 1$).

As distribuições densidade de probabilidade, além de serem necessárias para a construção das curvas ROC (*Receiver operating characteristic*), também são usadas para obter o valor de nível descritivo (p) na análise de dados reais. O valor de p associado a um *voxel* da imagem, indica a probabilidade de ocorrer eventos tão ou mais extremos do que o observado considerando a hipótese nula verdadeira (ou seja, considerando que o *voxel* apresenta só ruído). Um valor de $p = 0,05$, significa que temos 5% de chance de obter o valor observado ou mais extremo. Na fig. 3.6 mostramos o valor de nível descritivo (p) para um valor observado (x_o).

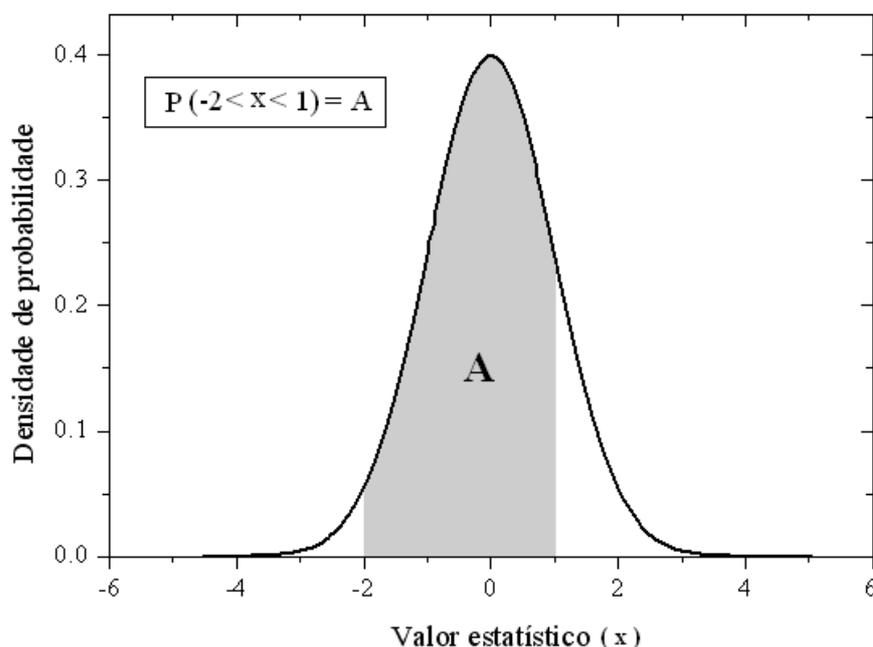


Figura 3.5: Densidade de probabilidade da estatística x . O valor da área sombreada (A) é a probabilidade de x estar entre -2 e 1 .

Para obtermos a distribuição de x sob H_0 , simulamos séries temporais contendo apenas ruído e , em seguida, calculamos a estatística x para cada série simulada. De forma análoga, obtemos a distribuição de x sob H_1 ao simularmos séries contendo sinal acrescido de ruído (dada uma SNR), da estatística x supondo H_1 verdadeira. Para cada valor de SNR simulado obtemos uma distribuição de x (sob H_1) diferente.

A região de rejeição é delimitada pelo valor de corte (x^{α}). O resultado do teste é classificado como sinal (+) se a estatística x pertencer à região de rejeição. Caso contrário, é classificado como ruído (j). A sensibilidade é definida como a probabilidade $P(+|+)$ de verdadeiros positivos, ou seja, rejeitar a hipótese nula dada que ela é falsa. A especificidade é dada pela probabilidade $P(j|j)$ de verdadeiros negativos, ou seja, a probabilidade de aceitar a hipótese nula dada que ela é verdadeira, então:

$$P(+|+) = \int_{x^{\alpha}}^{\infty} f(x|+) dx \quad (3.5)$$

e

$$P(j|j) = \int_{-\infty}^{x^{\alpha}} f(x|j) dx, \quad (3.6)$$

em que $f(x|+)$ e $f(x|j)$ são as densidades de probabilidade obtidas de simulações com sinal (para uma dada SNR) e com ruído ($SNR \rightarrow \infty$), respectivamente.

O complementar da especificidade ($1 - \text{especificidade}$) é o nível de significância α . Seu

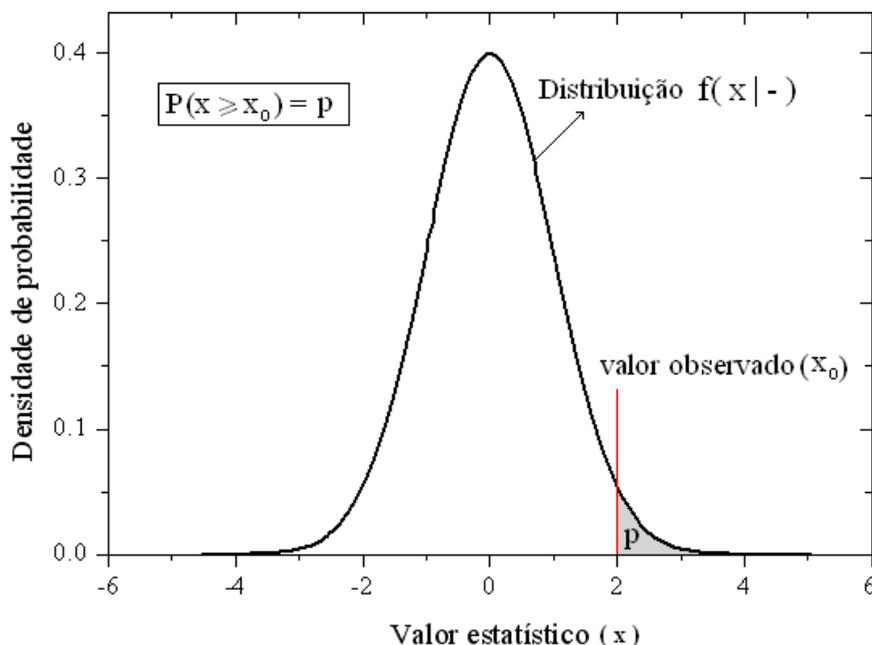


Figura 3.6: Valor do nível descritivo p , obtido a partir da função densidade de probabilidade de x , supondo (H_0) verdadeira.

valor é a probabilidade de falsos positivos $P(+j|j^-)$, ou seja, concluir que é sinal dado que é ruído (erro do tipo II). Analogamente, o complementar da sensibilidade é β , que é a probabilidade de falsos negativos $P(j^-|j^+)$, concluir que é ruído dado que é sinal (erro tipo II). A fig. 3.7, mostra a especificidade, sensibilidade, nível de significância e β para uma estatística x , dados um valor crítico x^α e uma SNR.

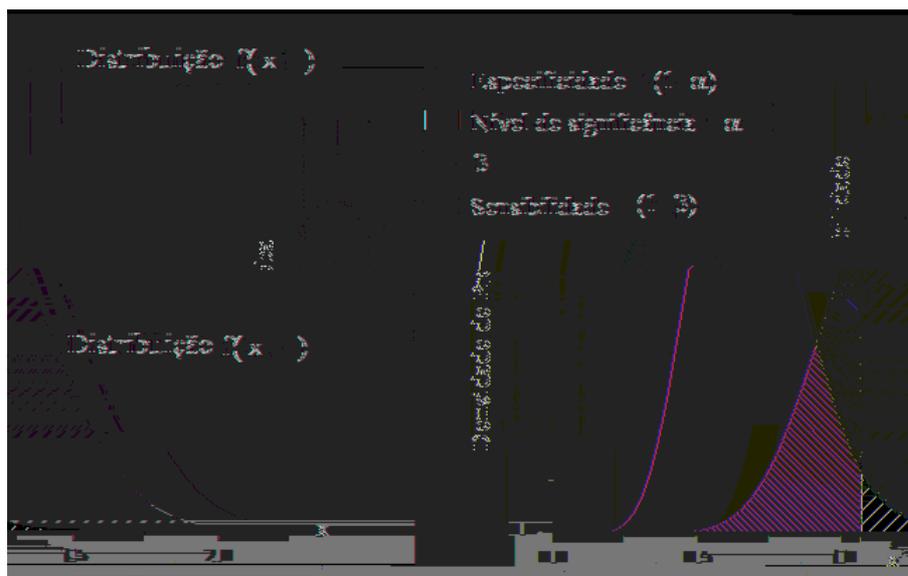


Figura 3.7: Valores de sensibilidade, especificidade, nível de significância e β dados um valor crítico x^α e uma SNR.

Ao aumentarmos o valor de x^α , aumentamos a especificidade, mas, por outro lado, dimi-

nuímos a sensibilidade. Assim, quando queremos diminuir o erro do tipo I (falsos positivos) estamos aumentando o erro de tipo II (falsos negativos). Portanto, cabe ao experimentador estabelecer qual tipo de erro quer minimizar. Na fig. 3.8 são mostrados os tipos de erros associados à sensibilidade e especificidade.

	H ₀ Verdadeira	H ₀ Falsa
Aceita H ₀	DECISÃO CORRETA Probabilidade: $1-\alpha$ <i>ESPECIFICIDADE</i>	ERRO TIPO II Probabilidade: β
Rejeita H ₀	ERRO TIPO I Probabilidade: α <i>NÍVEL DE SIGNIFICÂNCIA</i>	DECISÃO CORRETA Probabilidade: $1-\beta$ <i>SENSIBILIDADE</i>

Figura 3.8: Tipos de erros, sensibilidade e especificidade, associados aos testes de hipóteses.

3.3 Curvas ROC

A curva ROC [24, 25] é o gráfico da sensibilidade (probabilidade de um verdadeiro positivo) *versus* ($1 - \text{especificidade}$) (probabilidade de um falso positivo), resultante da variação do limiar x^r . A qualidade do método empregado na análise é avaliada por sua habilidade em detectar áreas cerebrais ativas com alta sensibilidade e especificidade. Portanto, quanto maior for a concentração de pontos no canto superior esquerdo do gráfico, melhor é a qualidade do método. A área sob a curva ROC é uma forma de quantificar este desempenho. O poder do teste aumenta conforme a área ROC varia de 0,5 (menor valor) até 1 (ótimo desempenho). Nossas simulações mostram que a área ROC (e, portanto, o poder do teste) aumenta conforme cresce o valor de SNR. A fig. 3.9 mostra algumas curvas ROC e suas respectivas áreas. A tabela 3.1 associa a qualidade do método a faixas de valores da área ROC.

Área abaixo da curva ROC	Qualidade
0.5 ! 0.6	Péssima
0.6 ! 0.7	Ruim
0.7 ! 0.8	Regular
0.8 ! 0.9	Boa
0.9 ! 1.0	Excelente

Tabela 3.1: Valores da área abaixo da curva ROC associadas às respectivas qualidades.

A sensibilidade varia de 0 a 1 e está relacionada à probabilidade de rejeitar a hipótese

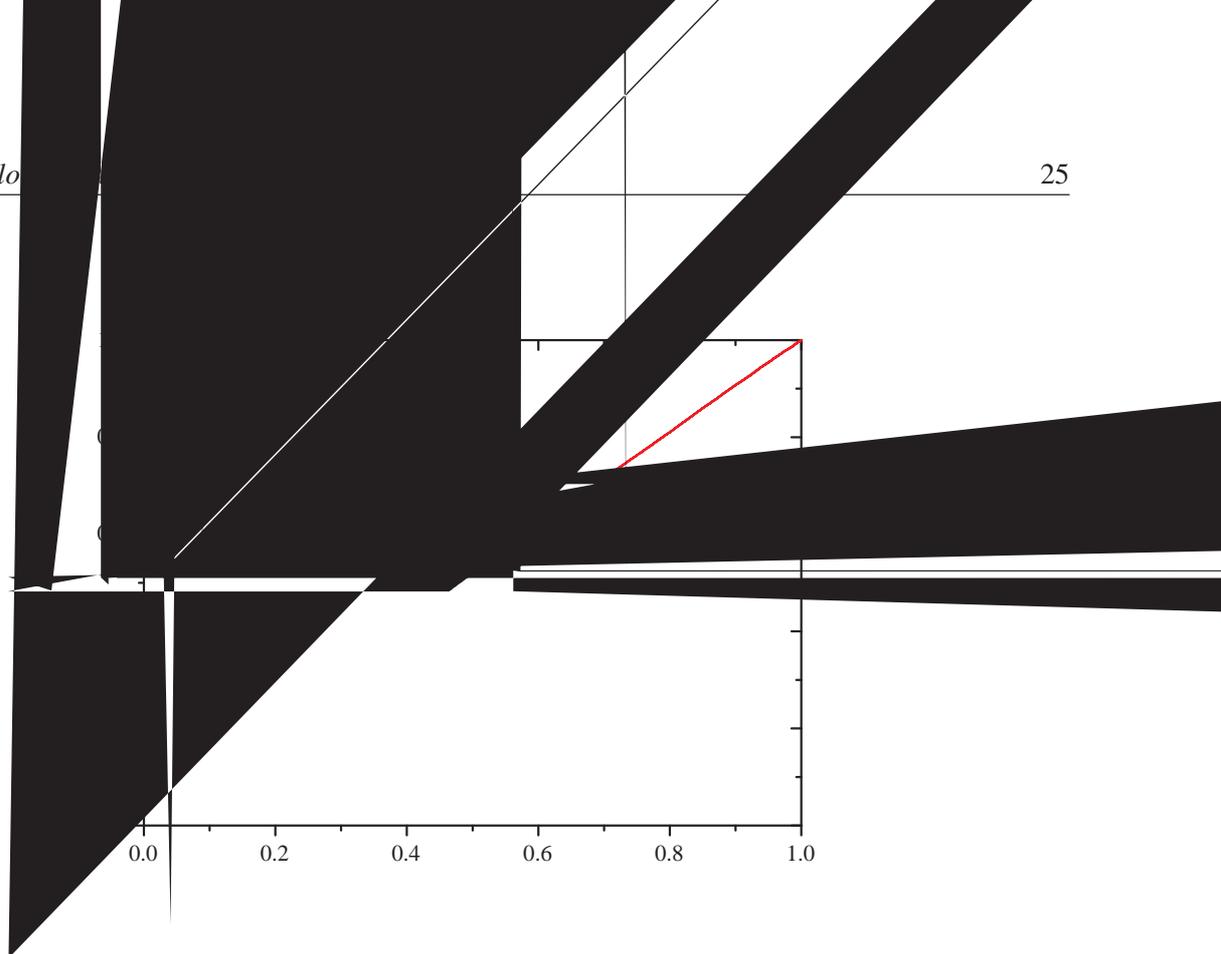


Figura 3.9: Curvas ROC e as áreas associadas a cada uma delas.

nula (H_0), dada que ela é falsa, ou seja, a fração esperada de verdadeiros negativos (FVN). A especificidade, também varia de 0 a 1 e está relacionada à probabilidade de aceitar a hipótese nula, dada que ela é verdadeira, ou seja, a fração esperada de verdadeiros positivos (FVP). Especificamente em fMRI, a hipótese nula é de que temos ruído e a hipótese alternativa (H_1), sinal (ativação neural).

3.4 Protocolo de aquisição de dados reais

As simulações implementadas neste trabalho foram realizadas com base nos dados reais disponíveis para análise. Esses dados foram obtidos a partir de um paradigma de evento relacionado em que o sujeito movimentava ligeiramente os dedos das duas mãos (finger-tapping) durante 3s e em seguida ficava em repouso por 20s. Três fatias axiais posicionadas sobre porções do córtex motor primário foram adquiridas. Em cada época, que compreende um período de estímulo e repouso, foram adquiridas 14 imagens (3 durante o estímulo e 11 no repouso), totalizando 336 imagens por fatia. Os parâmetros usados na obtenção das imagens foram, TR: 1.68s; TE: 118ms; Matriz: 128x128; FOV: 210mm; Dimensão do *voxel*: 1.64 x 1.64 x 4.00mm.

3.5 Apresentação dos mapas estatísticos

Depois de calcular o valor estatístico correspondente a cada *voxel*, associamos esse valor a um nível descritivo (p). Se o valor de p é menor que o nível de significância estabelecido (α), então o *voxel* é considerado ativo e é representado na imagem com uma intensidade proporcional ao valor estatístico obtido. Caso contrário, recebe uma coloração fora da escala de ativação, de forma que o diferencie dos ativos.

As "escalas de fogo" são as mais usadas para representar ativação. Nesta escala, os menores valores da estatística são representados em vermelho, os intermediários em amarelo e os maiores em branco cf. fig. 3.10. Quanto mais significativo, mais claro é o *voxel*.

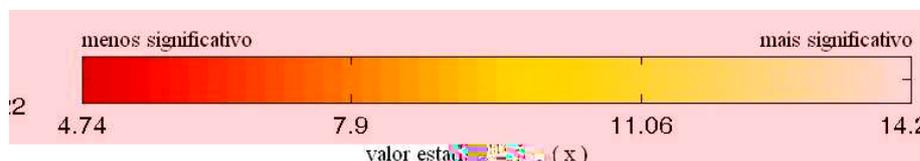


Figura 3.10: Escala de cores para representar ativação.

Os mapas estatísticos são sobrepostos à imagens anatômicas de alta resolução com contraste ponderado em T_1 . Porém, neste trabalho, utilizamos as próprias imagens EPI obtidas para análise da série temporal, usando-as apenas como guia, uma vez que a localização em detalhe da ativação não foi nosso objeto de estudo. De acordo com as normas radiológicas, apresentaremos os mapas de forma que o lado esquerdo da imagem anatômica representa o lado direito do indivíduo e vice versa.

Ao realizarmos muitos testes estatísticos, esperamos a ocorrência de muitos erros em termos absolutos. Para uma imagem com resolução de 128×128 *voxels*, temos 16384 *voxels*, assim para $\alpha = 0,05$ esperamos que em média 5% desse valores sejam falsos positivos, ou seja, 820 erros do tipo I. Para minimizar esse tipo de erro, alguns recursos são utilizados.

3.5.1 Correção de Šidák - Bonferonni

O nível de significância (α) usado na análise de dados reais é fixo para cada teste (*voxel*), independente do número de testes realizados. Chamaremos este valor de *alfa por teste* (α_t). A correção proposta por Šidák e Bonferonni [26], minimiza os erros do tipo I, através de um nível de significância (α_f) que leva em conta o número de testes realizados (família de testes).

A probabilidade de não cometer nenhum erro do tipo I em C testes considerados independentes é dado por $(1 - \alpha_t)^C$. Logo a probabilidade de cometer pelo menos 1 erro do tipo I é

dado pelo complementar: $1 - (1 - \alpha)^C$. Assim para um valor de $(\alpha_t = 0,05)$ em $C = 100$ testes, temos 99,4% de chance de cometer pelo menos um erro do tipo I. O valor de α_f , é exatamente o valor da probabilidade de cometer pelo menos um erro do tipo I, isto é,

$$\alpha_f = 1 - (1 - \alpha_t)^C. \quad (3.7)$$

Neste trabalho, utilizamos o nível de significância por família de testes e, a partir de seu valor, calculamos o nível de significância a ser usado por cada teste, de acordo com a equação

$$\alpha_t = 1 - (1 - \alpha_f)^{1/C}. \quad (3.8)$$

Utilizando esse mecanismo, diminuimos de forma controlada o número de falsos positivos. Entretanto, é importante ter em mente que estamos também aumentando os erros do tipo II, ou seja, os falsos negativos, uma vez que fica mais difícil classificar um *voxel* como ativo.

3.6 Equipamentos e aplicativos utilizados

No processo que parte da aquisição das imagens até o processamento dos dados, são utilizados muitos equipamentos e aplicativos. Podemos dividir esse processo em três etapas:

3.6.1 Aquisição das imagens

As imagens foram adquiridas em um *scanner* de 1.5T (*Siemens, Magnetom Vision*)¹ com bobina de quadratura transmissora/receptora de cabeça e polarização circular.

3.6.2 Pré-processamento

O pré-processamento dos dados foram feitos pelo *software* Brain Voyager *QXTM* (Versão 1.3.8; Brain Innovation, Maastricht, Holanda). Foram realizadas a correção de movimento, correção do tempo entre fatias, suavização espacial (filtro gaussiano com FWHM² de 4 mm) e filtragem temporal (filtro passa alta de 3 ciclos/s).

¹Equipamento em atividade regular no HCFMRP - Serviço de radioterapia.

²Sigla em inglês para Full Width at Half Maximum

3.6.3 Processamento

- ² Aplicativo - *MATLAB*[®], usado na análise de dados reais, construção de gráficos e obtenção dos mapas de ativação.
- ² Aplicativo - *FORTTRAN*, usado para simular séries temporais e calcular a estatísticas dos métodos implementados a partir das séries simuladas.
- ² Aplicativo - *Mathematica*[®], usado para auxiliar na solução de equações analíticas.
- ² Equipamentos - *Cluster* de 20 computadores com velocidade de relógio superior a 1GHz cada um, usados para executar as simulações.

4 *Métodos Clássicos em fMRI*

Os dados experimentais de fMRI geralmente possuem baixa SNR. Dessa forma, os testes de significância realizados devem avaliar se as séries temporais apresentam alguma componente decorrente da manipulação experimental (estímulo).

O teste t de Student é apropriado quando se deseja avaliar se a série temporal de um *voxel* possui diferentes valores médios, em intervalos de tempo consecutivos, dadas diferentes condições experimentais. Por outro lado, a correlação será indicada se houverem previsões específicas sobre a forma da HRF.

Neste capítulo, apresentaremos métodos estatísticos simples (teste t e correlação), para detectar ativação neural. Realizamos simulações para verificar o poder discriminante de cada teste através das curvas ROC. Em seguida, utilizamos os testes para analisar dados reais.

4.1 O Teste t

A estatística t foi primeiramente proposta por William Sealy Gosset para monitorar a qualidade no preparo de cerveja. O pseudônimo *Student* foi utilizado por Gosset para não ter problemas com a divulgação de dados confidenciais da empresa em que trabalhava, a cervejaria *Guinness* na Irlanda.

A partir de duas amostras independentes, o teste t pode ser usado para testar se a média das respectivas populações são diferentes. Em imagens de fMRI, temos para cada *voxel* uma série temporal que compreende períodos de estímulo e de repouso intercalados. Ao compararmos o subconjunto de pontos referente ao estímulo com o de repouso, podemos concluir que há sinal se existir uma diferença significativa entre esses subconjuntos de pontos. Ao aplicarmos o teste t de Student, podemos inferir que a série temporal apresenta características da ativação neural se o valor da estatística t for grande o suficiente para rejeitarmos a hipótese nula de que as médias populacionais μ_x e μ_y são iguais ($\mu_x = \mu_y$), dado um nível de significância (α) e o número de graus de liberdade (v). O número de graus de liberdade é dado pela soma do número de valores

de cada amostra subtraído do número de amostras. Como exemplo, suponhamos duas amostras: a primeira proveniente da população X com n_x elementos, média \bar{x} e desvio padrão S_x ; a segunda proveniente da população Y com n_y elementos, média \bar{y} e desvio padrão S_y . Se $n_x = n_y = 10$, temos neste caso $v = 18$ graus de liberdade. Para as duas amostras, o valor da estatística t é calculado por:

$$t = \frac{(\bar{x} - \bar{y})}{S_{xy}}, \quad (4.1)$$

em que

$$S_{xy} = \sqrt{\left(\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}\right)}, \quad (4.2)$$

e a variância amostral combinada (*pooled variance* - S_p^2) é

$$S_p^2 = \frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}, \quad (4.3)$$

com as variâncias das amostras S_x^2 e S_y^2 ,

$$S_x^2 = \frac{\sum_{i=1}^{n_x} (x_i - \bar{x})^2}{(n_x - 1)}, \quad (4.4)$$

$$S_y^2 = \frac{\sum_{i=1}^{n_y} (y_i - \bar{y})^2}{(n_y - 1)}. \quad (4.5)$$

Se a hipótese nula $H_0 : \mu_x = \mu_y$ for verdadeira, então,

$$t = t^0 = \frac{\bar{x} - \bar{y}}{S_{xy}}. \quad (4.6)$$

Ao considerarmos a hipótese nula como verdadeira, a distribuição da variável t , descreve a diferença esperada entre duas médias amostrais retiradas de uma mesma população. A fig. 4.1, mostra algumas distribuições da variável t para alguns valores de graus de liberdade.

Podemos observar que a distribuição t é muito semelhante à distribuição normal. À medida que aumentamos o número de graus de liberdade, a distribuição t se aproxima da distribuição normal com média ($\mu = 0$) e desvio padrão ($\sigma = 1$). Tal distribuição é conhecida como *distribuição normal padrão* e é um caso particular da distribuição t quando o número de graus de liberdade tende ao infinito. Para propósitos práticos, os valores da distribuição t aproximam-se

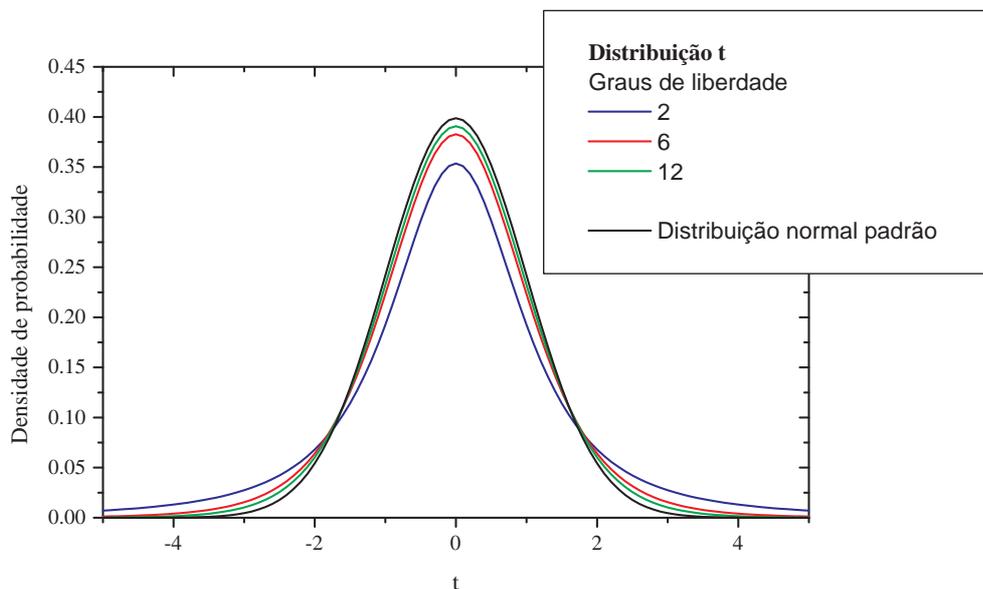


Figura 4.1: Distribuições da variável t para diferentes valores de graus de liberdade (v) e distribuição normal padrão.

dos valores da distribuição normal padrão relativamente depressa, de forma que quando $v = 30$, as distribuições são quase idênticas.

Em testes de hipóteses, o que se busca é basicamente aceitar ou rejeitar a hipótese nula (H_0) com base no valor da estatística obtida e o nível de significância (α). O valor de α indica a probabilidade de rejeitar H_0 dado que ela é verdadeira, ou seja, probabilidade de ocorrer um erro do *tipo I*. Assim, ao obtermos a estatística t , associamos a ela um valor de nível descritivo (p), definido como a probabilidade de ocorrer eventos tão ou mais extremos do que o observado, considerando que H_0 é verdadeira. Assim, um nível descritivo baixo (menor que α), indica que seria pouco provável observar determinado resultado se H_0 fosse verdadeira.

Há dois tipos de testes estatísticos, o unilateral e o bilateral. No unilateral, a hipótese alternativa (H_1) pode ser dada por: $\mu_x > \mu_y$ ou $\mu_x < \mu_y$. Devemos nesse caso considerar apenas um extremo da distribuição t , ao obtermos o valor p . No teste bilateral, a hipótese alternativa (H_1) considera $\mu_x \neq \mu_y$, neste caso consideramos os dois extremos da distribuição.

Definido o tipo de teste (unilateral ou bilateral), o valor de p pode ser obtido a partir de tabelas encontradas na literatura.

O teste t pode, portanto, ser usado para avaliar se as duas amostras correspondentes aos períodos de estímulo e ao repouso, são provenientes de uma mesma população. Em caso positivo o *voxel* é considerado inativo, ou seja, o estímulo não provocou modificações na intensidade da

série temporal. Caso contrário, consideramos o *voxel* ativo, a série temporal correspondentes ao estímulo e ao repouso são diferentes.

A seguir, mostraremos a avaliação do teste t através de simulações e em seguida o teste aplicado a dados reais.

4.1.1 Teste t - Simulações

Para avaliar o teste t, realizamos simulações conforme descrito na seção 3.1. Séries temporais, contendo 24 épocas de 14 pontos cada foram utilizadas. Associamos o subconjunto de pontos $X = \{x_i; i = 1, 2, \dots, n_x\}$ ao estímulo e $Y = \{y_i; i = 1, 2, \dots, n_y\}$ ao repouso. Com $n_x = n_y = 168$, todos os valores associados à ativação e $n_y = 168$, todos os valores associados ao repouso.

Nas simulações de HRF com ruído, a estatística t^θ avaliada é, de acordo com (4.1) e (4.6), dada por:

$$t^\theta = \frac{\bar{x} - \bar{y}}{S_{xy}} = t + \frac{(\mu_x - \mu_y)}{S_{xy}}. \quad (4.7)$$

Pela característica da HRF simulada, a média do sinal relativo ao estímulo é maior que a média dos valores do repouso ($\mu_x > \mu_y$). Portanto há um deslocamento à direita da distribuição t . Esse deslocamento será maior quanto maior a SNR, uma vez que a diferença entre os dados relativos ao estímulo e ao repouso, se tornam mais evidentes. Essa característica da HRF nos permite utilizar um teste de hipótese unilateral à direita, na análise de dados reais que será apresentada na seção seguinte. A fig.4.2, mostra a distribuição da variável t^θ para alguns valores de SNR.

A distribuição de t^θ obtida para ruído simulado, deve reproduzir a distribuição t de Student, uma vez que para o ruído, as médias μ_x e μ_y são iguais. A fig. 4.3 mostra a distribuição da variável t^θ calculada para ruído simulado com os parâmetros ($n_x = n_y = 10$) e a linha cheia representa a distribuição t com número de graus de liberdade ($v = 18$).

A medida que aumentamos a SNR, os valores de t^θ ficam mais distantes da distribuição do ruído, pois a quantidade $(\mu_x - \mu_y)/S_{xy}$ aumenta. Para avaliar o quão distante as distribuições de t^θ (para uma dada SNR) estão da distribuição de t^θ correspondente a ruído somente, construímos as curvas ROC (fig. 4.4).

As curvas ROC mostram que a medida que aumentamos a SNR, a especificidade e a sensibilidade aumentam. Mesmo para um valor muito baixo de SNR como (-8dB), o método se mostra eficiente em apontar a presença da HRF.

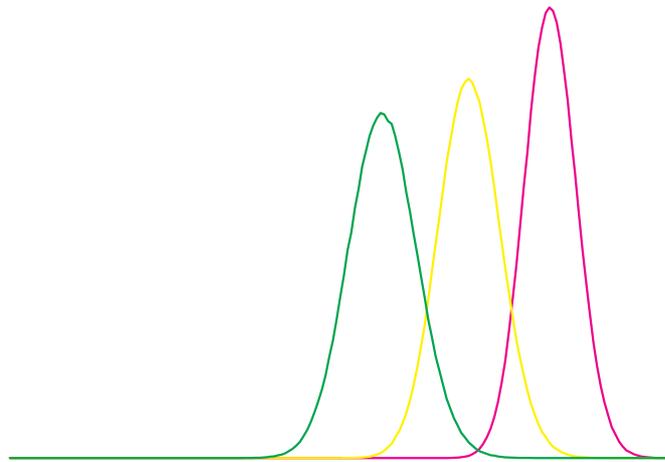


Figura 4.2: Distribuição da estatística t^{θ} para diferentes valores de SNR e para ruído simulado.

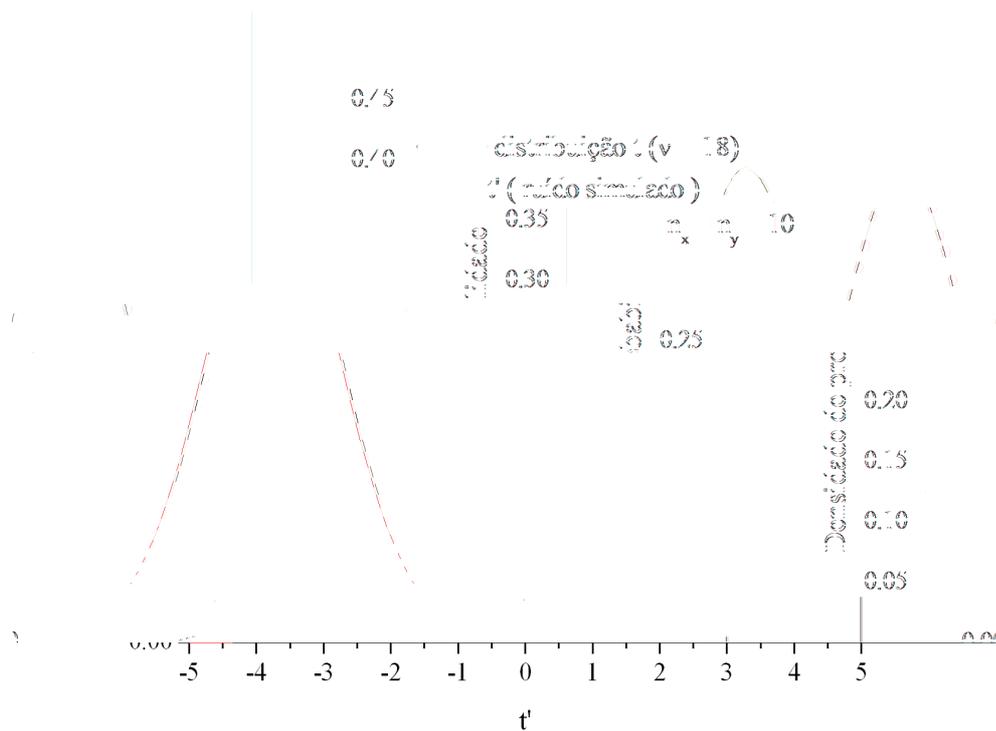


Figura 4.3: Distribuição t com número de graus de liberdade ($v = 18$) sobreposta a distribuição da variável t' calculada para ruído simulado com os parâmetros ($n_x = n_y = 10$).

4.1.2 Teste t - Dados reais

Utilizamos a estatística t para obter os mapas de ativação a partir de dados reais. Para cada série temporal de cada *voxel*, calculamos o valor de t , utilizando os subconjuntos de pontos

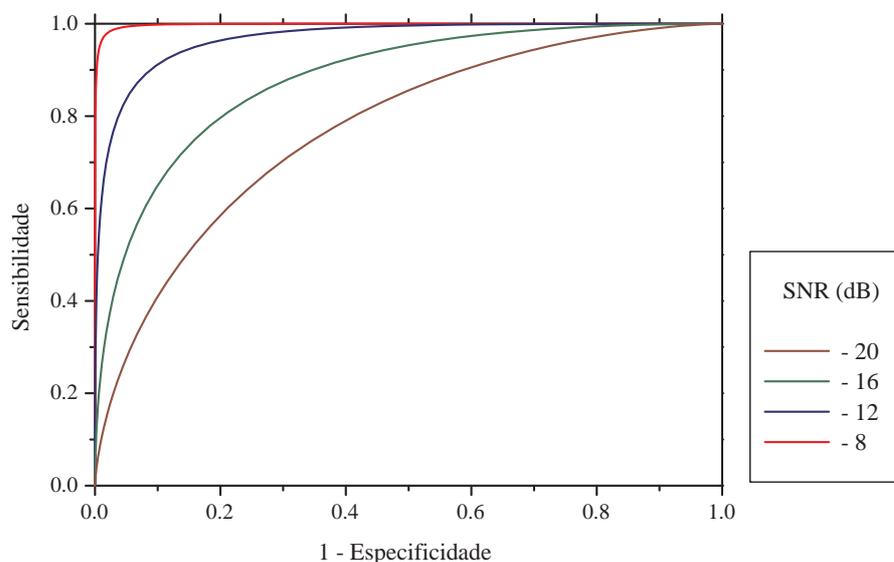


Figura 4.4: Curvas ROC para a estatística t^l para diferentes valores de SNR do sinal simulado. Note que os valores de SNR maiores que -8 dB não estão representados uma vez que as curvas para tais valores tocam as extremidades dos eixos.

associados ao estímulo e ao repouso. Como cada série temporal contém 336 pontos (24 épocas, cada uma com 7 pontos relativos ao estímulo e 7 ao repouso), temos 334 graus de liberdade. A partir do valor de t obtido em cada *voxel*, obtivemos o nível descritivo p , a partir da distribuição t com 334 graus de liberdade. Como já discutido anteriormente, adotamos um teste de hipótese unilateral. Os *voxels* com valor de nível descritivo (p) menor que o nível de significância (corrigido para múltiplos testes, $\alpha_f = 0,05$), estão representados em "escala de fogo" sobrepostos a uma imagem EPI.

Os mapas apresentam ativação em regiões do córtex motor primário e secundário, de acordo com estímulo aplicado.

É importante destacar que para a obter os mapas, não foi necessário o uso do modelo da HRF. A única suposição que fizemos foi de que, os valores da série temporal correspondente à ativação são maiores que os valores correspondentes ao repouso (por isso o uso de um teste unilateral). O teste bilateral pode ser feito, mas temos que ter em mente que ao fazer isso, estamos considerando que a ativação neural pode também causar uma diminuição no sinal. Isso gera uma discussão acerca dos efeitos fisiológicos que levam a essa diminuição, o que está além dos objetivos deste trabalho.

A partir dos resultados da simulação e dos mapas de ativação, verificamos que o teste t apesar de simples, é efetivo para análise de sinais de fMRI. Porém, alguns aspectos devem ser

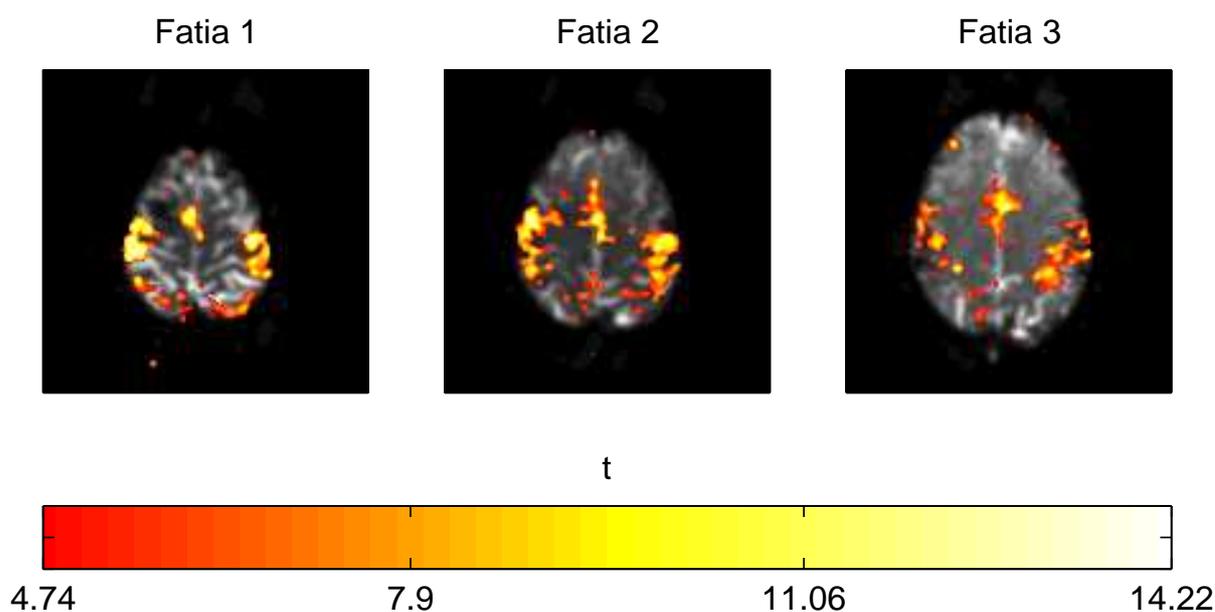


Figura 4.5: Mapa de ativação obtido através do teste t. Nível de significância: $\alpha_f = 0,05$; $\alpha_t = 3,13 \times 10^{-6}$, valor t de corte: $t_c = 4,74$.

ponderados ao implementar tal método:

1. a decisão de quais pontos correspondem a uma condição experimental (estímulo) e a outra (repouso) é essencial para o sucesso do teste;
2. qualquer artefato que cause uma mudança sistemática entre as condições experimentais, o movimento da cabeça por exemplo, pode comprometer o teste;
3. caso o objetivo seja medir se há diferença entre duas condições experimentais, mas cuja série temporal correspondente a uma condição tenha média muito próxima da correspondente a outra condição, o teste t pode não ser o mais indicado.

Outro teste clássico que pode ser usado na análise de sinais de fMRI é o teste da correlação. Ao contrário do teste t, a correlação usa um modelo de HRF para analisar dados reais.

4.2 Correlação

Em teoria de probabilidade e estatística, a correlação mede o quanto duas variáveis aleatórias estão linearmente relacionadas. O *produto de Pearson*, é obtido dividindo-se a covariância das duas variáveis pelo produto dos seus desvios-padrão. Uma covariância positiva significa que à medida que os valores dos dados experimentais aumentam (diminuem), os dados previstos também aumentam (diminuem) da mesma forma. Para covariância negativa ocorre o

contrário, à medida que os valores dos dados experimentais aumentam, os valores dos dados previstos diminuem e vice-versa. A divisão da covariância pelo produto dos desvios-padrão tem como finalidade normalizar os valores do produto de Pearson entre -1 e +1.

Quando calculado com base em uma amostra, o coeficiente de correlação é designado pela letra "r", conhecido como "r de Pearson". O valor de r recebe as seguintes interpretações:

- ² $r = -1$) há uma relação linear negativa perfeita entre as variáveis;
- ² $r = 0$) não há relação alguma entre as variáveis;
- ² $r = +1$) há uma relação linear positiva perfeita entre as variáveis.

Ao analisarmos os dados de fMRI, podemos usar a correlação para quantificar a correspondência entre a série temporal observada com um modelo de resposta esperada (HRF), e inferir que quanto maior o valor de r calculado para um voxel, mais provável aquele voxel estar ativo. Os primeiros estudos de fMRI usando correlação foram feitos por Bandettini e colegas em 1993 [27].

Considerando o conjunto de pontos da uma série temporal de um voxel: $D = fd_i; i = 1...ng$ e o conjunto de pontos do modelo: $M = fm_i; i = 1...ng$, o coeficiente de correlação de Pearson é calculado por:

$$r = \frac{1}{(n_j - 1)} \sum_{i=1}^n \frac{(d_i - \bar{d})(m_i - \bar{m})}{S_d S_m}, \tag{4.8}$$

sendo os desvios-padrão S_d e S_m , dados por:

$$S_d = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{(n_j - 1)}} \tag{4.9}$$

e

$$S_m = \sqrt{\frac{\sum_{i=1}^n (m_i - \bar{m})^2}{(n_j - 1)}}. \tag{4.10}$$

Assim, a expressão para o cálculo de r pode ser reescrita como:

$$r = \frac{\sum_{i=1}^n (d_i - \bar{d})(m_i - \bar{m})}{S_d S_m}$$

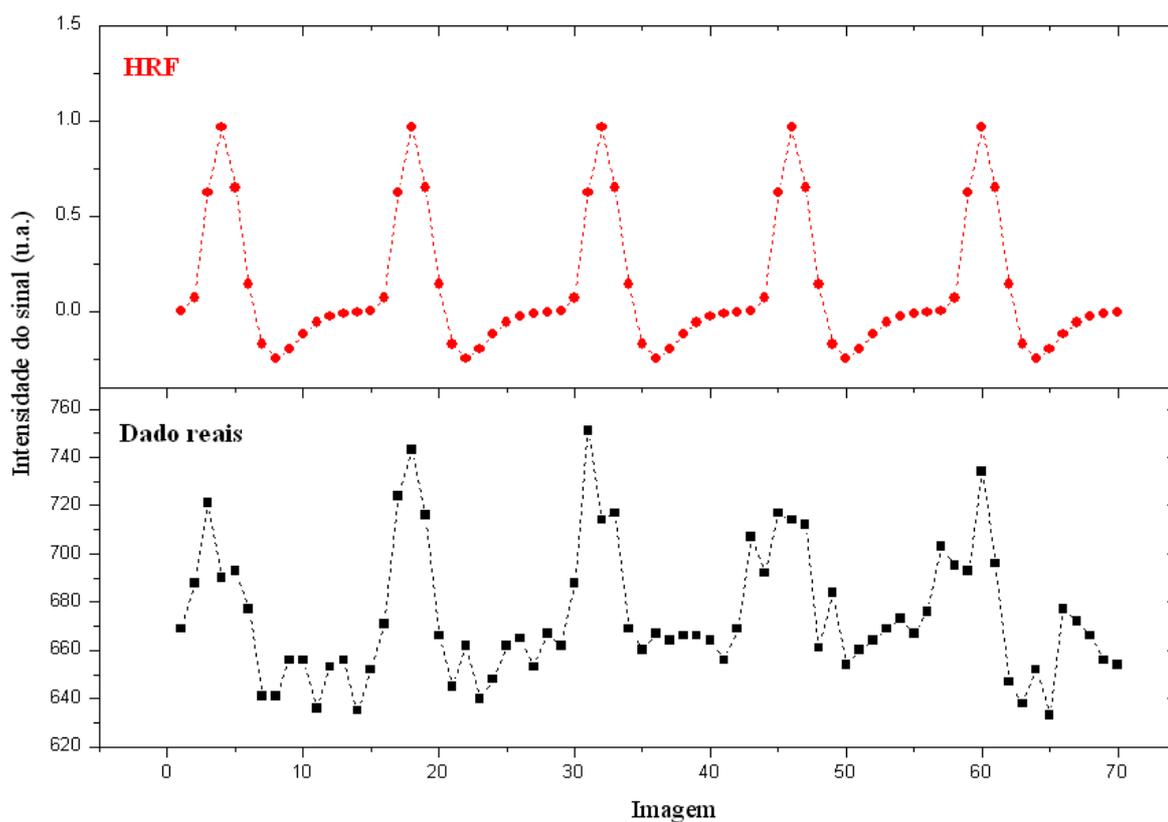


Figura 4.6: Cinco épocas de uma série temporal de um *voxel* em comparação com o modelo de resposta hemodinâmica. A correlação quantifica a "semelhança" entre as duas.

entre os dados é clara, tratando-se possivelmente de um *voxel* ativo.

Simulamos série temporais para verificar a eficiência da estatística r em análise de sinais de fMRI. Os resultados obtidos serão apresentados a seguir.

4.2.1 Correlação - Simulações

Calculamos o valor de r para séries temporais simuladas. É importante lembrar que o modelo de HRF usado para simular as séries temporais é o mesmo que o utilizado para calcular a correlação. Assim, para valores altos de SNR, teremos a correlação da HRF com mesma função, fazendo com que o valor de r fique muito próximo de 1. A fig. 4.7, mostra as distribuições da estatística r para alguns valores de SNR e para o ruído.

Podemos perceber que a medida que aumentamos a SNR, as distribuições se distanciam da distribuição do ruído, chegando a uma separação completa para valores por volta de -8dB. As curvas ROC da fig. 4.8 mostram que a partir desse valor, a sensibilidade e a especificidade são máximas.

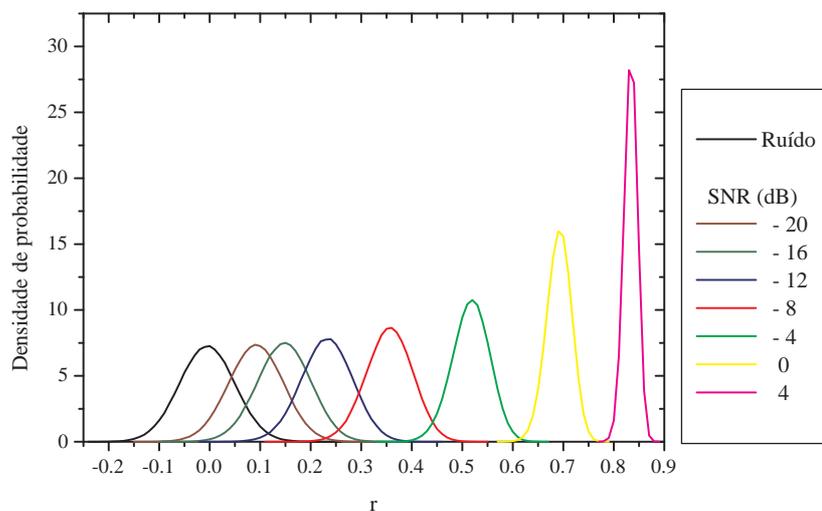


Figura 4.7: Distribuição da estatística r para diferentes valores de SNR e para o ruído.

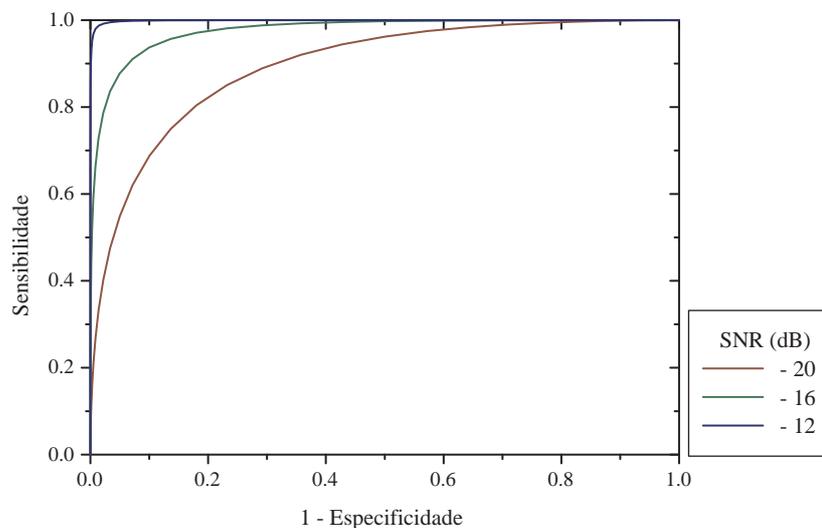


Figura 4.8: Curvas ROC para o método da correlação para alguns valores de SNR.

O teste da correlação começa a perder sensibilidade e especificidade apenas para valores muito baixos de SNR (inferiores a -8dB). Portanto, se o modelo da HRF utilizado representar com fidelidade a resposta esperada, a correlação pode ser um método excelente para detecção de ativação.

4.2.2 Correlação - Dados reais

Para a análise de dados reais, calculamos a correlação das séries temporais de cada *voxel* com o modelo de HRF. A partir da distribuição de r calculada para ruído, obtivemos o nível descritivo p para cada estatística r calculada. A fig. 4.9, mostra os mapas de ativação obtidos a partir da correlação.

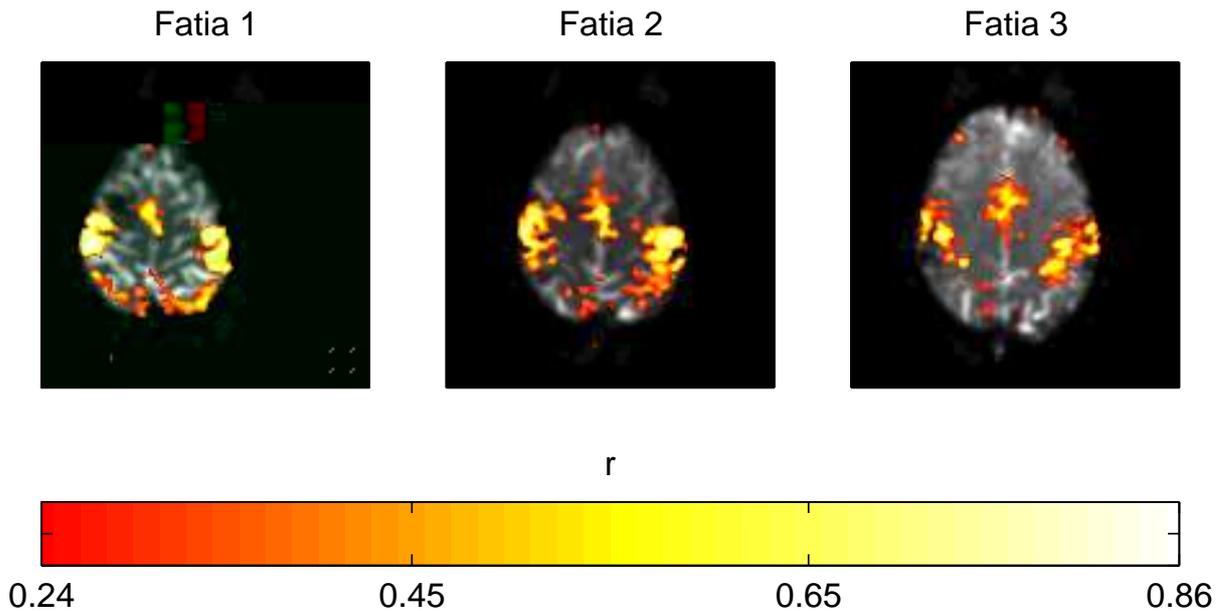


Figura 4.9: Mapas de ativação obtidos a partir do método da correlação. O valor de corte ($r_c = 0,242$) foi obtido a partir da distribuição de r correspondente a ruído somente, dado $\alpha_f = 0,05$.

Assim como no teste t , os mapas obtidos através da correlação apresentam ativação em regiões do córtex primário e secundário, de acordo com o estímulo aplicado.

Vale ressaltar que para implementar a correlação, utilizamos a distribuição da estatística r através de simulações do ruído, e com ela pudemos obter o valor de corte (r_c), dado um nível de significância por teste (α_f). Por outro lado, a variável r está relacionada à variável t de Student [28] através da equação:

$$t = r \sqrt{\frac{2n_j - 2}{1 - r^2}}. \quad (4.12)$$

Esta relação também pode ser usada para obtenção do nível descritivo, bastando considerar a distribuição t com $(2n_j - 2)$ graus de liberdade. Para verificar se as simulações estão de acordo com a relação da eq. (4.12), simulamos 10^6 séries temporais (experimentos) contendo apenas ruído com $n = 12$ elementos. Calculamos o valor da correlação r para cada experimento e associamos a este um valor de t . Obtemos a distribuição de t calculada dessa forma e comparamos

com a distribuição de t com $(v = 2n_j - 2 = 22)$ graus de liberdade. Como mostram as figs. 4.10 e 4.11, as simulações reproduzem muito bem a relação dada pela eq. (4.12).

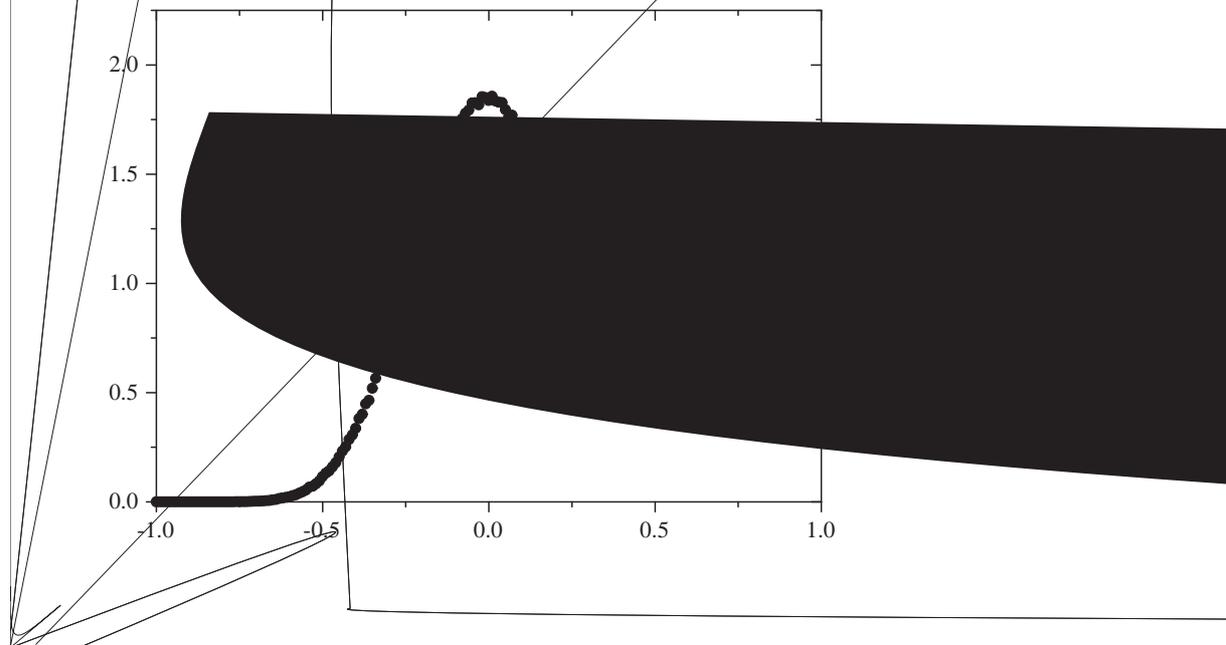


Figura 4.10: Distribuição r obtida a partir de ruído simulado.

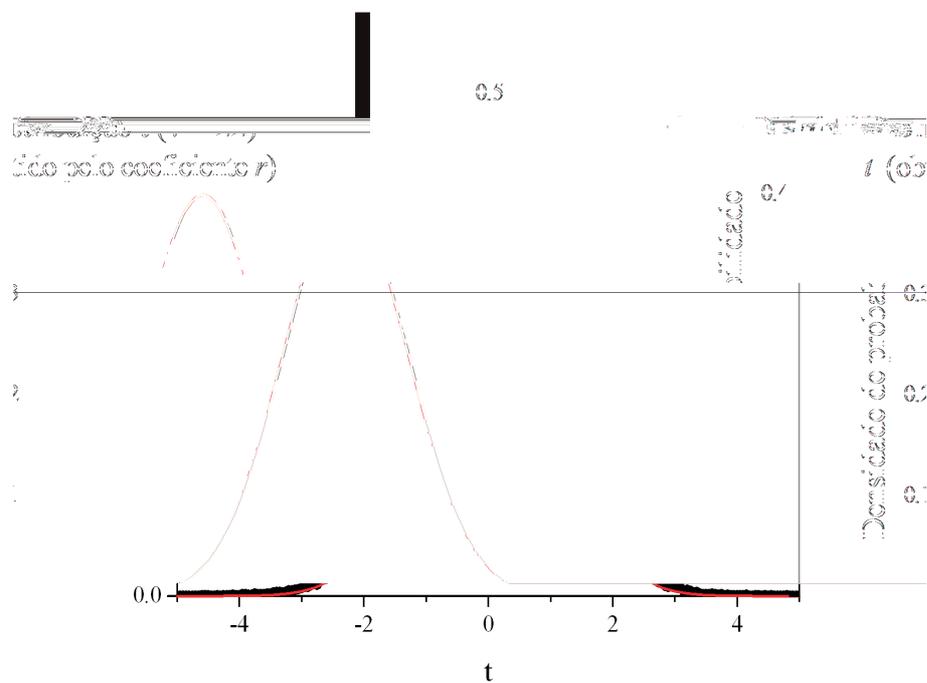


Figura 4.11: Distribuição t com $(v = 22)$ graus de liberdade, sobreposta a distribuição t obtida a partir do coeficiente de correlação r .

Pelos resultados apresentados, vemos que a correlação pode ser eficiente ao detectar ati-

vação neural desde que se tenha um conhecimento prévio da função resposta hemodinâmica. Alguns trabalhos [29, 30] mostram que a HRF pode mudar, dependendo da região cerebral, indivíduo ou estímulo estudados. Assim, a necessidade de um modelo preditor para a HRF pode dificultar ou até mesmo inviabilizar o método. Além disso, alguns cuidados devem ser tomados (tal como no teste t):

1. Estabelecer quais pontos correspondem a uma condição experimental (estímulo) e a outra (repouso);
2. Evitar artefatos que causem alguma mudança sistemática nas condições experimentais.

5 *Distâncias de Kullback-Leibler em fMRI*

Alguns métodos de análise, como a correlação, precisam de um modelo da função resposta hemodinâmica (HRF) para serem aplicados. Ao se adotar um modelo padrão para tentar identificar se há ou não ativação, surge a questão quanto à natureza realística do modelo empregado. Por outro lado, métodos que não necessitam de um conhecimento prévio da forma da HRF são mais flexíveis quanto aos tipos de variações possíveis. Nesse contexto, aparecem os métodos baseados em teoria da informação. As entropias de Shannon, Tsallis e a entropia relativa (conhecida também como *distância de Kullback-Leibler*), são medidas de informação que têm sido empregadas na análise de sinais de fMRI. Essas medidas quantificam a entropia ("desordem") do sinal inferindo a presença ou não de ativação neural.

A entropia de Shannon foi utilizada por *Araujo et al.* [7], mostrando que os métodos entrópicos podem de fato ser uma boa alternativa para análise em fMRI. Os métodos entrópicos também foram estudados por *Sturzbecher* [31] neste contexto.

Neste capítulo, aplicamos a distância de Kullback-Leibler na análise de sinais de fMRI, obtendo sua distribuição de probabilidade através de simulações numéricas. Obtemos as curvas ROC, que demonstram a sensibilidade e especificidade do método empregado e também analisamos dados reais. Finalmente, estendemos este estudo, considerando a distância de Kullback-Leibler generalizada.

5.1 Entropia de Shannon

O conceito de entropia na teoria da informação tem origem no artigo *A Mathematical Theory of Communication* [32] escrito pelo matemático e engenheiro eletrônico *Claude Elwood Shannon* (1916 - 2001). Em teoria da informação, a entropia de Shannon é uma medida da incerteza associada a uma variável aleatória (X) que apresenta L valores, $X = \{x_i, i = 1, 2, \dots, L\}$, com probabilidades $P = \{p(x_i), i = 1, 2, \dots, L\}$. O cálculo da entropia é dado por:

$$H(X) = -k \sum_{i=1}^L p(x_i) \log_2[p(x_i)], \quad (5.1)$$

em que k é uma constante positiva.

Para $k = 1$, esta quantidade mede a informação em *bits* contida em uma mensagem. O seu valor é a quantidade (comprimento) mínima de *bits*, necessária para transmitir informações. Quanto mais previsível é a mensagem, menor o comprimento necessário para transmiti-la e vice-versa.

Como exemplo, considere o lançamentos de moedas e a observação do resultado (cara ou coroa). A entropia da jogada é maximizada se a moeda é justa, isto é, se cara e coroa têm iguais probabilidades de ocorrer (50%). Esta é a situação de máxima incerteza, uma vez que é mais difícil prever o resultado da jogada. No entanto, se sabemos que a moeda não é justa, então há menos incerteza. Todo o tempo, um dos lados tem mais probabilidades de ocorrer. A incerteza e, conseqüentemente a entropia são reduzidas. O caso extremo é o de uma moeda totalmente viciada em que todas as jogadas revela apenas uma face. Não havendo incerteza. A entropia é zero: cada jogada da moeda não fornece nenhuma informação.

Portanto, para estados equiprováveis a entropia é máxima, enquanto que para situações extremas (probabilidade máxima para um estado e mínima para os outros), a entropia é mínima. Considere a variável aleatória (X) que pode assumir dois estados, x_1 e x_2 , com probabilidades $p(x_1)$ e $p(x_2) = [1 - p(x_1)]$ o valor da entropia $H(X)$ em função de $p(x_1)$, está ilustrado no gráfico da fig. 5.1.

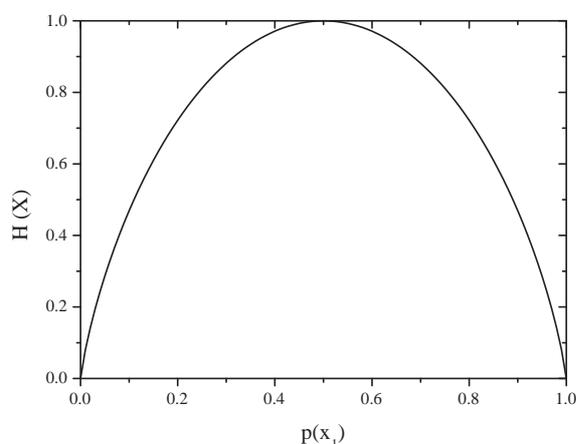


Figura 5.1: Entropia de Shannon $H(X)$ para diferentes valores de $p(x_1)$. A entropia é máxima para $p(x_1) = p(x_2) = 0,5$ e mínima para os casos extremos $[p(x_1) = 0; p(x_2) = 1$ e $p(x_1) = 1; p(x_2) = 0]$.

No estudo publicado por *Araujo et al.* [7], dada a série temporal com períodos intercalados

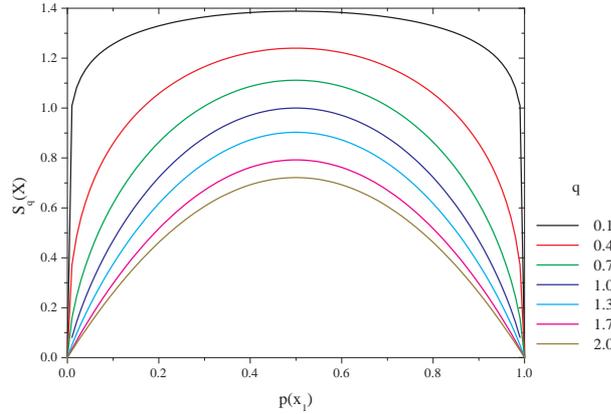


Figura 5.3: Entropia de Tsallis $S_q(X)$ em função de $p(x_1)$ para diferentes valores de q . Para q positivo, a entropia é máxima para $[p(x_1) = p(x_2) = 0.5]$ e mínima para $[p(x_1) = 0; p(x_2) = 1$ e $p(x_1) = 1; p(x_2) = 0]$. Quando $q = 0$ a entropia é constante independente do valor de $p(x_1)$, e para $q = 1$ obtemos a entropia de Shannon.

A entropia de Tsallis foi utilizada por *Sturzbecher* [31], e neste estudo, a introdução do parâmetro q possibilitou uma otimização do método pela escolha apropriada desse valor.

5.3 Distância de Kullback-Leibler

Em teoria da informação, entropia relativa é uma medida de "distância" entre distribuições de probabilidade p_1 e p_2 de duas variáveis discretas X_1 e X_2 , respectivamente. Esta medida de informação é conhecida como distância de Kullback-Leibler (D) [34] e é definida por:

$$D(p_1 || p_2) = \sum_{j=1}^L p_{1j} \log_2 \left(\frac{p_{1j}}{p_{2j}} \right), \quad (5.4)$$

com p_{ij} a probabilidade de X_i assumir o j -ésimo valor de seu conjunto de L elementos. O logaritmo na base 2 é usado de forma a obter o valor de entropia em *bits*.

A entropia relativa não é simétrica, isto é, $D(p_1 || p_2) \neq D(p_2 || p_1)$ e seu valor é sempre não-negativo. É zero se e somente se $p_1 = p_2$. Para simplificar a notação, neste trabalho usaremos o termo D ao nos referir a $D(p_1 || p_2)$. Para o cálculo de D , devemos também considerar os seguintes limites:

$$\lim_{\substack{p_{1j} \rightarrow 0 \\ p_{2j} \neq 0}} p_{1j} \log_2 \left(\frac{p_{1j}}{p_{2j}} \right) = 0, \quad (5.5)$$

e

$$\lim_{\substack{p_{2j} \rightarrow 0 \\ p_{1j} \neq 0}} p_{1j} \log_2 \left(\frac{p_{1j}}{p_{2j}} \right) = \infty. \quad (5.6)$$

Dada a série temporal, para o cálculo da entropia de Kullback-Leibler, associamos cada época do experimento a uma janela W . Essa janela é dividida ao meio, de forma que a primeira parte W

Pela eq. (5.6), notamos que a entropia relativa pode divergir no limite $p_{2j} \rightarrow 0$ (para $p_{1j} \neq 0$). Na prática, esta indesejável divergência seria suscetível de ocorrer se a função de probabilidade p_2 puder assumir valores nulos. Para evitar esta divergência, introduzimos um pequeno parâmetro positivo δ ($0 < \delta < 1$) na definição das probabilidades dos níveis da série temporal, tal que,

$$p_{ij} = \frac{n_{ij} + \delta}{n_i + L\delta}, \quad (5.7)$$

em que $n_{ij} = \sum_{j=1}^L I_{ij}$ é o número de pontos na janela W_i ($i=1,2$) e nível I_j ($j=1,2,\dots,L$) e $n_i = \sum_{j=1}^L n_{ij}$ é o número total de pontos na janela W_i . Na definição acima, a condição $\sum_{j=1}^L p_{ij} = 1$ é claramente satisfeita enquanto que se $\delta \rightarrow 0^+$ então $p_{ij} \rightarrow n_{ij}/n_i$, recuperando o significado usual da probabilidade p_{ij} de ocupação do nível I_j . Além disto, eq. (5.7) pode ser interpretada como a frequência relativa de pontos no nível I_j (dentro de W_i) dado que o "ponto fracionário" (de peso δ) é igualmente colocado em cada nível I_j de ambas as janelas W_1 e W_2 . Neste sentido, a introdução do parâmetro δ não influencia na comparação entre as distribuições p_1 e p_2 uma vez que ela permanece balanceada.

5.3.1 Distância de Kullback-Leibler - Simulações

Realizamos 10^6 experimentos com $N = 24$ épocas contendo ($n_1 = n_2 = 7$) pontos cada. Para cada experimento, obtemos 24 valores de D , correspondentes a cada época. As distribuições obtidas para ruído simulado e para as SNRs -4dB, 0dB e 4dB, estão representadas nas figs. 5.5, 5.6, 5.7 e 5.8.

Se por um lado a distribuição da variável D é discreta, apresentando picos para alguns valores, a distribuição da variável \bar{D} , calculada pela média das N épocas, pode ser aproximada por uma distribuição contínua. A média amostral \bar{D} é escolhida como a estatística para testar a hipótese nula de que um *voxel* é inativo (ruído). De acordo com o teorema central do limite, a distribuição dos valores de \bar{D} tende a uma distribuição Gaussiana para $N \rightarrow \infty$. Porém para valores menores de N ($N < 30$), verificamos que as funções densidade de probabilidade $f(\bar{D})$ podem ser muito bem aproximadas por distribuições Gamma,

$$f(\bar{D}) \approx \frac{\bar{D}^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} \exp\left(-\frac{\bar{D}}{\beta}\right), \quad (5.8)$$

em que α e β são parâmetros da distribuição relacionados com a média $\mu = \alpha\beta$ e variância $\sigma^2 = \alpha\beta^2$. A fig. 5.9 apresenta as densidades de probabilidade $f(\bar{D})$ correspondentes a simulações do ruído ($SNR \rightarrow -\infty$) e alguns valores finitos de SNR (-12dB, -8dB, -4dB, 0dB e 4dB); as curvas

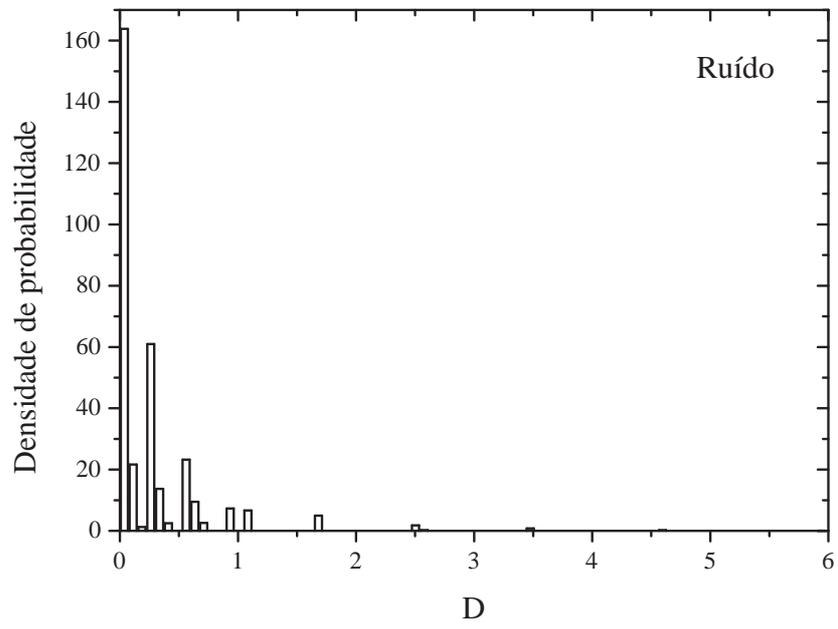


Figura 5.5:

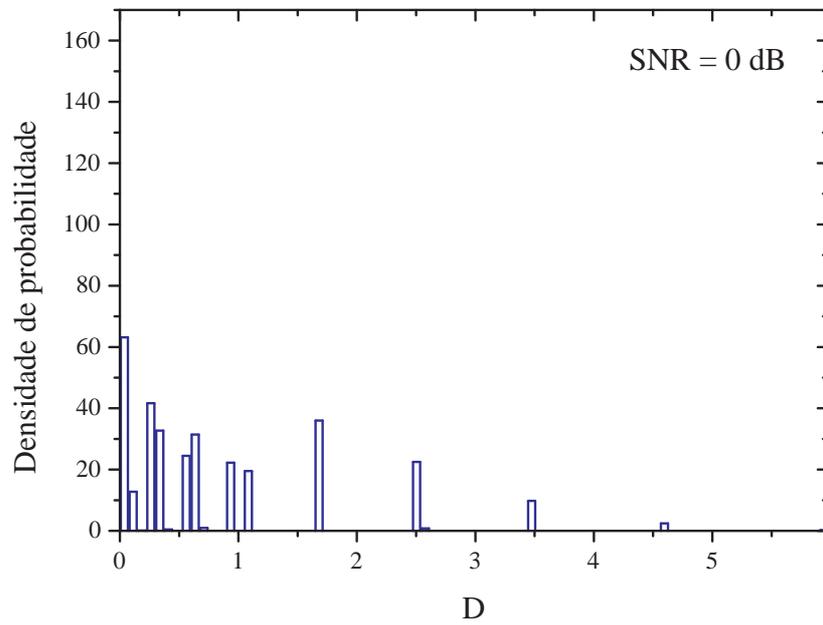


Figura 5.7: Distribuição da variável D calculada para $SNR = 0dB$; parâmetros: $\delta = 0,1$ e $L = 2$.

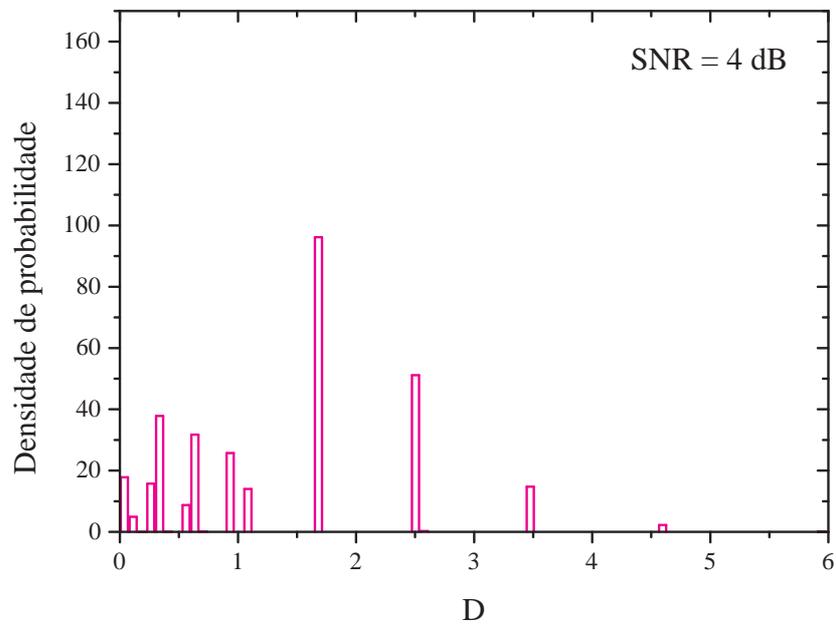


Figura 5.8: Distribuição da variável D calculada para $SNR = 4dB$; parâmetros: $\delta = 0,1$ e $L = 2$.

contínuas correspondem ao ajuste pela distribuição Gamma ($\chi^2 < 0,0009$) que se sobrepõe aos pontos simulados. Os parâmetros do ajuste são mostrados na tabela 5.1.

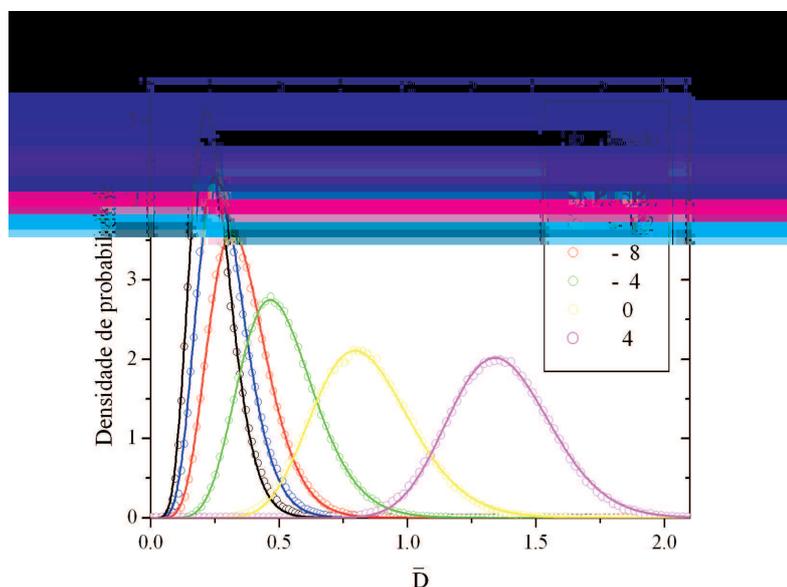


Figura 5.9: Densidade de probabilidade $f(\bar{D})$ para simulações de ruído, e alguns valores finitos de SNR; usando os parâmetros $N = 24$, $n_1 = n_2 = 7$, $L = 2$ e $\delta = 0.1$. As curvas contínuas correspondem à aproximação pela função Gamma ($\chi^2 < 0,0009$) dos valores simulados.

Parâmetros do ajuste				
SNR	α	β	χ^2	R^2
Ruído	8,96(2)	0,02731(5)	0,0009	0,99844
-12	8,92(1)	0,03250(5)	0,0005	0,99893
-8	9,37(1)	0,03800(4)	0,0002	0,99949
-4	11,494(8)	0,04446(3)	0,0006	0,99979
0	18,86(2)	0,04461(6)	0,0001	0,99937
4	47,17(7)	0,02906(4)	0,0002	0,99904

Tabela 5.1: Parâmetros do ajuste da distribuição de \bar{D} , pela distribuição Gamma. Simulações realizadas utilizando os parâmetro ($N = 24$, $n_1 = n_2 = 7$; $L = 2$; $\delta = 0, 1$).

5.3.1.1 Otimização do número de níveis (L) e do parâmetro δ

Nas simulações realizadas, otimizamos os parâmetros L e δ . Os valores otimizados são posteriormente empregados na análise de dados reais.

Calculamos a área abaixo da curva ROC para diferentes valores de SNR. Quanto maior a área ROC, melhor é a relação entre sensibilidade e especificidade. Para um número fixo de níveis e diferentes SNRs, encontramos o valor de δ que maximiza a área abaixo da curva ROC, cf. figs. 5.10, 5.11, 5.12 e 5.13.

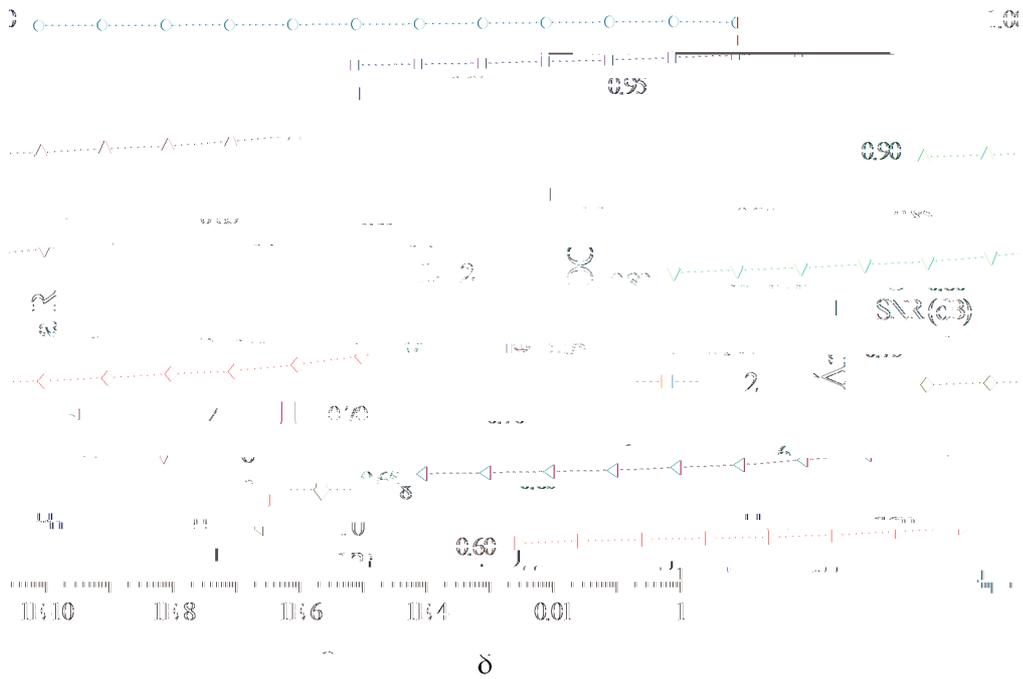


Figura 5.10: Otimização do parâmetro δ para os números de níveis ($L = 2$). O valor de δ que maximiza as áreas ROC é $\delta = 0,1$.

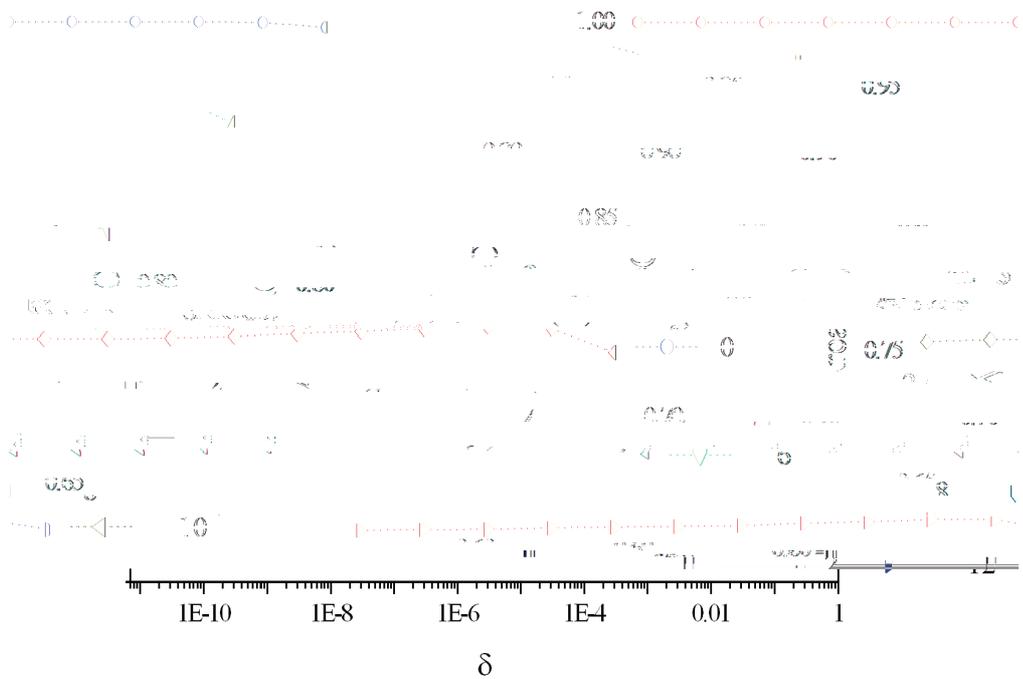


Figura 5.11: Otimização do parâmetro δ para os números de níveis ($L = 3$). O valor de δ que maximiza as áreas ROC é $\delta = 0,01$.

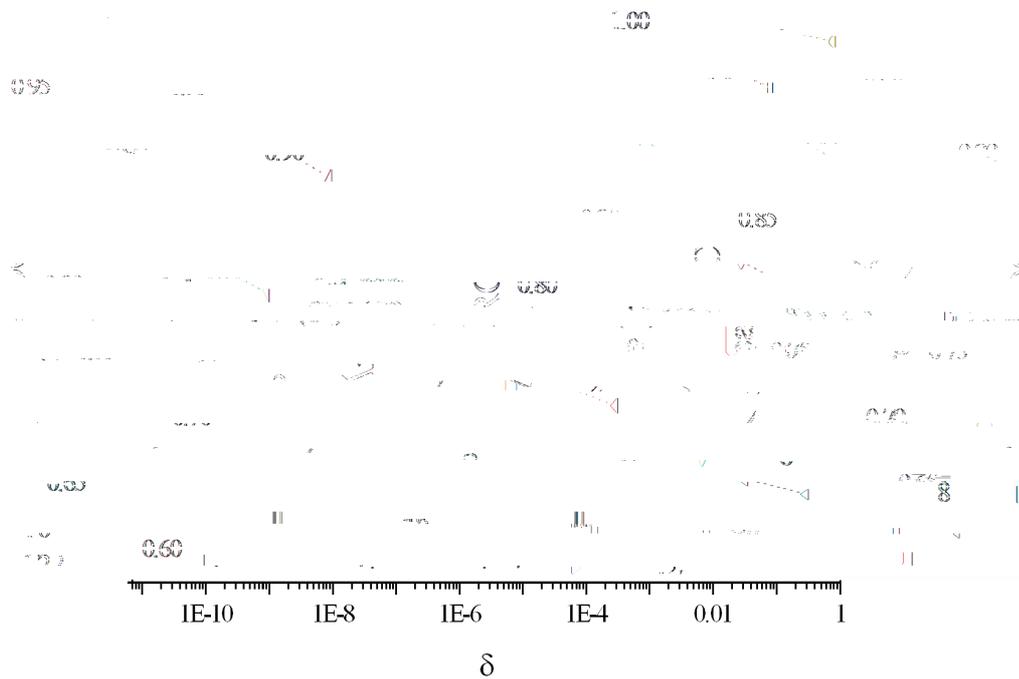


Figura 5.12: Otimização do parâmetro δ para os números de níveis ($L = 4$). O valor de δ que maximiza as áreas ROC é $\delta = 0,001$.

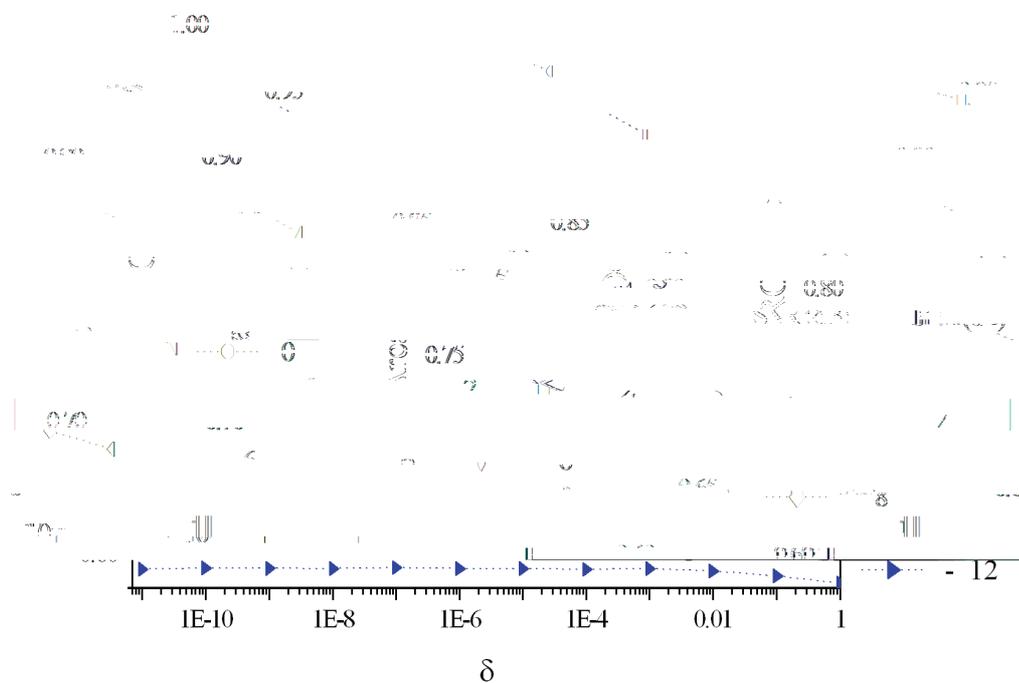


Figura 5.13: Otimização do parâmetro δ para os números de níveis ($L = 5$). O valor de δ que maximiza as áreas ROC é $\delta = 0,0001$.

O valor de δ otimizado depende do número de níveis. Para $L = 2$, o valor de δ em que a área ROC é máxima está por volta de 0,1. Para $L = 3$ este valor é $\delta = 0,01$. Para $L = 4$ e $L = 5$ o valor da área ROC varia muito pouco abaixo de certo valor de δ (variação na terceira casa), adotamos $\delta = 10^{-3}$ (para $L = 4$) e $\delta = 10^{-4}$ (para $L = 5$) como valores otimizados.

A fig. 5.14 apresenta o comportamento da área sob a curva ROC, em função da SNR, para diferentes níveis ($L=2,3,4$ e 5), e parâmetros δ otimizados correspondentes. Tal gráfico permite identificar o número de níveis ideal para o estudo.

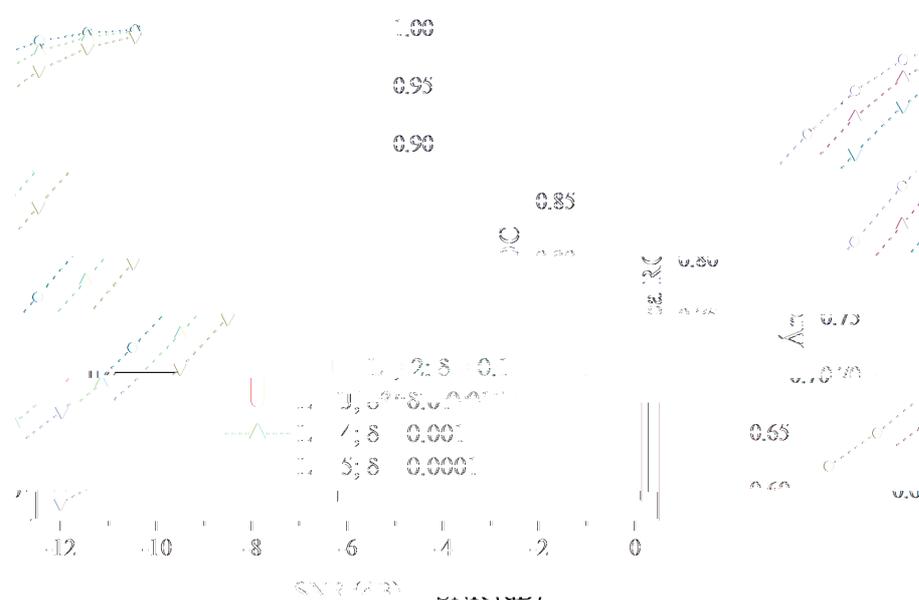


Figura 5.14: Otimização do número de níveis L . A área abaixo da curva ROC é maior ao usar os parâmetros $L = 2$ e $\delta = 0,1$ para uma ampla faixa de valores de SNR.

A medida que aumentamos o número de níveis, a área ROC diminui. Verificamos que para $L = 2$ e $\delta = 0,1$, a área ROC é maior para diferentes valores de SNR. Portanto, utilizaremos os valores otimizados $L = 2$ e $\delta = 0,1$ para a análise de dados reais.

5.3.2 Distância de Kullback-Leibler - Dados reais

Usando os valores de δ e L otimizados, analisamos os dados reais com o método da entropia de Kullback-Leibler. Para obter o valor do nível descritivo p , utilizamos a distribuição da variável \bar{D} calculada para ruído simulado. A fig. 5.15 mostra o mapa de ativação obtido.

O mapa apresenta ativação em regiões do córtex motor primário e secundário de acordo com o estímulo aplicado. Usando um nível de significância por família de testes ($\alpha_f = 0,05$), o

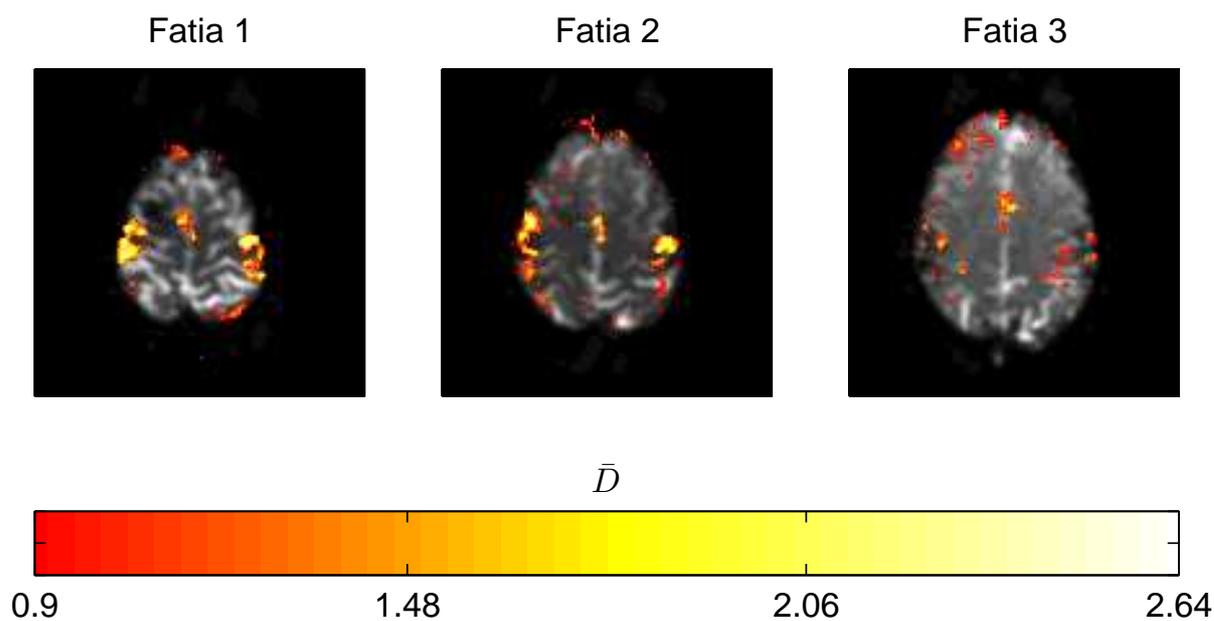


Figura 5.15: Mapa de ativação obtido a partir da variável \bar{D} calculada para ruído simulado. Foram utilizados os parâmetros otimizados $L = 2$ e $\delta = 0.1$.

valor de corte obtido pela distribuição do ruído simulado foi de $\bar{D}^{\alpha} \approx 0,88$.

Apenas como comparação, utilizamos valores não otimizados de L e δ para verificar a diferença no mapa de ativação. O mapa a seguir (fig. 5.16) apresenta o mapa obtido com os valores não-otimizados ($L = 5$; $\delta = 0.1$).

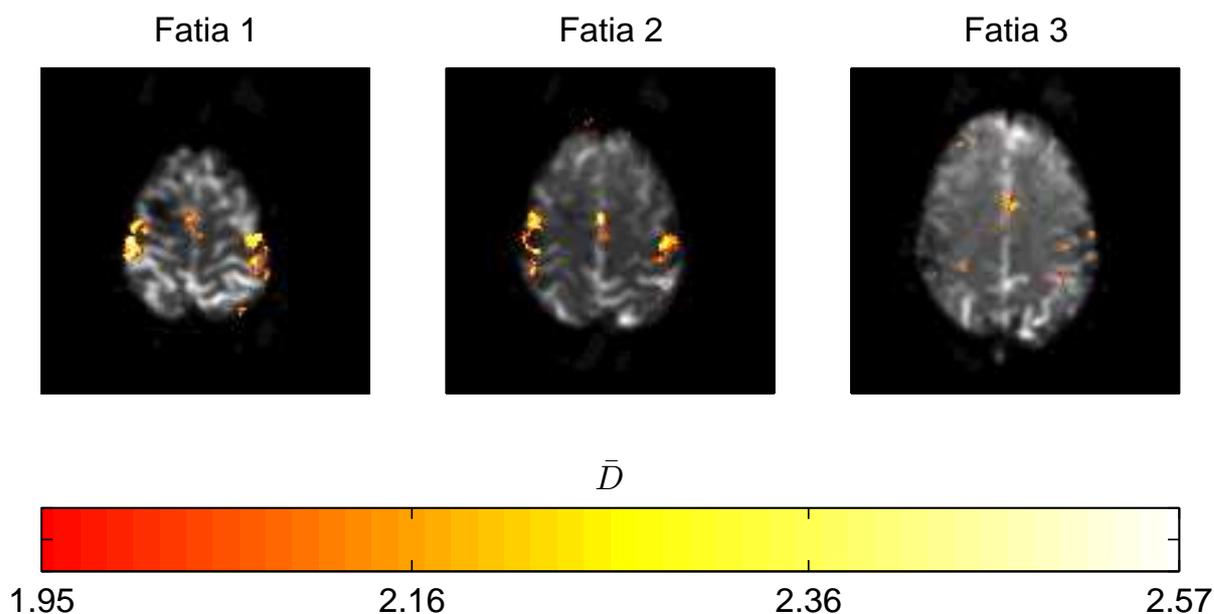


Figura 5.16: Mapa de ativação obtido a partir da variável \bar{D} calculada para ruído simulado. Foram utilizados os parâmetros não otimizados $L = 5$ e $\delta = 0,1$.

Utilizando o mesmo nível de significância ($\alpha_f = 0,05$) e os parâmetros não-otimizados,

obtemos o valor crítico $\bar{D}^* \approx 1,95$. Ao comparar os mapas das figs. 5.15 e 5.16, observamos que ao utilizar parâmetros não-otimizados, o número de *voxels* considerados ativos caiu significativamente. Isto era esperado, uma vez que ao usar o mesmo valor de nível de significância (α_f) para ambos os mapas, fixamos a fração de erros do tipo I (falsos positivos). O que muda com a utilização de parâmetros não otimizados é a fração de erros do tipo II (falsos negativos), fazendo com que menos *voxels* sejam considerados ativos, mesmo que eles apresentem ativação. Para ilustrar esse comportamento, a fig. 5.17 apresenta as distribuições de \bar{D} calculado para ruído simulado e as distribuições para (sinal + ruído) simulado com SNR=2dB e 4dB. Usando o mesmo valor de nível de significância ($\alpha_f = 0,05$; $\alpha_t = 3,13 \cdot 10^{-6}$) para os parâmetros otimizados ($L = 2$ e $\delta = 0,1$) e para os não-otimizados ($L = 5$ e $\delta = 0,1$), podemos verificar a diferença esperada na fração de erros do tipo II (falsos negativos).

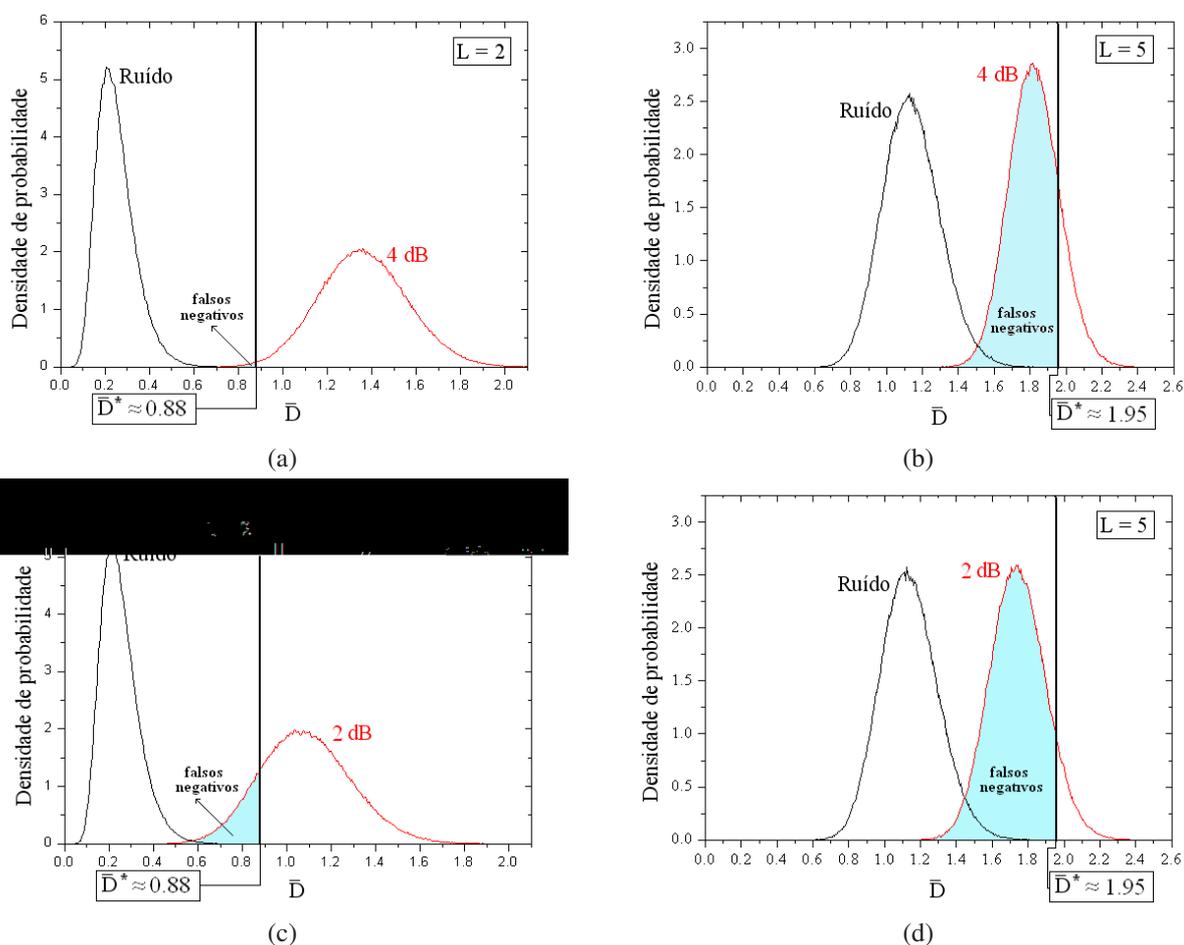


Figura 5.17: Gráficos ilustrando o número de falsos negativos usando parâmetros otimizados ($L = 2$; $\delta = 0,1$) e não otimizados ($L = 5$; $\delta = 0,1$) para dois valores de SNR (2dB e 4dB). Parâmetros não otimizados, geram uma quantidade de erros do tipo II maior em relação ao uso dos parâmetros otimizados. Como o nível de significância é o mesmo para ambos os casos ($\alpha_f = 0,05$), o número de falsos positivos é o mesmo.

O uso dos parâmetros não-otimizados sempre acarretará um maior número de erros do

tipo II. Porém, é importante lembrar que os gráficos da fig. 5.17 apresentam uma diferença específica para $\text{SNR}=2\text{dB}$ e $\text{SNR}=4\text{dB}$. Para outros valores de SNR essa diferença pode ser maior ou menor. Nos dados reais, como existem vários valores de SNR, o número de erros do tipo II a mais (por não utilizar os parâmetros otimizados), vai depender desses valores de SNR.

As curvas ROC, fornecem informação a respeito da relação especificidade-sensibilidade. Como vimos anteriormente, a área abaixo da curva ROC é maior ao usar os parâmetros otimizados. O gráfico da fig. 5.18 mostra esse comportamento para alguns valores de SNR, comparando o uso de parâmetros otimizados com os não-otimizados.

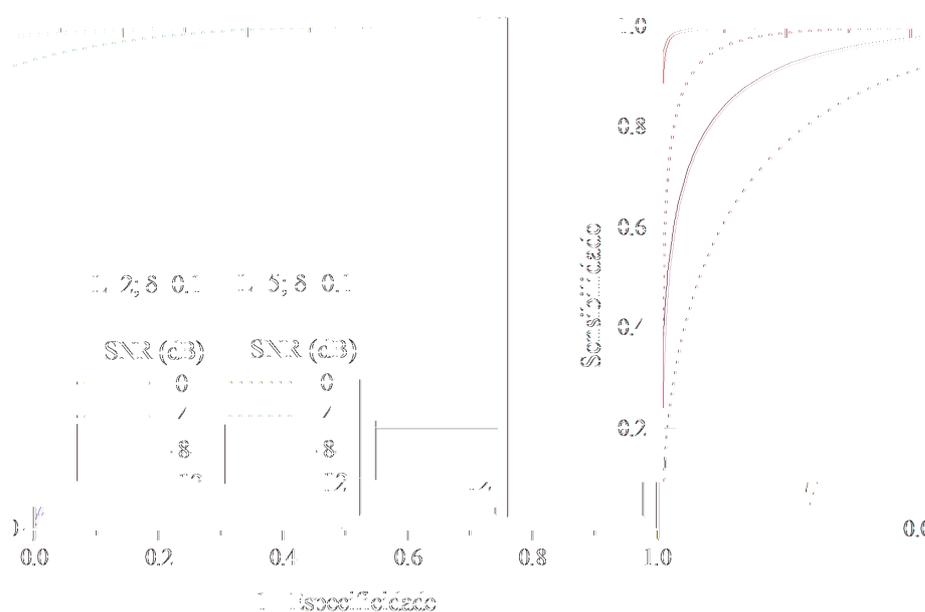


Figura 5.18: Curvas ROC para os parâmetro otimizados ($L = 2$; $\delta = 0, 1$) e para os parâmetros não-otimizados ($L = 5$; $\delta = 0, 1$).

Conforme mostramos, a entropia de Kullback-Leibler pode ser utilizada na análise de dados reais sem utilizar um modelo de resposta hemodinâmica pré-definido. Os mapas de ativação podem ser obtidos apenas pela distribuição densidade de probabilidade da quantidade \bar{D} calculada para o ruído simulado. Tal distribuição pode ser obtida eliminando a divergência associada ao cálculo de \bar{D} através da inserção do parâmetro (δ).

5.4 Distância de Kullback-Leibler generalizada

Nesta seção, estendemos nosso estudo, utilizando a distância de Kullback-Leibler generalizada. Assim como a entropia de Shannon, a distância de Kullback-Leibler também pode ser

generalizada através da definição do q -logaritmo¹ [35, 36]:

$$\ln_q(p) = \frac{p^{1-i} - q}{1-i} \quad (5.9)$$

Como função do q -logaritmo, distância de Kullback-Leibler é escrita como.

$$\begin{aligned} D_q &= k \sum_{j=1}^L p_{1j} \ln_q \left(\frac{p_{1j}}{p_{2j}} \right) \\ &= \frac{k}{(1-i)} \sum_{j=1}^L p_{1j}^q \left(p_{1j}^{1-i} - q \right) \left(p_{2j}^{1-i} - q \right), \end{aligned} \quad (5.10)$$

em que k é uma constante para ajustar a unidade dos valores de D_q . Se $k=(1/\ln 2)$, a entropia relativa é dada em *bits* (para $q \neq 0$), mantendo a definição original da distância de Kullback-Leibler no contexto de teoria da informação.

O parâmetro q pode assumir qualquer valor real ($q \in \mathfrak{R}$), porém para valores menores do que 0 e maiores de que 1, a quantidade D_q pode divergir. Para verificarmos essa condição, consideremos o somatório da eq. (5.10): se $q < 0$, então p_{1j}^q pode divergir para $p_{1j} = 0$. Por outro lado, se $q > 1$, então os termos p_{1j}^{1-i} e p_{2j}^{1-i} podem divergir. Assim, adotamos os valores ($0 < q < 1$) para evitar qualquer tipo de divergência e possibilitar a construção das distribuições densidade de probabilidade da quantidade D_q . O cálculo da quantidade D_q foi realizado da mesma forma apresentada na seção anterior (sec. 5.3), associando a cada época do experimento uma janela W dividida em duas outras janelas W_1 e W_2 . Contudo, com a forma generalizada da entropia de Kullback-Leibler, podemos simplesmente calcular as probabilidades dos níveis do sinal como

$$p_{ij} = \frac{n_{ij}}{n_i}, \quad (5.11)$$

sem a necessidade de se introduzir o parâmetro δ , uma vez que D_q não diverge no intervalo ($0 < q < 1$).

O valor de D_q para ($L = 2$) em função das probabilidades $p_{1,1}$ e $p_{2,1}$ são mostrados na fig. 5.19 para valores de $q = 0, 1; 0, 3; 0, 6$ e $0, 9$. Lembrando que ($p_{1,1} + p_{1,2} = 1$) e ($p_{2,1} + p_{2,2} = 1$).

¹O índice q não se refere a base do logaritmo e sim ao parâmetro da generalização do logaritmo na base natural e .

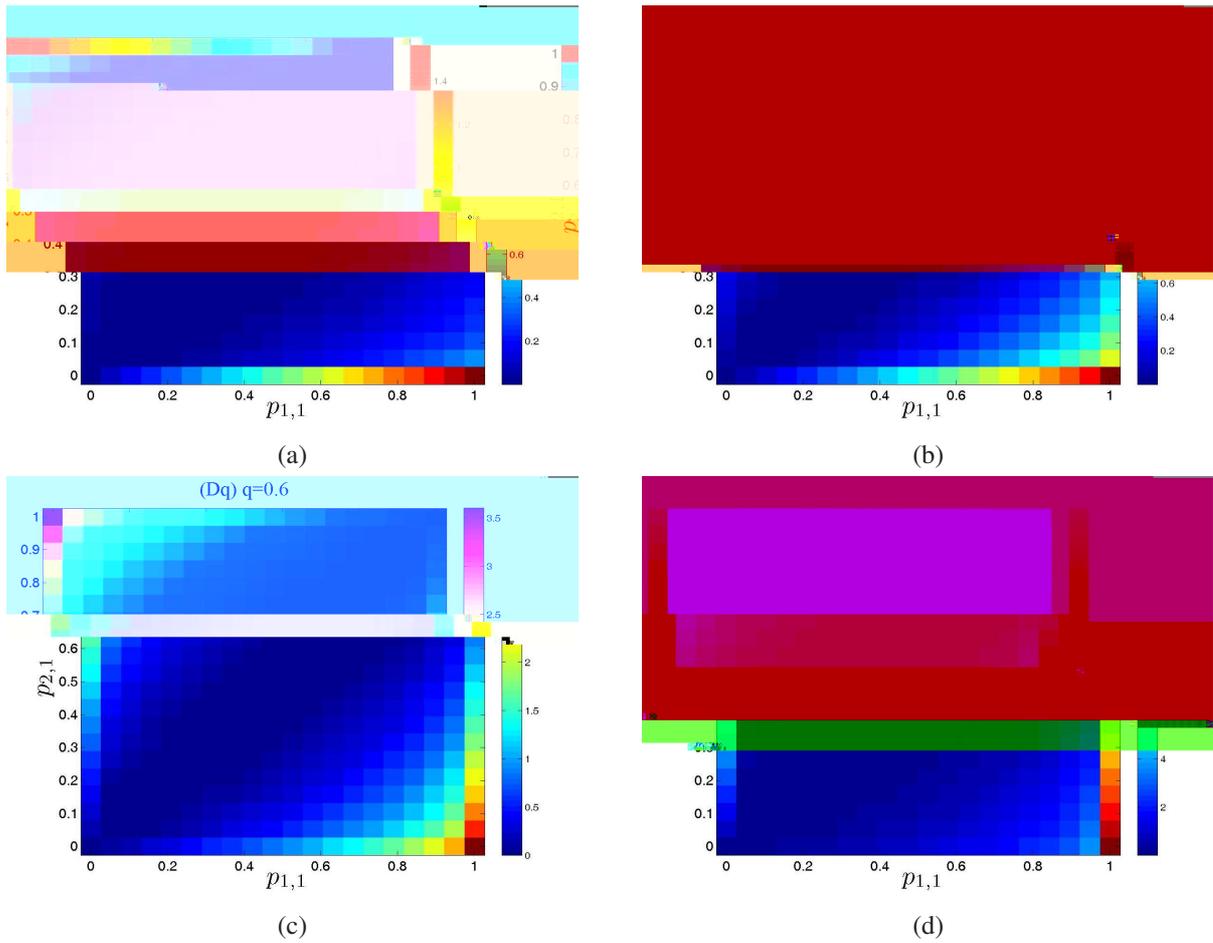


Figura 5.19: Valores de D_q com $L = 2$, para diferentes valores de q . (a) $q=0,1$; (b) $q=0,3$; (c) $q=0,6$; (d) $q=0,9$. Os valores mais altos de D_q ocorrem quando $(p_{2,1} = 1$ e $p_{1,1} = 0)$ ou $(p_{2,1} = 0$ e $p_{1,1} = 1)$.

5.4.1 Distância de Kullback-Leibler Generalizada - Simulações

Realizamos 10^6 experimentos com $N = 24$ épocas contendo $(n_1 = n_2 = 7)$ pontos cada. Para cada experimento, obtemos 24 valores de D_q , correspondentes a cada época. As distribuições obtidas para ruído simulado e para as SNRs -4dB, 0dB e 4dB, estão representadas nas figs. 5.20, 5.21, 5.22 e 5.23.

Assim como no caso da variável D , a quantidade D_q possui distribuição discreta. Porém, ela também pode ser aproximada por uma distribuição Gamma, cf. eq. (5.8) se considerarmos a distribuição de \bar{D}_q calculada como média de D_q ao longo das N épocas. A fig. 5.24 apresenta as densidades de probabilidade $f(\bar{D}_q)$ correspondentes a simulações do ruído ($SNR \rightarrow \infty$) e alguns valores finitos de SNR (-12dB, -8dB, -4dB, 0dB e 4dB). As curvas contínuas correspondem à aproximação pela distribuição Gamma ($\chi^2 < 0,0001$) que sobrepõe os pontos. Os parâmetro do ajuste são mostrados na tabela 5.2.

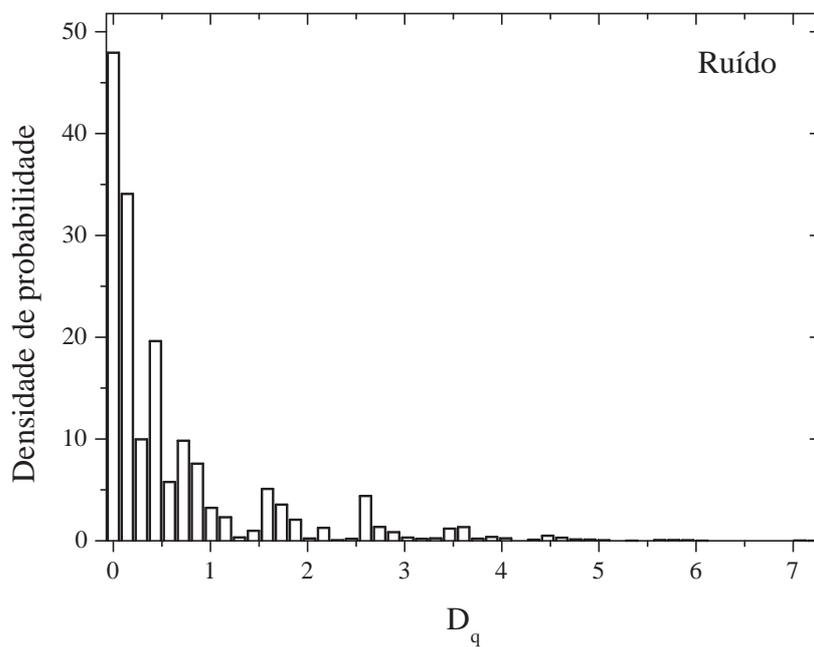


Figura 5.20: Distribuição da variável D_q calculada para ruído; parâmetros: $q = 0,8$ e $L = 3$.

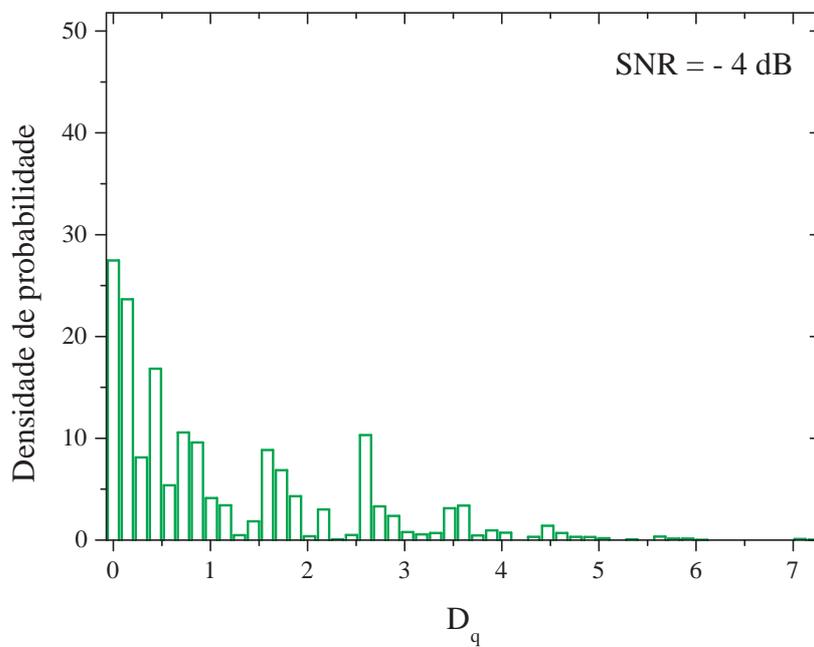


Figura 5.21: Distribuição da variável D_q calculada para $SNR = -4$ dB; parâmetros: $q = 0,8$ e $L = 3$.

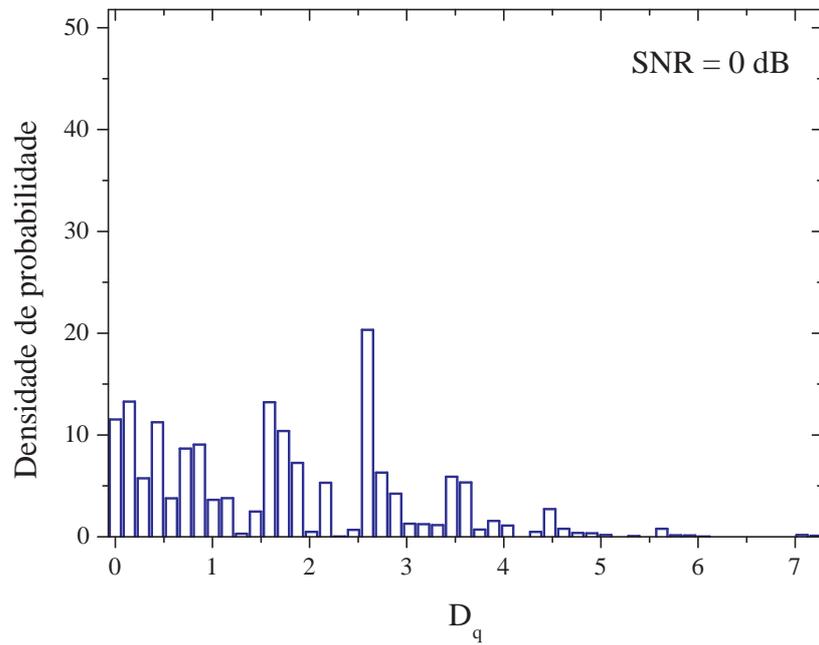


Figura 5.22: Distribuição da variável D_q calculada para $SNR = 0 \text{ dB}$; parâmetros: $q = 0,8$ e $L = 3$.

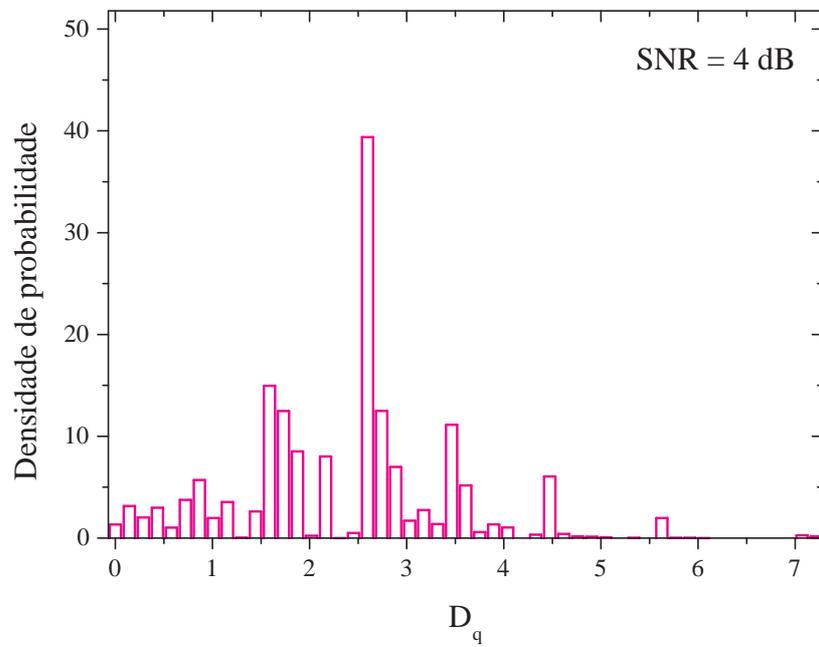


Figura 5.23: Distribuição da variável D_q calculada para $SNR = 4 \text{ dB}$; parâmetros: $q = 0,8$ e $L = 3$.

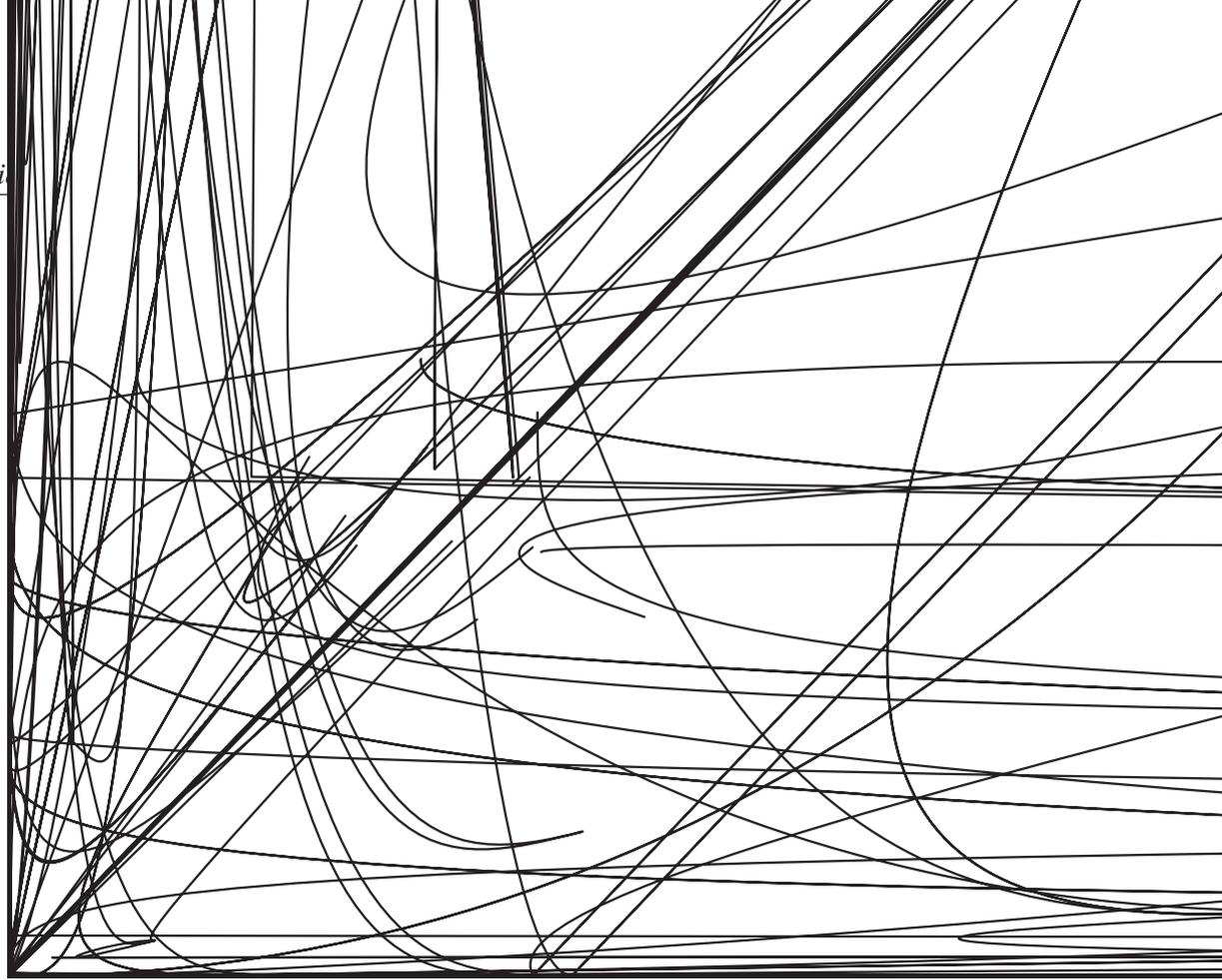


Figura 5.24: Densidade de probabilidade $f(\bar{D}_q)$ para simulações de ruído, e alguns valores finitos de SNR; usando os parâmetros $N = 24$, $n_1 = n_2 = 7$, $L = 3$ e $q = 0,8$. As curvas contínuas correspondem à aproximação pela função Gamma ($\chi^2 < 0,0001$) dos valores simulados.

SNR	Parâmetros do ajuste			
	α	β	χ^2	R^2
Ruído	13,33(2)	0,05148(6)	0,00003	0,99985
-12	14,42(2)	0,0537(5)	0,00006	0,99967
-8	16,51(3)	0,0547(4)	0,00007	0,99958
-4	22,90(5)	0,0524(1)	0,00009	0,99939
0	45,1(1)	0,0392(6)	0,0001	0,99924
4	121,1(2)	0,01994(4)	0,00006	0,99963

Tabela 5.2: Parâmetros do ajuste da distribuição de \bar{D}_q simulado, pela distribuição Gamma. Simulações realizadas utilizando os parâmetro ($N = 24$, $n_1 = n_2 = 7$; $L = 3$; $q = 0,8$).

5.4.1.1 Otimização do número de níveis (L) e do parâmetro q

Como o cálculo da distância de Kullback-Leibler Generalizada depende de dois parâmetros (L e q), realizamos simulações para verificar quais os parâmetros que maximizam a área abaixo das curvas ROC para diferentes valores de SNR. Inicialmente fixamos o número de níveis (L), encontrando o valor de q que maximiza a área abaixo da curva ROC para diferentes valores de SNR, cf. figs. 5.25, 5.26, 5.27, 5.28.

O valor de q otimizado depende do número de níveis. Para os níveis considerados, temos os seguintes valores de q que maximizam a área ROC: $L = 2$ ($q = 0,70$), $L = 3$ ($q = 0,80$),

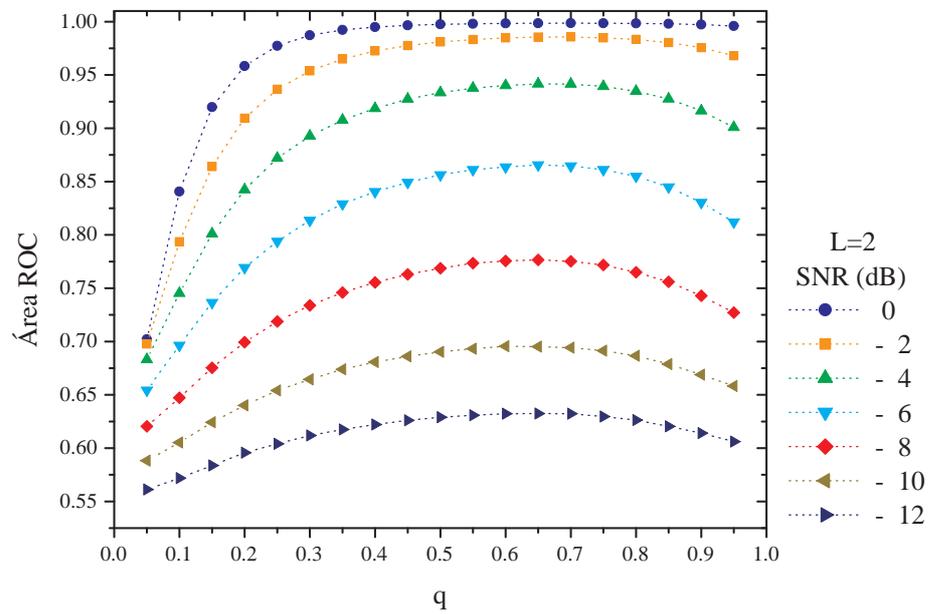


Figura 5.25: Otimização do parâmetro q para o números de níveis ($L = 2$). O valor de q que maximiza as áreas ROC é $q = 0,70$.

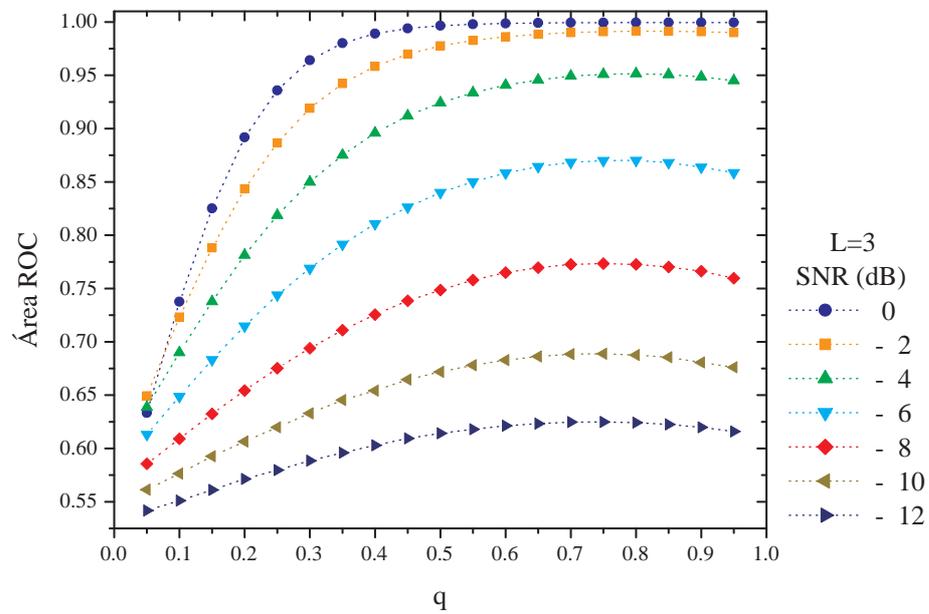


Figura 5.26: Otimização do parâmetro q para o números de níveis ($L = 3$). O valor de q que maximiza as áreas ROC é $q = 0,80$.

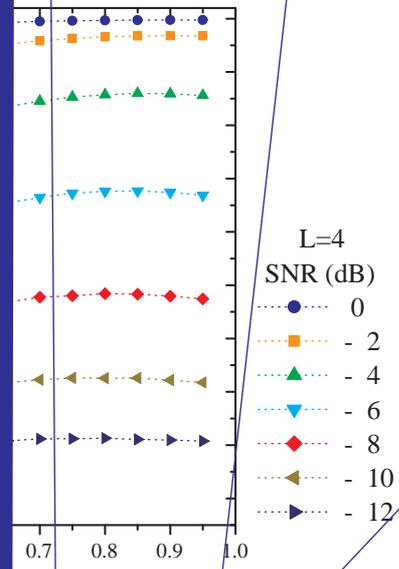


Figura 5.27: Otimização do parâmetro q para o número de níveis ($L = 4$). O valor de q que maximiza as áreas ROC é $q = 0,95$.

de níveis ($L = 4$). O valor de q que maximiza

Figura 5.28: Otimização do parâmetro q para o número de níveis ($L = 5$). O valor de q que maximiza as áreas ROC é $q = 0,95$.

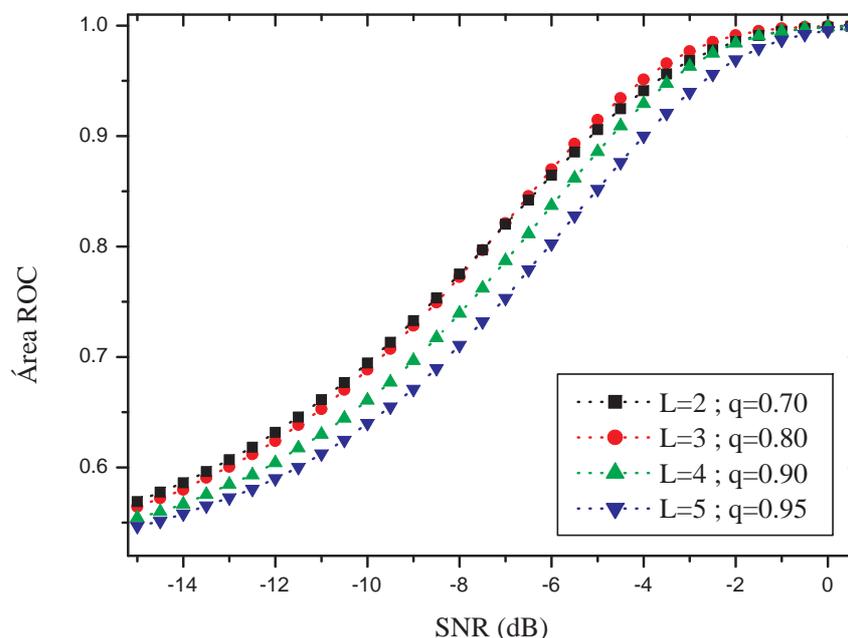


Figura 5.29: Otimização do número de níveis L . A área abaixo da curva ROC é maior ao usar os parâmetros $L = 3$ e $q = 0,80$ para a faixa de valores de $SNR > 8dB$.

$L = 4$ ($q = 0,90$), $L = 5$ ($q = 0,95$). Para verificar qual o número de níveis ideal, comparamos os valores da área ROC para os diferentes níveis ($L=2,3,4$ e 5), usando o valor de q otimizado, cf. fig. 5.4.1.1.

Para valores altos de SNR ($> 8dB$), a melhor escolha é $L = 3$ ($q = 0,8$), mas para baixos valores de SNR ($< 8dB$), os parâmetros $L = 2$ ($q = 0,7$) apresentam valores ligeiramente maiores de área ROC. Porém, para baixos valores de SNR, o teste não é confiável (área abaixo da curva ROC menor que 0,75), e por essa razão os parâmetros otimizados escolhidos foram $L = 3$ e $q = 0,80$.

5.4.2 Distância de Kullback-Leibler generalizada - Dados reais

Usando os valores de q e L otimizados, calculamos a entropia de Kullback-Leibler generalizada para a série de dados reais. Para obter o valor do nível descritivo p , utilizamos a distribuição da variável \bar{D}_q calculada para ruído simulado. A fig. 5.31 mostra o mapa de ativação obtido.

O mapa apresenta ativação em regiões do córtex motor primário e secundário de acordo com o estímulo aplicado. Usando um nível de significância por família de testes ($\alpha_f = 0,05$), o

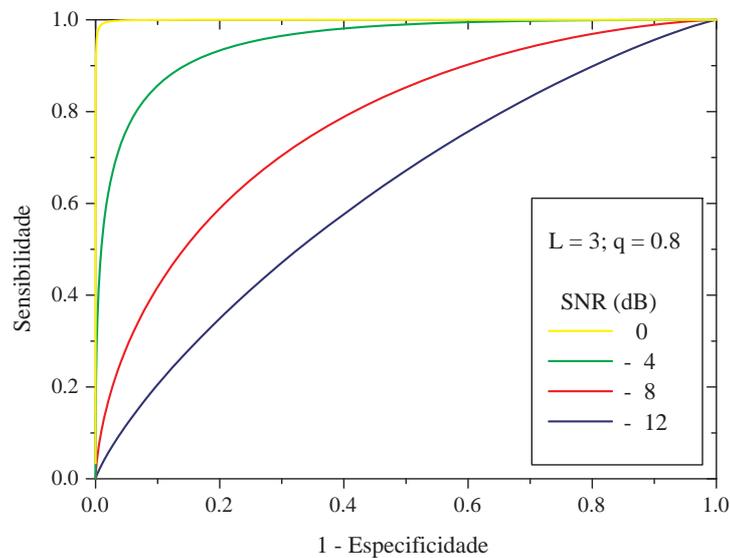


Figura 5.30: Curvas ROC para os parâmetro otimizados ($L = 3; q = 0,8$).

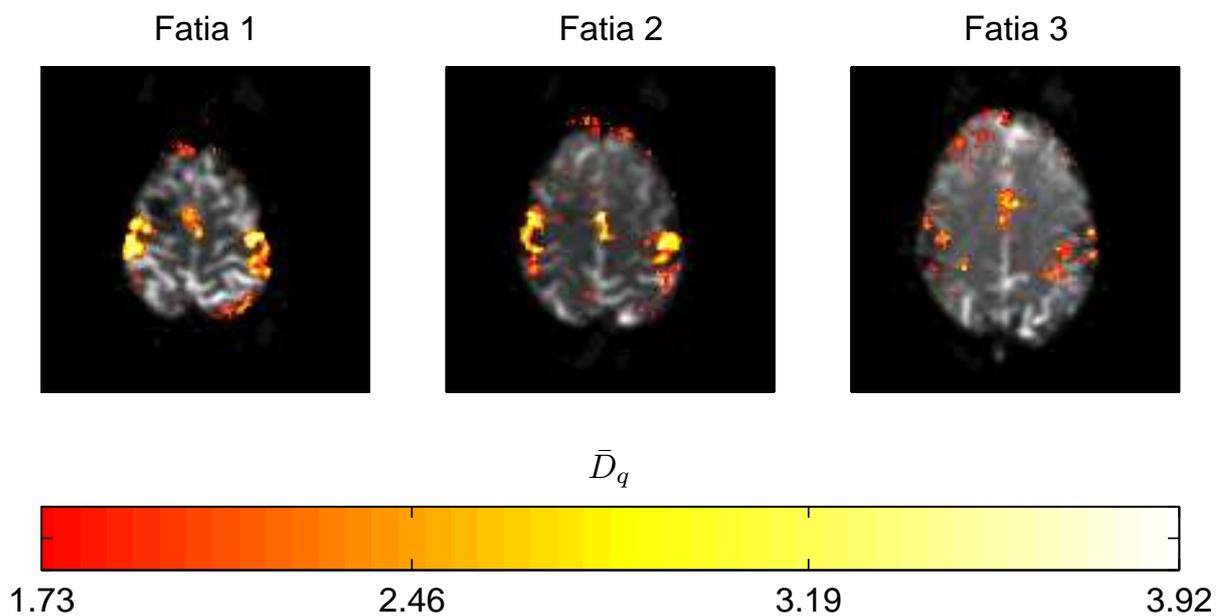


Figura 5.31: Mapa de ativação obtido a partir do método da distância de Kullback-Leibler Generalizada. Para o cálculo de \bar{D}_q , foram utilizados os parâmetros $L = 3$ e $q = 0,8$. O valor de corte $\bar{D}_q^{\alpha} \approx 1,73$ foi obtido considerando o nível de significância por família de testes ($\alpha_f = 0,05$).

valor de corte obtido pela distribuição da variável \bar{D}_q para ruído simulado foi de $\bar{D}_q^{\alpha} \approx 1,73$.

6 *Estatística Bayesiana em fMRI*

O teorema de Bayes fornece um sistema de inferência estatística adaptado a uma construção iterativa de um modelo, partindo-se de um conhecimento *a priori*. O trabalho desenvolvido por Amaral et al. [10, 11], mostra como os métodos Bayesianos podem ser aplicados em análise de sinais de fMRI. Os autores estudaram dois métodos: O método do pixel independente e o de multiescala. O primeiro trata os *voxels* (ou *pixel*) de forma independente, enquanto que o segundo utiliza informação global da imagem.

Neste capítulo abordaremos o método do pixel independente com um tratamento ligeiramente diferente no que tange à estimativa de parâmetros, quando comparado ao estudo supra citado. Outrossim, definimos como estatística do teste, a probabilidade *posteriori* $P(H_1|s_v)$ de a hipótese H_1 ser verdadeira, dado o conjunto discreto s_v correspondente ao sinal no *voxel* v . Esta é expressa em termos da probabilidade priori $P(H_1)$ de H_1 ser verdadeira e das verossimilhanças $P(s_v|H_k)$, $k = 0, 1$. Através de simulações numéricas, estimamos a distribuição das probabilidades *posteriori* para experimentos do tipo evento-relacionado. Deste modo, obtemos as curvas ROC para vários valores de SNR. Investigamos a influência do *priori* quanto ao desempenho do método e, finalmente, obtemos os mapas de ativação para um conjunto de dados reais.

6.1 O método do pixel independente

O método do pixel independente, estudado por Amaral et al. [10, 11], consiste na aplicação do teorema de Bayes para cada *voxel* v da imagem, de modo independente dos demais.

Considere que o conjunto discreto

$$\begin{aligned} s_v &= f_{s_v}(t), t = 1, \dots, Tg \\ &= f_{s_v(1), s_v(2), \dots, s_v(T)}g \end{aligned} \tag{6.1}$$

representa a série temporal no *voxel* v da imagem. De modo geral,

$$s_v(t) = b_v + A_v \cos(\omega t + \phi) + n_v(t)$$

em que b_v é uma constante para o *voxel* v , $n_v(t)$ é a função resposta herdada da seção 2.3, A_v é a amplitude do sinal e a_v é uma variável que indica a presença ($a_v = 1$) ou ausência de sinal ($a_v = 0$) e $n_v(t)$ é o ruído extraído de um ruído Normal $N(0, \sigma^2)$

As hipóteses nula e alternativa são expressas como:

$$H_0: a_v=0 \text{ (ruído);}$$

$$H_1: a_v=1 \text{ (sinal + ruído)}$$

O teorema de Bayes é usado para atualizar a opinião sobre uma hipótese ($k = 0; 1$), dado o resultado de alguma informação I relevante

$$P(H_k|I) = \frac{P(H_k) P(s_v|H_k, I)}{P(s_v|I)}$$

Como todos os termos, esta será omitida doravante, por ser conhecida

$$P(H_k) \text{ e } P(s_v|H_k)$$

... a priori, ou simplesmente *priori* (H_k)

tório de densidades de probabilidade gaussianas $N(0, \sigma^2)$, dada a independência dos eventos:

$$\begin{aligned}
 P(s_v j H_k) &\propto p(\xi_v(1), \dots, \xi_v(T) j H_k) \\
 &= \prod_{t=1}^T \left[\frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(\xi_v(t))^2}{2\sigma^2}\right)\right] \\
 &= \frac{(2\pi)^{-T/2}}{\sigma^T} \prod_{t=1}^T \exp\left\{-\frac{1}{2\sigma^2} [s_v(t) - b_v - A_v h(t) a_v]^2\right\} \\
 &= \frac{(2\pi)^{-T/2}}{\sigma^T} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^T [s_v(t) - b_v - A_v h(t) a_v]^2\right\} \quad (6.7)
 \end{aligned}$$

Este resultado nos conduz às definições das funções *likelihood* $L_0(b_v, A_v, \sigma)$ e $L_1(b_v, A_v, \sigma)$ para os casos em que H_0 é verdadeira ($a_v = 0$) e H_1 é verdadeira ($a_v = 1$), respectivamente:

$$P(s_v j H_0) \propto L_0(b_v, A_v, \sigma) = \frac{1}{\sigma^T} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^T [s_v(t) - b_v]^2\right\}, \quad (6.8)$$

$$P(s_v j H_1) \propto L_1(b_v, A_v, \sigma) = \frac{1}{\sigma^T} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^T [s_v(t) - b_v - A_v h(t)]^2\right\}. \quad (6.9)$$

A eq. (6.6) pode então ser reescrita em termos das funções *likelihood*:

$$\frac{P(H_1 j s_v)}{P(H_0 j s_v)} = \frac{P(H_1) L_1(b_v, A_v, \sigma)}{P(H_0) L_0(b_v, A_v, \sigma)}, \quad (6.10)$$

considerando-se as relações

$$P(H_0 j s_v) + P(H_1 j s_v) = 1 \quad (6.11)$$

e

$$P(H_0) + P(H_1) = 1, \quad (6.12)$$

podemos reescrever a equação em termos somente do *priori* $P(H_1)$ e do *posteriori* $P(H_1 j s_v)$, obtendo finalmente,

$$P(H_1 j s_v) = \frac{L_1(b_v, A_v, \sigma) P(H_1)}{L_0(b_v, A_v, \sigma) + [L_1(b_v, A_v, \sigma) + L_0(b_v, A_v, \sigma)]} \quad (6.13)$$

com as condições $P(H_0) \neq 1$ e $P(H_1) \neq 1$.

6.1.1 Estimativa de parâmetros

Para calcular as funções *likelihood* eqs. (6.8) e (6.9), devemos estimar os valores dos parâmetros b_v , σ , A_v . Estes valores são estimados a partir dos dados reais s_v de cada *voxel*, independentemente. Para o cálculo da linha de base b_v , calculamos a média da série temporal s_v .

$$b_v = \frac{1}{T} \sum_{t=1}^T s_v(t). \quad (6.14)$$

A amplitude A_v pode ser obtida eliminando o parâmetro σ da função *likelihood* L_1 , utilizando a regra da marginalização. Como o valor de σ é restrito à valores positivos e além disso é um parâmetro de escala, expressamos completa ignorância sobre o valor de uma escala usando o prior de Jeffreys ($1/\sigma$) [37]. Ao multiplicar este valor de *priori* pela função *likelihood*, e integrando sobre todos os valores positivos de σ , obtemos as funções *likelihood* (L_1) independente do parâmetro σ :

$$L_1(b_v, A_v) = \int_0^{+\infty} \frac{L_1(b_v, \sigma, A_v) d\sigma}{\sigma}, \quad (6.15)$$

utilizando o valor de $L_1(b_v, \sigma, A_v)$ dado pela eq. (6.9), obtemos o resultado analítico para esta integral usando o aplicativo *Mathematica* [®].

$$L_1(b_v, A_v) = 2^{(j-1+\frac{T}{2})} \left\{ \sum_{t=1}^T [s_v(t) - b_v - A_v h(t)]^2 \right\}^{(j-1+\frac{T}{2})} \Gamma\left(\frac{T}{2}\right). \quad (6.16)$$

Estimamos então o parâmetro A_v , encontrando o máximo de L_1 pela derivada,

$$\frac{\partial L_1}{\partial A_v} = 0, \quad (6.17)$$

obtendo a seguinte estimativa para A_v :

$$A_v = \frac{j-1 + \sum_{t=1}^T h(t) + \sum_{t=1}^T s_v(t)h(t)}{\sum_{t=1}^T [h(t)]^2}. \quad (6.18)$$

Estimamos o desvio padrão σ através da expressão

$$\sigma = \sqrt{\frac{\sum_{t=1}^T [s_v(t) - b_v - A_v h(t)]^2}{T - 1}}. \quad (6.19)$$

De posse de todos os parâmetros (A_v , b_v , σ), podemos obter as funções *likelihood* eqs. (6.8) e (6.9), e calcular o valor de posteriori $P(H_1 | s_v)$ pela eq. (6.13), este valor é escolhido como a estatística para testar a hipótese nula de que um *voxel* é inativo.

6.2 Estatística Bayesiana - Simulações

Para avaliar o desempenho do método do pixel independente, realizamos 10^6 simulações de séries temporais contendo apenas ruído e outras 10^6 para séries contendo sinal+ruído para os valores de SNR = -4, -12, -16 e -20 dB. Inicialmente fixamos o valor de *priori* $P(H_1) = 0,1$. Para cada série simulada, calculamos o valor de *posteriori* $P(H_1 | s_v)$ de acordo com a eq. (6.13). As distribuições de $P(H_1 | s_v)$ para os valores citados de SNR e para o ruído estão representadas nas figs. 6.1; 6.2; 6.3; 6.4. Pudemos notar uma nítida separação entre as distribuições de $P(H_1 | s_v)$ calculadas para os casos em H_0 ou H_1 verdadeiros. Para revelar a relação entre sensibilidade e especificidade do método obtemos as curvas ROC, mostradas na fig 6.5.

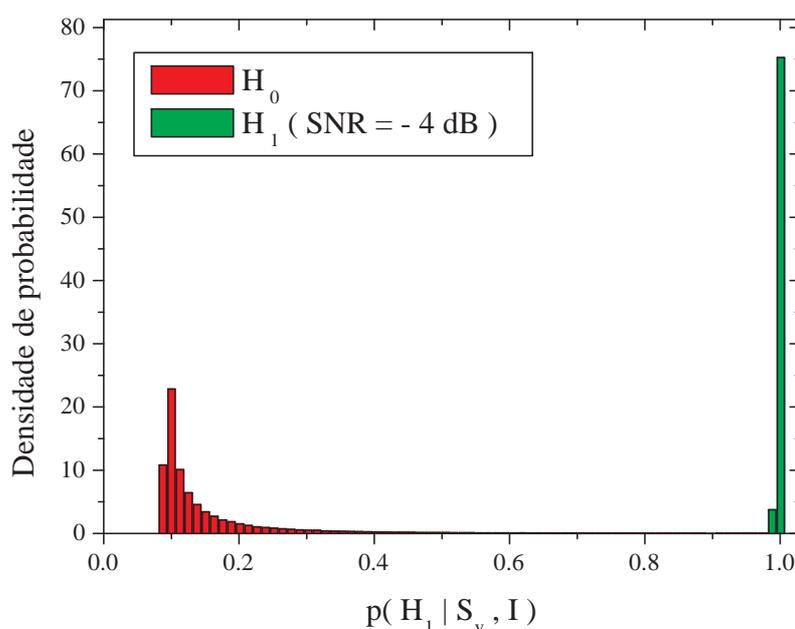


Figura 6.1: Gráficos mostrando as distribuições densidade de probabilidade da estatística $p(H_1 | s_v)$, para SNR=-4dB e para o ruído.

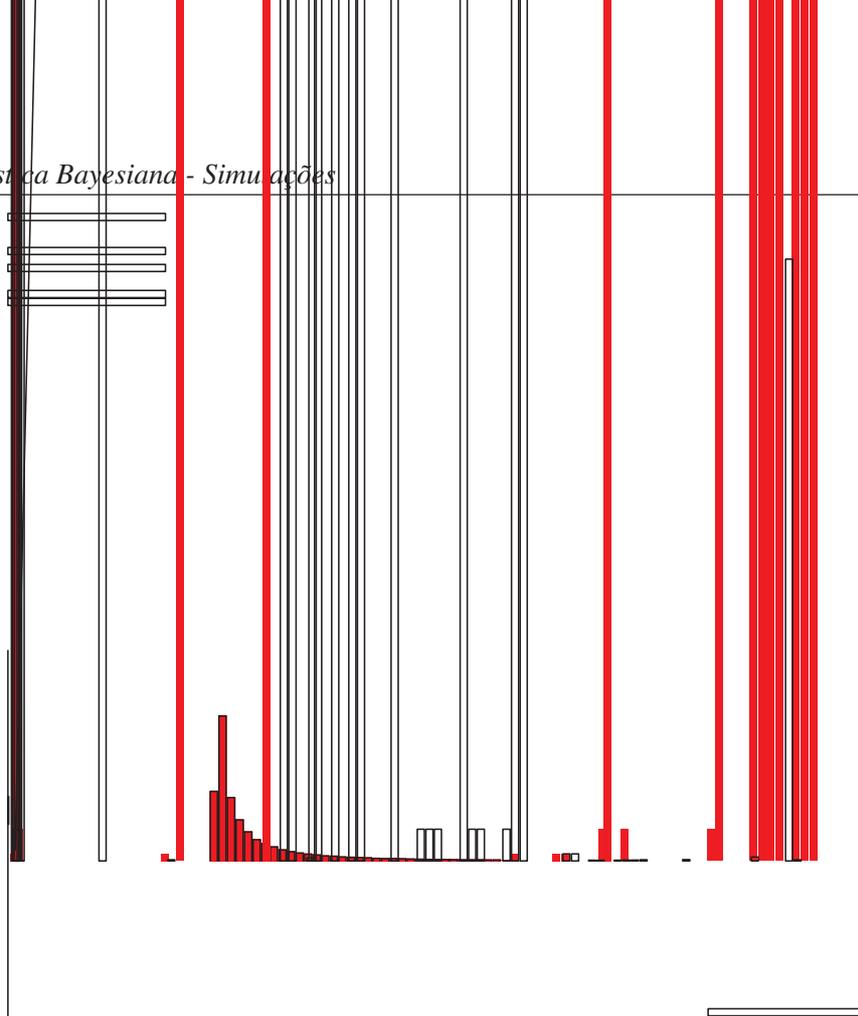


Figura 6.2: Gráficos mostrando as distribuições densidade de probabilidade da estatística $p(H_1|s_v)$, para SNR=-12dB e para o ruído.

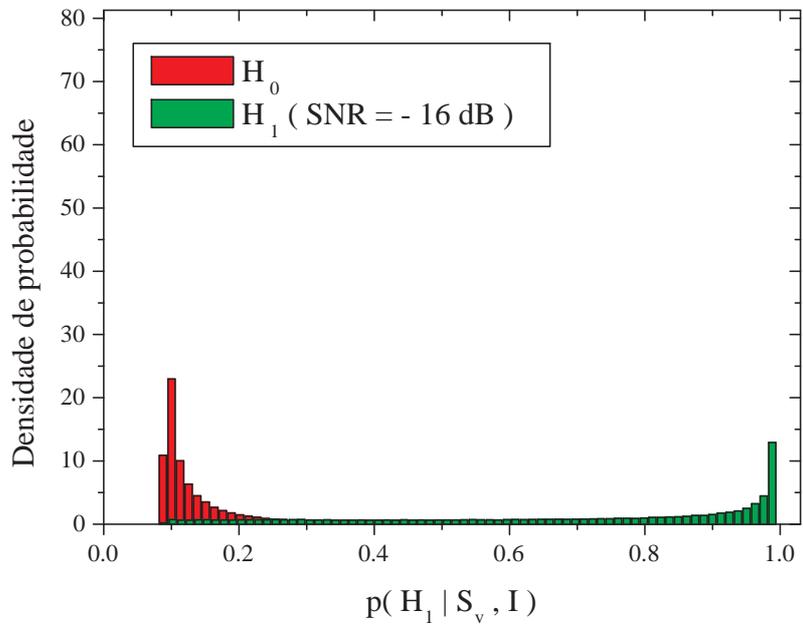


Figura 6.3: Gráficos mostrando as distribuições densidade de probabilidade da estatística $p(H_1|s_v)$, para SNR=-16dB e para o ruído.

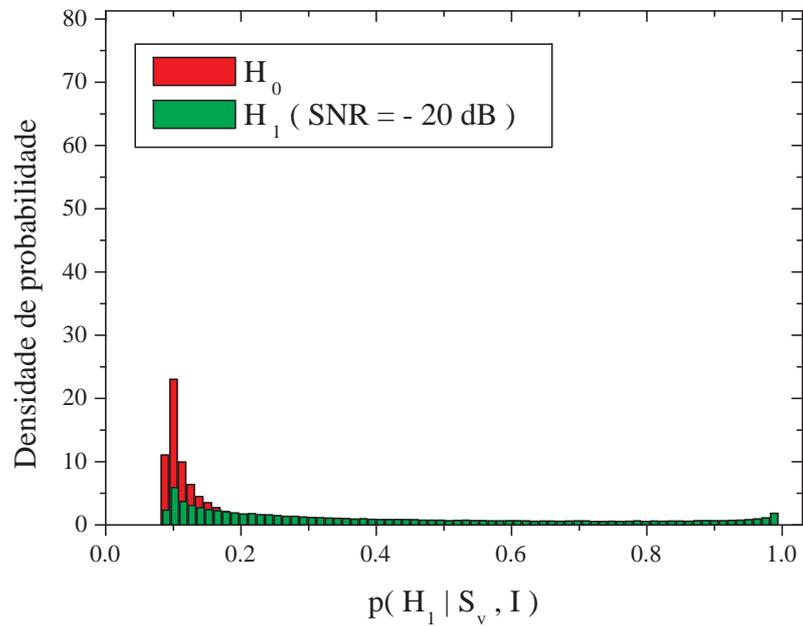


Figura 6.4: Gráficos mostrando as distribuições densidade de probabilidade da estatística $p(H_1|s_v)$, para SNR=-20dB e para o ruído.

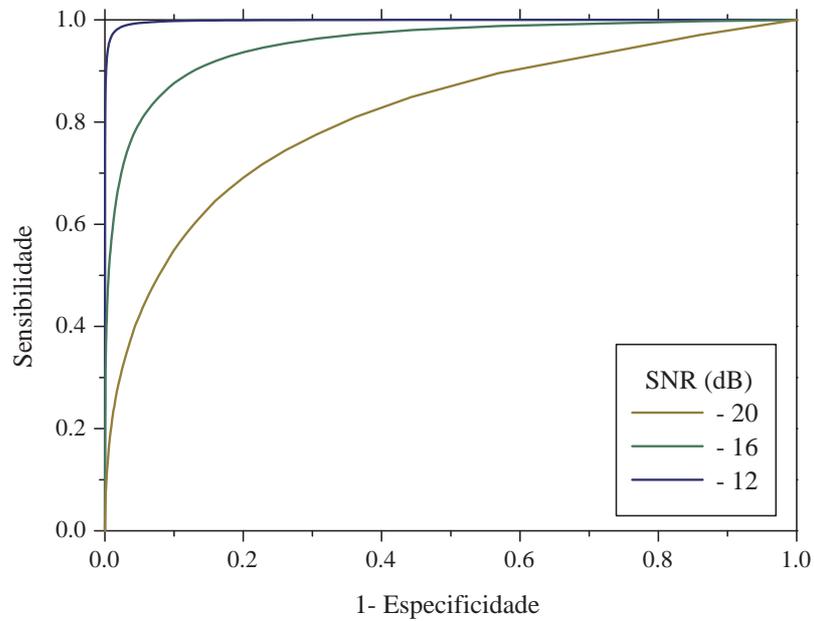


Figura 6.5: Curvas ROC para o método Bayesiano com SNR=-12, -16 e -20dB.

As curvas ROC mostram que o método é muito bom, mesmo para valores baixos de SNR ($-12dB$), a especificidade e sensibilidade são altas (≈ 1).

Obtemos as distribuições dos *posteriori* para diferentes valores de *priori* para investigar a influência desse parâmetro nas distribuições. Pelas figs. 6.6; 6.7; 6.8; 6.9, percebemos que a medida que aumentamos os valores de *priori*, os valores dos *posteriori* também aumentam, deslocando as distribuições para a direita. Porém, pelo fato desse deslocamento ocorrer da mesma maneira, tanto para a distribuição H_0 quanto para a distribuição H_1 , a sensibilidade e a especificidade do método não muda, conforme mostra a fig. 6.10 que relaciona a área ROC para cada valor de *priori*.

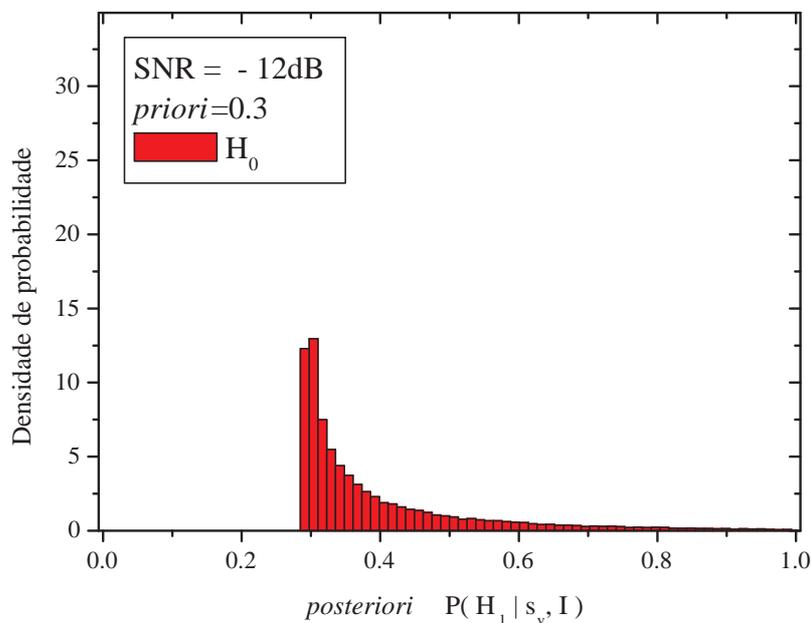


Figura 6.6: Distribuição densidade de probabilidade do *posteriori* $P(H_1 | s_v)$ para o valor de *priori* $P(H_1) = 0,3$

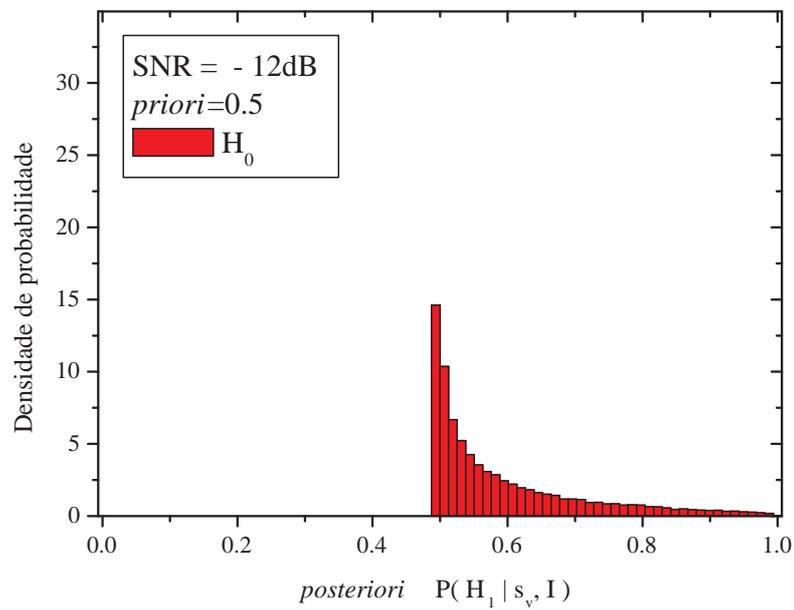


Figura 6.7: Distribuição densidade de probabilidade do *posteriori* $P(H_1 | s_v)$ para o valor de *priori* $P(H_1) = 0,5$

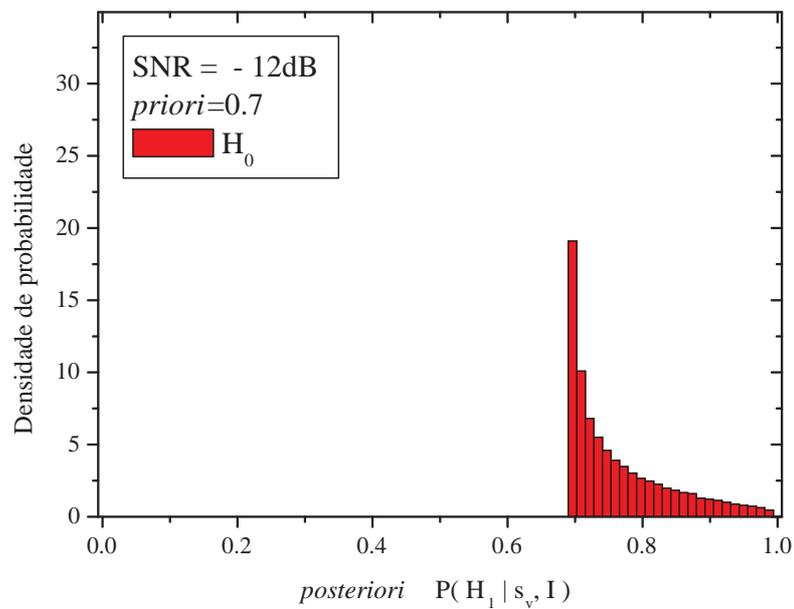


Figura 6.8: Distribuição densidade de probabilidade do *posteriori* $P(H_1 | s_v)$ para o valor de *priori* $H_1) = 0,7$



Figura 6.9: Distribuição densidade de probabilidade do *posteriori* $P(H_1|s_v)$ para o valor de *priori* $P(H_1) = 0,9$

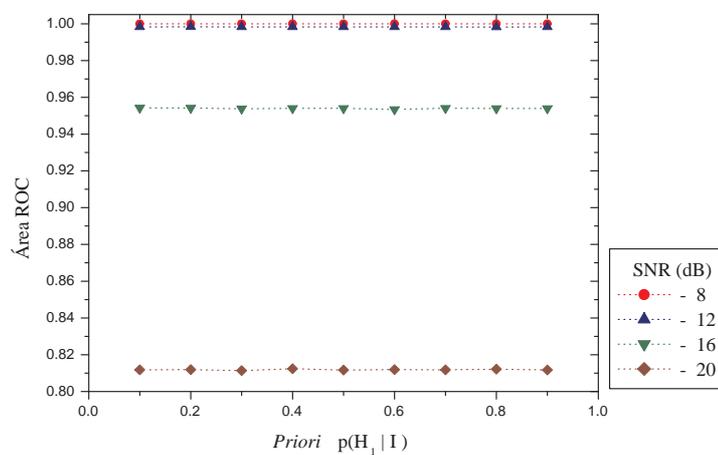


Figura 6.10: Comportamento da área ROC para (SNR= -8, -12,-16 e -20 dB) em função do *priori*.

6.3 Estatística Bayesiana - Dados reais

Utilizamos o método Bayesiano para analisar o conjunto de dados reais. Para obter o valor do nível descritivo p , utilizamos a distribuição da variável $P(H_1|s_v)$ para o *priori* $P(H_1) = 0,1$, calculada para ruído simulado. A fig. 6.11 mostra o mapa de ativação obtido usando um *priori* $P(H_1) = 0,3$. Usamos um nível de significância por família de testes ($\alpha_f = 0,05$), para tal, o

valor de corte obtido pela distribuição do ruído simulado foi de $P(H_1|s_v) \approx 0,999912$. Para verificar a influência do valor de *priori* na análise de dados reais, obtemos também o mapa de ativação para $P(H_1) = 0,8$. Neste caso, utilizamos a distribuição do *posteriori* para o *priori* $P(H_1) = 0,8$. O valor de corte neste caso foi de $P(H_1|s_v) \approx 0,999995$. Independentemente do valor de *priori*, a relação entre sensibilidade e especificidade permanece inalterada. Porém, para a análise de dados reais, deve-se utilizar a distribuição do *posteriori* $P(H_1|s_v)$ correspondente ao *priori* usado na análise.

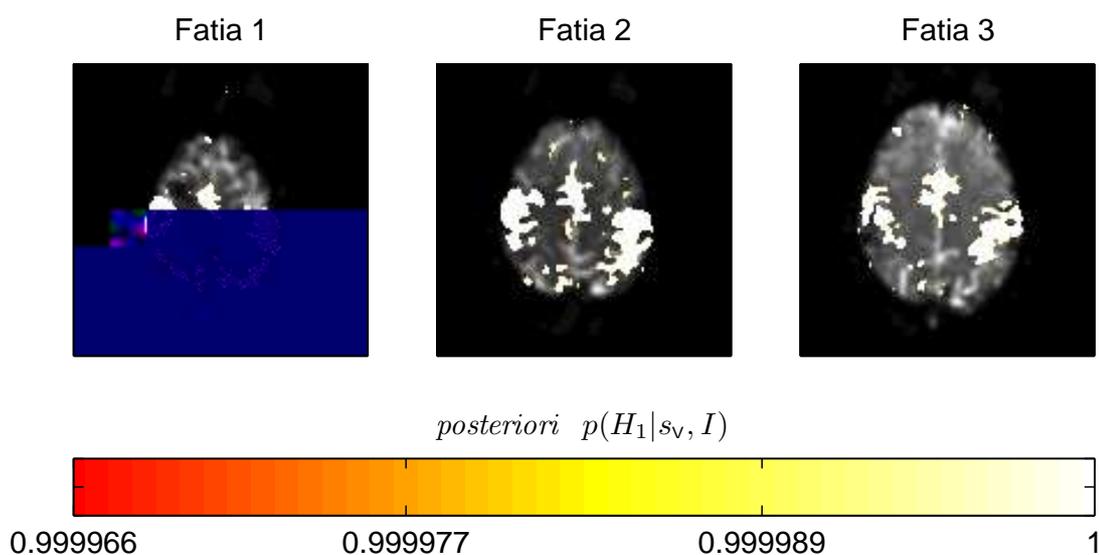


Figura 6.11: Mapa de ativação obtido a partir do método Bayesiano usando um *priori* $P(H_1) = 0,3$. O valor de corte $P(H_1|s_v) \approx 0,999966$ foi obtido considerando o nível de significância por família de testes ($\alpha_f = 0,05$).

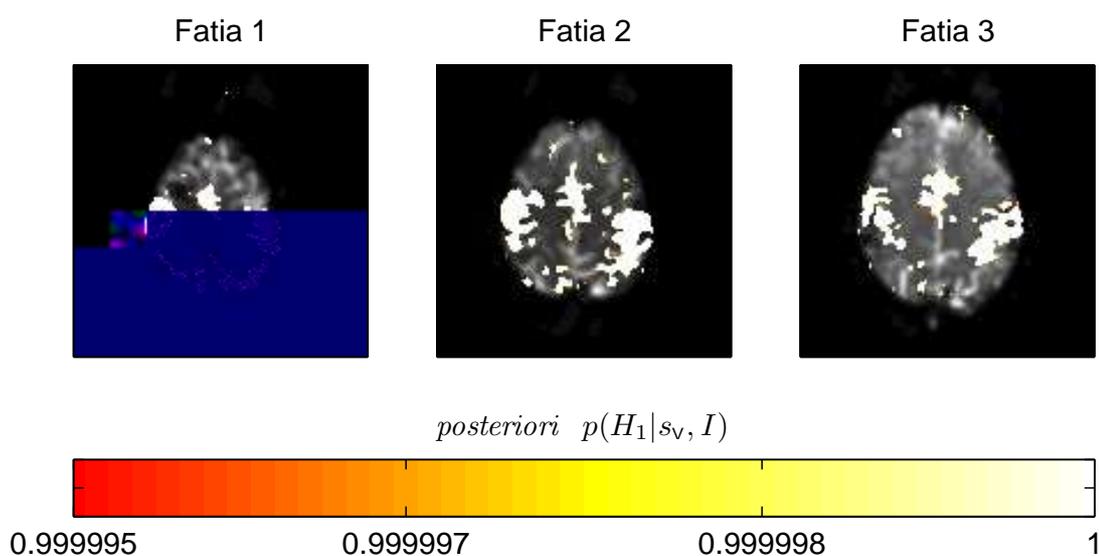


Figura 6.12: Mapa de ativação obtido a partir do método Bayesiano usando um *priori* $P(H_1) = 0,8$. O valor de corte $P(H_1|s_v) \approx 0,999995$ foi obtido considerando o nível de significância por família de testes ($\alpha_f = 0,05$).

7 *Comparação entre Métodos*

Comparamos os métodos estudados neste trabalho através do comportamento da área abaixo da curva ROC como função da SNR (variando de j 50dB até 5dB). Quanto maior o valor da área ROC, melhor é o método ao distinguir sinal de ruído, ou seja, melhor a "relação" entre sensibilidade e especificidade.

Cada método empregado tenta encontrar características na série temporal, que refletem a ativação neural. O teste t de Student fornece informação a respeito da diferença entre as médias dos valores da série temporal relativa ao estímulo e ao repouso. Caso a diferença seja significativa, o valor da estatística t é alto, caso não haja uma diferença grande, t é baixo. A inferência, portanto, se baseia no valor estatístico t , usado para testar a hipótese nula H_0 de que o voxel é inativo. Quanto maior o valor de t mais provável ser sinal.

A correlação faz uso de uma função preditora para inferir se a série temporal observada se assemelha a essa função. Ao estabelecermos o modelo de função resposta hemodinâmica (HRF) como função preditora, o coeficiente de correlação de Pearson (r) fornecerá um valor (j $1 < r < 1$) que indicará a semelhança da série observada com a HRF. Quanto mais próximo de 1 (um) for o valor de r , mais provável que esta série temporal apresente sinal.

A distância de Kullback-Leibler D e sua forma generalizada D_q quantificam a similaridade entre as distribuições p_1 e p_2 dos níveis do sinal nos períodos de estímulo e repouso, respectivamente. Quanto maiores os valores destas entropias, maior a probabilidade de se detectar o sinal correspondente à ativação neural.

O método Bayesiano do pixel independente, calcula a probabilidade *posteriori* $P(H_1|s_v, I)$, a partir de uma valor *priori* e da razão entre as funções de verossimilhança (L_1/L_0). A verossimilhança é a probabilidade de ao considerarmos uma hipótese como verdadeira (H_0 ou H_1), obter o conjunto de valores observados (série temporal - s_v). Assim, ao calcularmos a razão entre as verossimilhanças, esse valor será maior quanto mais provável for a hipótese H_1 e vice-versa. Com isso podemos inferir não só qual a hipótese mais provável, mas também o quanto ela é mais provável. Ao multiplicar essa razão pela razão dos *prioris* [$p(H_1|I)/p(H_0|I)$] ob-

7 Comparação entre Métodos

temos o valor de *posterior* inatividade do voxel.

O comportamento de

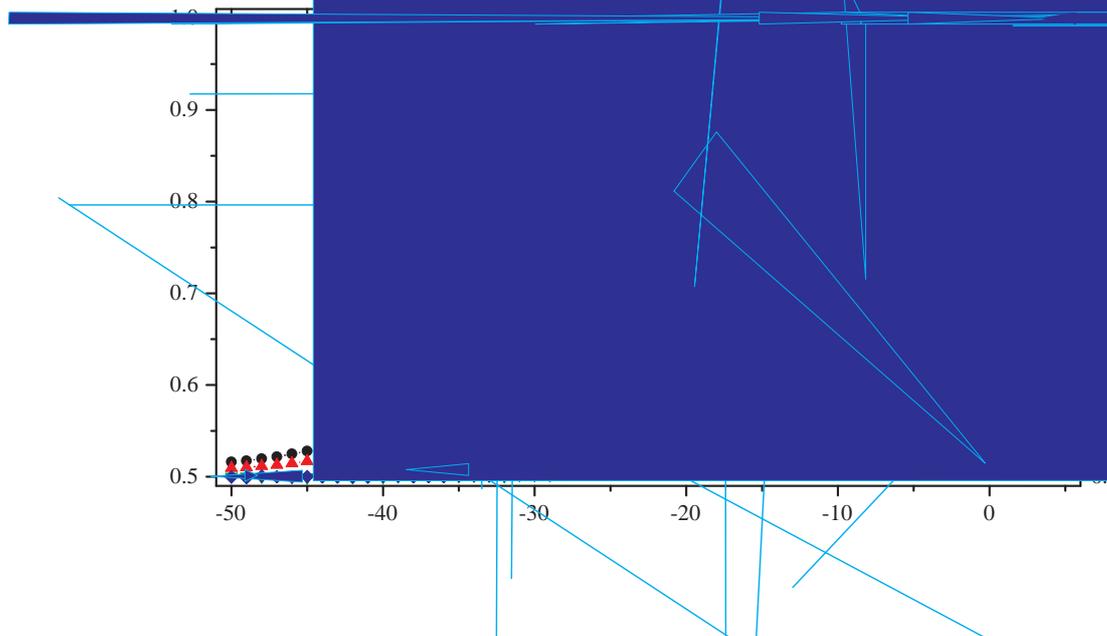


Figura 7.1: Gráfico comparando os métodos através do comportamento da área ROC.

Os métodos que necessitam de um modelo de resposta hemodinâmica (correlação e Bayesiano) apresentaram um desempenho melhor que outros métodos testados. Para valores de SNR maiores do que -15dB , não existe diferença significativa em relação ao poder dos dois testes. Porém, abaixo desse valor, os métodos Bayesianos começam a perder desempenho em relação a correlação.

Dentre os métodos que não necessitam do modelo para a HRF, o teste t apresentou melhores resultados que as distâncias de Kullback-Leibler. Isso indica que, ao menos para dados de fMRI, a simples comparação entre as médias dos valores da série temporal relativa ao estímulo e ao repouso, é mais efetiva do que a comparação entre as distribuições. Porém, uma vantagem da distância de Kullback-Leibler é que ela pode a princípio identificar diferenças entre distribuições mesmo que suas médias sejam muito próximas, o que não ocorre para o teste t .

8 *Conclusões e Perspectivas*

Neste trabalho, utilizamos conceitos de inferência estatística para aplicação e comparação de diferentes métodos de análise em fMRI.

Em especial, mostramos a viabilidade do uso da distância de Kullback-Leibler e da distância de Kullback-Leibler generalizada como métodos alternativos para a análise das séries temporais. As distribuições de probabilidade das distâncias médias de Kullback-Leibler e de sua forma generalizada (\bar{D} e \bar{D}_q) foram determinadas por simulações numéricas. A sensibilidade e especificidade dos métodos foram avaliadas pela construção das curvas ROC para diferentes valores de SNR. Foram obtidos os mapas funcionais correspondentes a dados reais provenientes de um voluntário assintomático submetido a um experimento motor de evento relacionado. Os resultados da análise das simulações e dos dados reais indicam que estas técnicas podem de fato ser uma boa opção para análise em fMRI.

A fMRI tem sido cada vez mais usada em avaliações pré-cirúrgicas de pacientes com epilepsia [38]. A função resposta hemodinâmica possui uma variabilidade maior em tecidos patológicos como regiões epileptogênicas cerebrais [39]. Desta forma, métodos que não utilizam um modelo de HRF têm a vantagem de poder detectar regiões ativas em que a forma da HRF é variável. Mais do que isso, métodos baseados em teoria da informação, como a distância de Kullback-Leibler, pode a princípio identificar diferenças entre distribuições mesmo que suas médias sejam muito próximas, o que não ocorre para o teste t, por exemplo. Além disso, o parâmetro q da distância de Kullback-Leibler generalizada pode ser utilizado na otimização do método para diferentes tipos e formas de sinais. Novos estudos utilizando sinais com diferentes características devem ser realizados.

Em outro estudo, utilizamos a inferência Bayesiana para testar a hipótese nula (H_0) de que um pixel está inativo versus a hipótese alternativa (H_1) de que o mesmo se encontra ativo. Como estatística do teste, calculamos a probabilidade a posteriori $P(H_1|s_v)$ de H_1 ser verdadeira, dada a série temporal s_v correspondente ao sinal no voxel v . Esta é expressa em termos da probabilidade *a priori* $P(H_1)$ de H_1 ser verdadeira e das verossimilhanças $P(s_v|H_k)$, $k=0,1$. Através

de simulações numéricas, estimamos a distribuição das probabilidades *a posteriori*. Obtemos as curvas características de operação (ROC) que expressam a sensibilidade e especificidade do método, para vários valores de SNR. Investigamos a influência do *a priori* quanto ao desempenho do método e verificamos que ao utilizar a inferência estatística, a escolha do *priori* não afeta este desempenho . Também obtemos os mapas de ativação para um conjunto de dados reais.

Finalmente, enfatizamos que o procedimento adotado no presente estudo pode, em princípio, ser utilizado em outros métodos.

Como perspectiva de trabalhos futuros, propomos investigar, com a presente abordagem, o método Bayesiano multiescala estudado por Amaral et al. [10, 11].

APÊNDICE A – Imagem por Ressonância Magnética Nuclear

A técnica de obtenção de imagens por ressonância magnética (IRM) é usada principalmente na imagiologia médica para visualizar a estrutura e função do corpo. Ela fornece imagens detalhadas do corpo em qualquer plano. A IRM tem contraste muito melhor para tecidos moles do que a tomografia computadorizada (TC) tornando-a especialmente útil para o diagnóstico de doenças neurológicas, musculoesqueléticas, cardiovasculares e oncológicas. Diferentemente da TC ela não usa radiação ionizante. O *scanner* cria um poderoso campo magnético, que alinha a magnetização de átomos de hidrogênio no corpo. Ondas de rádio são utilizadas para alterar o alinhamento desta magnetização. Isto faz com que os átomos de hidrogênio emitam um fraco sinal de rádio frequência que é amplificado pelo *scanner*. Este sinal pode ser manipulado por campos magnéticos adicionais de forma a obter informação suficiente para reconstruir uma imagem do corpo. Os *scanners* utilizados na medicina têm um campo magnético típico de 0,2 a 3,0 Teslas.

A.1 Princípios físicos

Partículas sub-atômicas elementares, como prótons possuem a propriedade quântica de *spin*. Núcleos tais como 1H ou ^{31}P , com um número ímpar de nucleons (prótons + nêutrons), sempre têm um *spin* não-nulo e, portanto, um momento magnético. Os núcleos com *spin* diferente de zero só adquirem valores distintos de energia, se estiverem na presença de um campo magnético. Classicamente, aceita-se que um *spin* pode ser compreendido como um momento magnético que precessiona em torno de um eixo, ver fig A.1.

Quando estes não estão sujeitos a qualquer campo magnético, o eixo ao redor do qual o *spin* precessiona é completamente aleatório, de modo que a magnetização total é nula, ver fig A.2.

Nos núcleos com *spin* ($S=1/2$), quando um campo magnético é aplicado, os *spins* passam a rodar em torno do eixo do campo paralela (estado de energia mais baixa – *spin* $(+1/2)$ ou

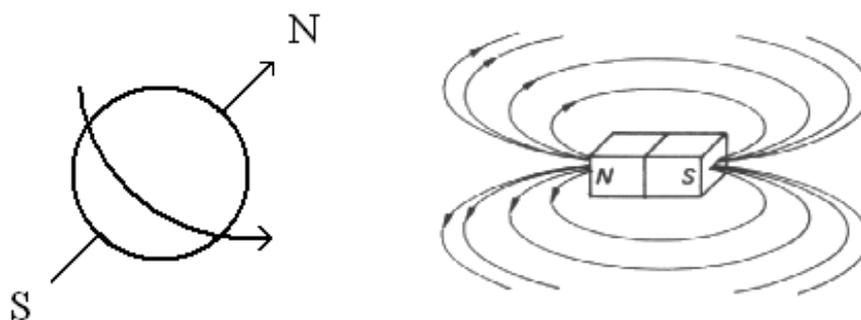


Figura A.1: Momento magnético que gira em torno de um eixo. Aproximação de um *spin* a um ímã(adapt. R.B. Lufkin, 1990).[1]



Figura A.2: Spins na ausência de campo magnético externo(adapt. R.B. Lufkin, 1990).

antiparalelamente (estado de energia mais alta – *spin* ($j = 1/2$)), conforme fig A.3.

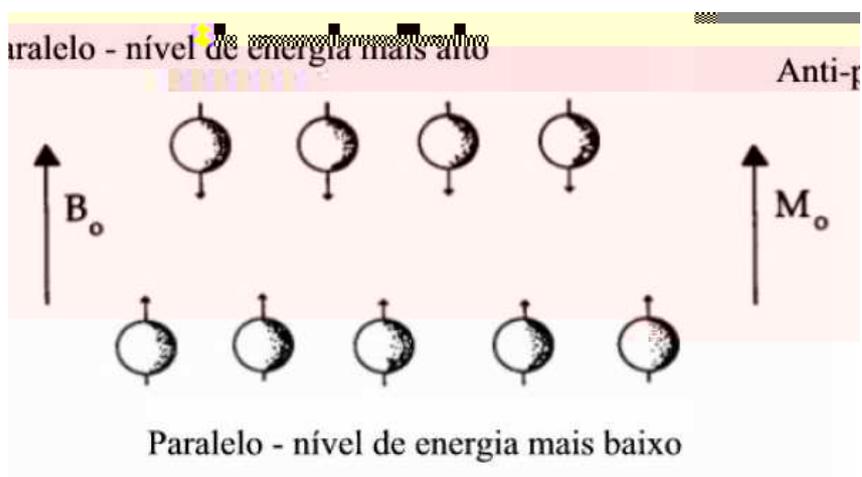


Figura A.3: *Spins* num meio onde se estabeleceu um campo magnético B_0 . M_0 tem o significado de magnetização total do meio(adapt. R.B. Lufkin, 1990).

Devido a diferença de energias dos dois estados, a população no estado de energia mais baixa é mais povoado do que o estado de energia mais alta. Por este motivo, a magnetização total deixa de ser nula e passa a ter a direção do campo z (fig. A.4), pois os *spins*, embora façam com o eixo do campo magnético externo um determinado ângulo, não estão em fase, encontrando-se aleatoriamente distribuídos sobre um cone, pelo que a sua componente xy se anula, restando apenas a componente z .

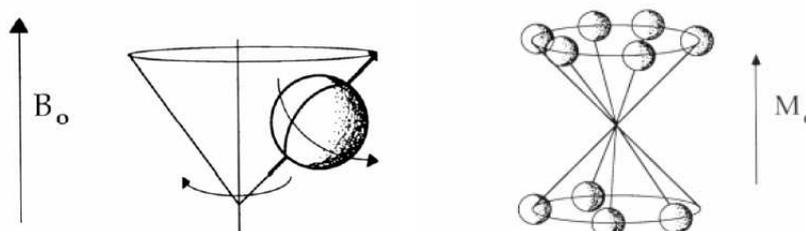


Figura A.4: Representação de *spins* a precessionarem em torno de um campo magnético externo (B_0) e magnetização total do meio (M_0) (adapt. R.B. Lufkin, 1990).

Uma das exigências para a utilização desta técnica é que os núcleos em estudo apresentem *spin* diferente de zero. Em imagens médicas, os núcleos utilizados são os de hidrogênio, uma vez que cumprem estas condições e são muito abundantes, o que permite obter um sinal de grande amplitude.

É possível definir a frequência à qual os núcleos giram (também chamada *frequência de Larmor*) como sendo proporcional a um parâmetro característico do núcleo (constante giromagnética - γ) e a amplitude do campo magnético externo aplicado (B_0):

$$f = \gamma B_0 \quad (\text{A.1})$$

Tendo em conta a ordem de grandeza da constante giromagnética dos núcleos e as amplitudes dos campos aplicados (cerca de 1T) a frequência de Larmor corresponde à faixa das rádio-freqüências (RF). Ao sujeitarmos os núcleos a um campo de rádio-freqüências interferiremos com estes, através de um fenômeno de ressonância. Estes campos colocam o movimento de precessão dos *spins* em fase. Nestas condições, a magnetização total muda de direção, passando a exibir componente xy . A amplitude e a duração dos pulsos de rádio-freqüência a que os spins são sujeitos, determinam os seus efeitos. Os chamados pulsos de 90° , são responsáveis por passar a magnetização da direção z para o plano xy (fig. A.5).

Logo após retirada do pulso de rádio-freqüência, a tendência natural do sistema é regressar ao estado inicial. Ou seja, haver reorganização do povoamento dos *spins* e a desfasagem dos mesmos. Estes dois processos são independentes e correspondem a diferentes fenômenos de

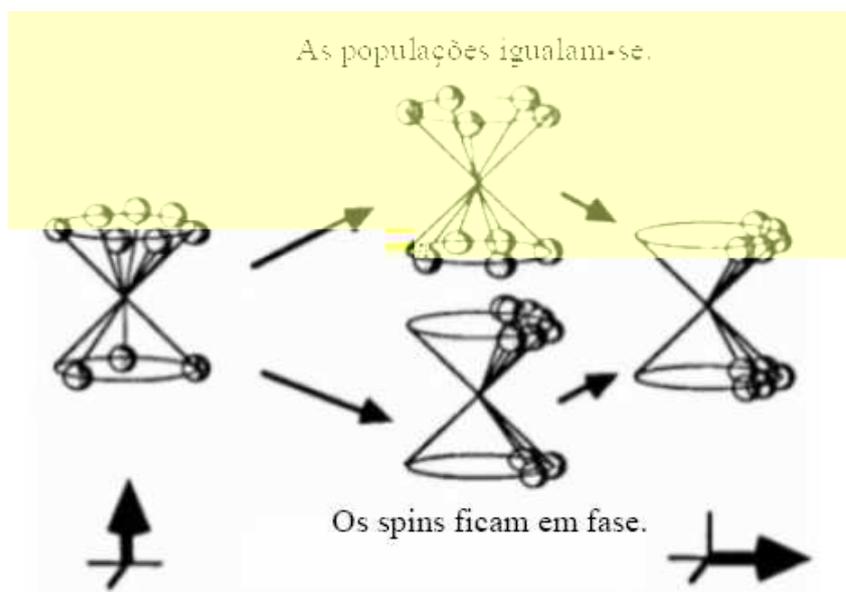


Figura A.5: Conseqüências da aplicação de um campo de rádio-freqüência na magnetização total. Exemplo de um pulso de 90° (adapt. R.B. Lufkin, 1990).

relaxação. Concentremo-nos no mecanismo de desfase dos *spins*. Se a freqüência de cada *spin* fosse exatamente a mesma, estes se manteriam em fase. Mas o que se verifica é que as freqüências de precessão dos *spins* vão ser ligeiramente diferentes. Esta circunstância deve-se, por um lado, ao fato de o campo magnético estático imposto não ser perfeitamente uniforme, apresentando heterogeneidades no espaço; por outro, o próprio meio onde os *spins* estão inseridos apresenta campos locais que são gerados pela presença de outros *spins*. Por este motivo, os *spins* irão se desfazer, a magnetização no plano *xy* vai tornando-se menor, o que corresponde a um decaimento no sinal medido (FID – *Free Induction Decay*), ver figs. A.6 e A.7.

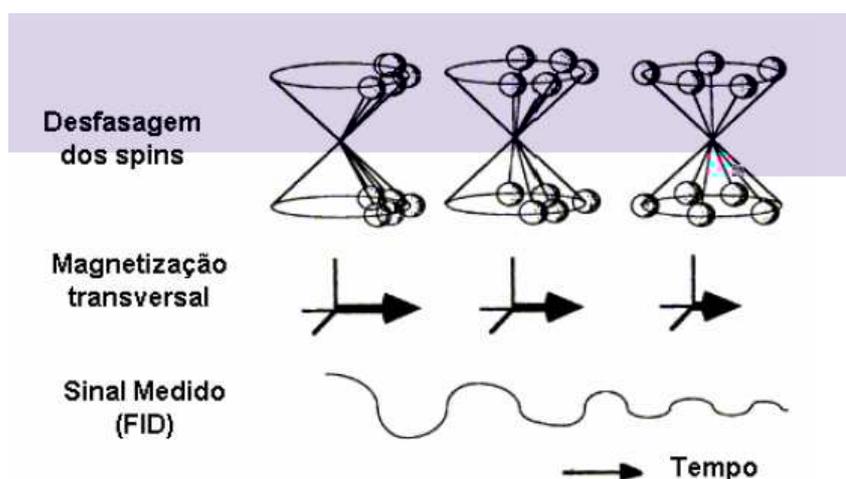


Figura A.6: 1.8 – Mecanismo de desfase dos spins, com conseqüente decaimento do sinal (adapt. R.B. Lufkin, 1990).

Verifica-se que este decaimento do sinal medido é exponencial. E, por conseguinte, é ca-

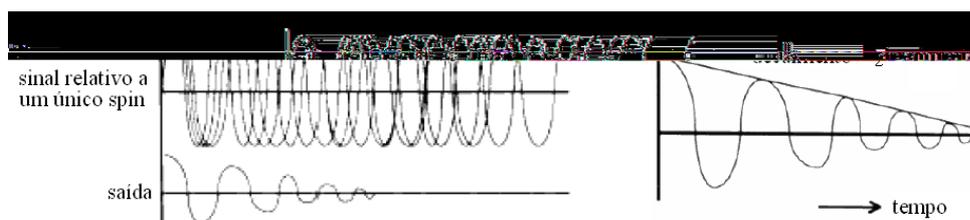


Figura A.7: Esquema do decaimento provocado pela desfasagem dos spins (adapt. R.B. Lufkin, 1990).

racterizado por uma constante de tempo. A esta constante de tempo dá-se o nome de T_2^* e é definida como o tempo necessário para que o sinal (magnetização transversal – perpendicular ao campo magnético estático) decaia para 37% do seu valor máximo. Observe-se, uma vez mais, que a grandeza T_2^* contém informação sobre as interações *spin/spin* (que é a que nos interessa, uma vez que está relacionada com a estrutura do tecido), mas está contaminada com as heterogeneidades do campo magnético estático, cujos efeitos são muito maiores do que os referentes aos campos locais, devido à presença dos *spins* vizinhos. Um pouco mais adiante, será referido um procedimento que nos permite separar estas duas componentes.

Como já foi anteriormente mencionado, existe ainda um outro mecanismo de relaxação que envolve troca de energia com o exterior, no sentido de repor as populações iniciais dos níveis energéticos de *spin*. Enquanto o pulso de rádio-frequência atua, existe excitação de *spins* que se encontravam no nível de energia mais baixo (paralelo com o campo magnético estático) para o estado de energia mais alto (anti-paralelo). A partir do momento que o pulso cessa, as populações tendem a reassumir a situação inicial, ou seja, a magnetização longitudinal (ao longo do campo magnético estático) retoma o valor inicial, ver fig. A.8.

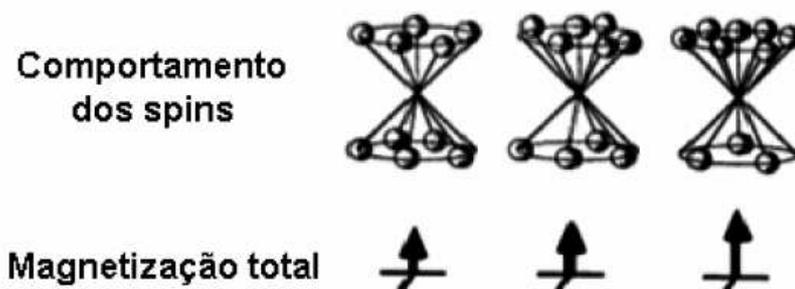


Figura A.8: Mecanismo de recuperação da magnetização longitudinal, devido à reorganização das populações de spin entre os estados energéticos, com conseqüente libertação de energia para o meio (adapt. R.B. Lufkin, 1990).

Este mecanismo ocorre através da transferência de energia para o meio e é caracterizado por uma constante de tempo T_1 , à qual se dá o nome de tempo de relaxação *spin/rede*. O tempo T_1 é, analogamente a T_2^* , o tempo que demora a magnetização longitudinal a recuperar 63% do seu

valor máximo. Como facilmente se compreende, também este parâmetro contém informação sobre os tecidos, uma vez que a maior ou menor facilidade com que os *spins* transferem energia para o meio, terá necessariamente que estar relacionada com a estrutura do meio onde estes se encontram. É desta forma que T_1 é utilizado para obter contraste entre os tecidos.

A.2 Formação da imagem

O tempo de relaxação T_2^* é encurtado pela presença de heterogeneidades do campo magnético estático que são constantes no tempo e cujo efeito se pretende anular. Alguns instantes após a ação do pulso de rádio-freqüência os *spins* já se encontram com diferentes velocidades angulares, devido as diferenças no valor do campo magnético a que cada um está sujeito. Se em determinado momento, for aplicado um novo pulso de rádio-freqüência, mas, desta vez, de 180° (ou seja, que faça a população de *spins* "virar" de 180°), então, inverte-se a posição relativa dos *spins* (os que estão girando com maior velocidade, encontram-se agora mais atrasados). Este procedimento implica, então, que passado algum tempo os *spins* se reagrupem (fiquem, novamente, em fase) sendo responsáveis por novo aumento na magnetização transversal (eco de spin), cf. fig. A.9.

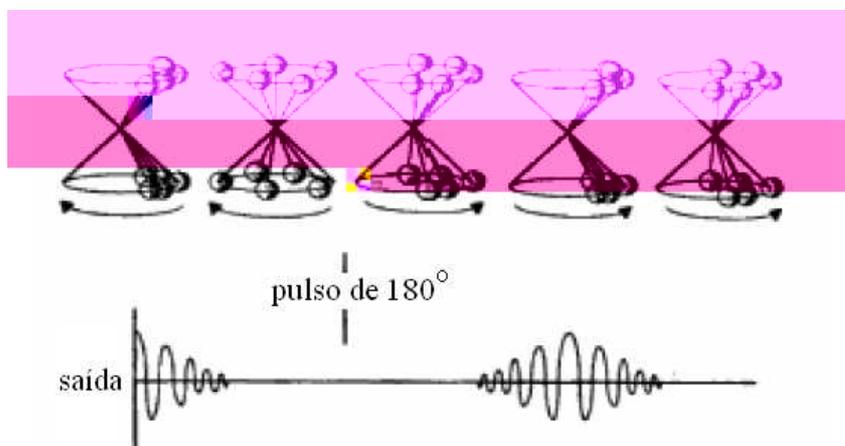


Figura A.9: Esquema da evolução da magnetização transversal com o comportamento dos *spins*, em resposta a um pulso de 180° (adapt. R.B. Lufkin, 1990).

Espera-se que a magnetização transversal seja completamente recuperada se, durante este processo, as velocidades angulares dos *spins* forem sempre constantes, o que não se passa. As interações entre *spins*, estão sujeitas a algumas oscilações pelo que os seus efeitos, ao contrario dos correspondentes as heterogeneidades do meio, prevalecem, e são responsáveis pela diminuição gradual da magnetização transversal, ver fig. A.10.

À constante de tempo que caracteriza este decaimento dá-se o nome de tempo de relaxação

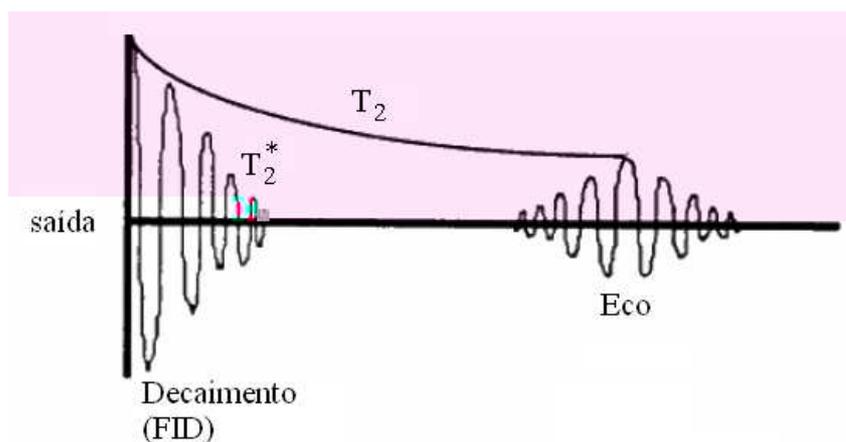


Figura A.10: Esquema explicativo sobre como, através da técnica de ecos de *spin*, é possível obter um sinal que é dependente apenas das interações entre os spins e não considera as heterogeneidades do campo magnético estático (adapt. R.B. Lufkin, 1990).

spin/spin cujo símbolo é T_2 . Este é, assim como a densidade de prótons e do tempo de relaxação *spin/rede* (T_1), um dos parâmetros responsáveis pela distinção entre os tecidos.

Atualmente, a formação de imagens de RMN implica seqüências de diversos pulsos que enfatizem os parâmetros de interesse. Uma destas seqüências, corresponde à repetição seqüencial de um pulso de 90° , seguido de vários de 180° . Nesta seqüência dá-se o nome de TE o tempo entre dois pulsos de 180° e de TR o de dois pulsos consecutivos de 90° .

A.3 Intensidade do sinal

O sinal medido em imagens de RMN é a magnetização transversal total dos tecidos. Então, a intensidade do sinal é tanto mais intensa quanto maior for a densidade de prótons (n); $I \propto n$. I também depende de T_2 segundo a eq. (A.2) e de T_1 segundo a eq. (A.3).

$$I \propto \left(e^{-TE/T_2} \right) \quad (\text{A.2})$$

$$I \propto \left(1 - e^{-TR/T_1} \right) \quad (\text{A.3})$$

Uma forma simples de compreender o comportamento da magnetização devido a T_1 e a T_2 e imaginando os casos limite em que $T_1 \gg T_2$ e $T_1 \ll T_2$. No primeiro caso o vetor magnetização começaria por girar do plano xy para o eixo z (direção do campo magnético estático) e, seguidamente, aumentaria a sua amplitude, cf. fig. ??, enquanto que no segundo, o módulo da magnetização aumentaria e só depois sofreria rotação cf. fig. ??.

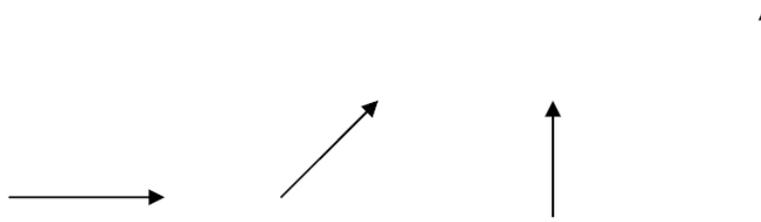


Figura A.11: Esquema explicativo do comportamento da magnetização devido a T_2 e devido a T_1 (adapt. R.B. Lufkin, 1990).



Figura A.12: Esquema simplificado do comportamento da magnetização no caso em que $T_1 \ll T_2$.

A faixa de valores de T_1 e T_2 em tecidos biológicos são, respectivamente, ($200ms < T_1 < 2000ms$) e ($50ms < T_2 < 500ms$). Assim, o sinal medido em RMN depende da densidade de prótons, do tempo de relaxação *spin/spin* e do tempo de relaxação *spin/rede* da seguinte forma:

(A.4)

Imagine-se que existem dois tecidos que possuem tempos de relaxação muito próximos, embora sejam caracterizados por densidades de prótons muito distintas. Nesse caso, sugere-se que se utilizem intervalos de tempo entre dois pulsos de 180° (TE) muito curtos e intervalos entre dois pulsos de 90° (TR) muito longos. Nesta situação, a primeira exponencial tende a saturar e a expressão entre parêntesis também. Pelo que, a forma de separar os tecidos é fundamentalmente através da densidade de prótons.

De forma semelhante, quando se pretende separar dois tecidos a partir da diferença entre os seus tempos de relaxação T_1 , deve-se utilizar TE e TR curtos. Na fig. A.13 é possível compreender este procedimento através da análise da intensidade do sinal em função do tempo, quando o tecido é caracterizado por um T_1 curto ou longo. Repare que para obter um maior

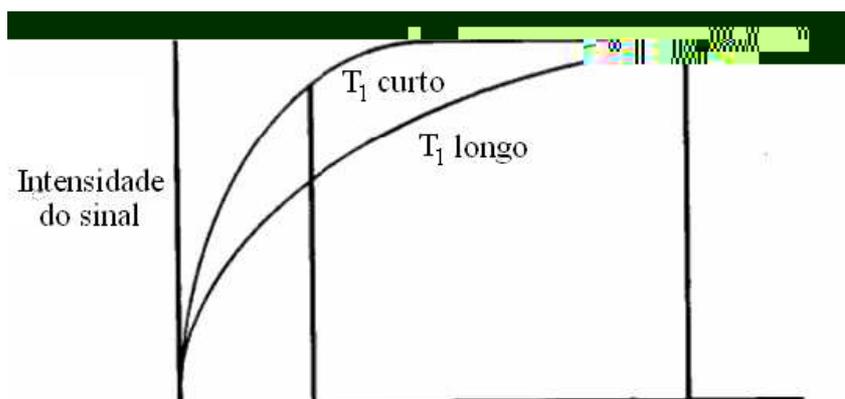


Figura A.13: Gráfico da intensidade do sinal em função do tempo para dois tecidos caracterizados por tempos de relaxação $spin/rede$ diferentes (adapt. R.B. Lufkin, 1990).

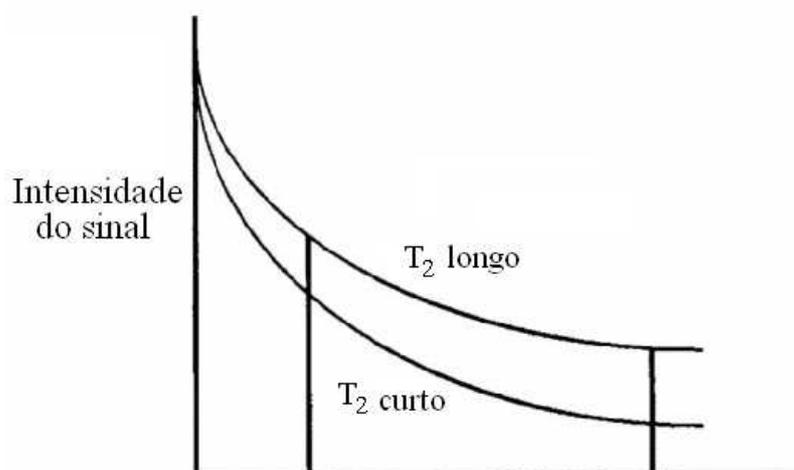


Figura A.14: Gráfico da intensidade do sinal em função do tempo para dois tecidos caracterizados por tempos de relaxação $spin/spin$ diferentes (adapt. R.B. Lufkin, 1990).

que ficou expresso nos parágrafos anteriores, é fácil admitir que regiões com uma densidade de prótons elevada aparecem mais brancas (intensidade de sinal elevada), como é o caso de tecidos gordos e fluidos. No outro extremo, encontram-se escuro as áreas com densidade de prótons baixa (calcificações, ar, tecidos fibrosos e osso cortical).

O valor de T_1 depende, como já foi descrito, da maior ou menor facilidade que o tecido tem de receber energia na faixa de rádio-freqüências adequada. Verifica-se que, enquanto a água apresenta um T_1 longo, o colesterol, por exemplo, apresenta um T_1 curto. Esta observação deve-se, fundamentalmente, ao fato de os movimentos no colesterol serem mais lentos e, por isso, mais próximos da freqüência de Larmor dos átomos de Hidrogênio. É interessante observar que, em muitas situações, a água que se encontra livre nos tecidos se liga (ainda que por ligações fracas) às fronteiras de muitas moléculas. Em tecidos em que este mecanismo ocorre, o tempo T_1 da água tende a diminuir. Na tabela A.1 estão apresentados os valores de T_1 e de T_2 para alguns

tecidos orgânicos, considerando uma frequência de 20 MHz.

	T_1 (ms)	T_2 (ms)
Sangue	900	200
Músculo	500	35
Gordura	200	60
Água	3000	3000

Tabela A.1: Valores de T_1 e T_2 para alguns tecidos biológicos.

No sangue, cujo principal constituinte é a água, o tempo de relaxação *spin/rede* é muito mais baixo do que o da água pura, porque se estabelecem as tais ligações que se referidas anteriormente, entre a água e os restantes constituintes sanguíneos.

Deste modo, para imagens com contraste em T_1 aparecem em branco tecidos como a gordura, fluidos com proteínas, moléculas lipídicas, hemorragias subagudas e a melanina. Em escuro apresentam-se regiões com neoplasmas, edemas, inflamações, fluidos puros e o líquido céfalo-raquidiano.

No tempo de relaxação *spin/spin* o fator determinante é a presença de campos magnéticos locais. Desta forma, nos sólidos e nas grandes moléculas, T_2 é, habitualmente, curto, uma vez que estas estruturas apresentam campos magnéticos intrínsecos. No extremo oposto encontram-se os fluidos cujas moléculas apresentam uma grande mobilidade e, por conseguinte, estes campos tendem para zero. Estas observações estão de acordo com a tabela A.1, onde tecidos como a água e o sangue apresentam T_2 superiores aos dos músculos, caracterizados por uma estrutura mais organizada ou das gorduras, constituídas por grandes moléculas. Assim, nas imagens em T_2 , aparecem em branco os tecidos com uma grande quantidade de água livre: neoplasmas ou inflamações e em escuro as substâncias que contêm ferro.

Na fig. A.15 é possível observar as diferenças obtidas na imagem do mesmo plano cerebral quando o contraste é feito em densidade de prótons, em T_1 e em T_2 . Esta é umas das grandes vantagens das imagens de RMN: uma estrutura que pode não ser visível com um dos contrastes, pode tornar-se extremamente nítida com outro.

A.4 Agentes de contraste

Além de apresentar um enorme potencial no tocante ao contraste, a RMN permite ainda a utilização de agentes de contraste que melhoram a visibilidade de determinado tecido. Contam-se como agentes de contraste substâncias que, devido à sua susceptibilidade magnética, interferem ao nível dos tempos de relaxação. No caso da RMN craniana, uma das substâncias mais

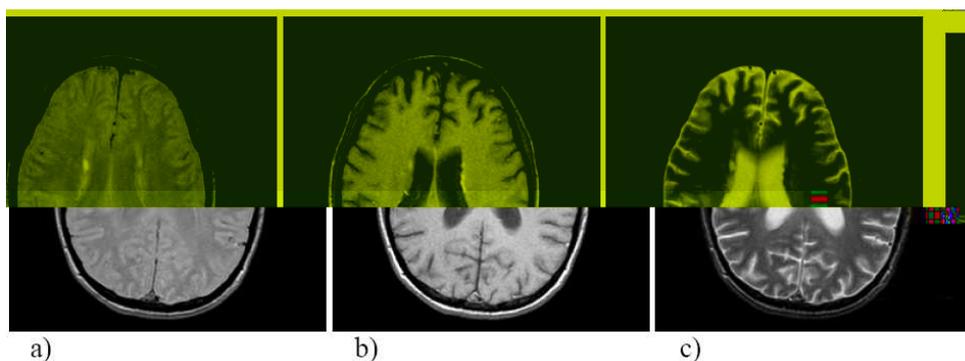


Figura A.15: Comparação entre as imagens de RMN obtidas através de contraste em: a) densidade de prótons, b) T_1 e c) T_2 . Repare-se que o osso aparece em todas as imagens a escuro (baixa densidade de prótons); o líquido céfalo-raquidiano é escuro na imagem a T_1 e branco na imagem em T_2 ; a mielina é branca nas imagens em T_1 e escura nas imagens em T_2 .

utilizadas para este efeito é o gadolínio. Esta substância, pelo fato de apresentar elétrons desemparelhados, contribui de um modo decisivo, para a alteração do tempo de relaxação *spin/spin* (T_1), visto que cria campos magnéticos locais fortes.

A.5 Gradientes de campo magnético

Para terminar, resta referir o modo como os pontos da imagem são determinados. Até aqui consideramos a magnetização total do tecido, não se referindo ao modo como a imagem é construída. Para isso é necessário um mecanismo capaz de distinguir os diversos pontos de um tecido. Esta questão é resolvida, com a aplicação de um gradiente de campo em substituição do campo magnético estático. Ao aplicar um gradiente de campo magnético numa determinada direção (x), os *spins* vão começar a girar com velocidades diferentes e, conseqüentemente, a frequência da radiação medida vai ser diferente para cada "fatia" perpendicular a x . Compreendido o significado do sinal de RMN, falta explorar o mecanismo através do qual se associa um determinado sinal a uma determinada posição, de modo a conseguir-se construir, efetivamente, uma imagem. Ao medir a magnetização transversal temos, acesso a informação referente a três parâmetros: densidade de prótons, T_1 e T_2 , relativos a todo o sistema em estudo. Iremos, em seguida, explicar a forma como e possível obter informação associada exclusivamente a um elemento de volume (comummente referido como *voxel*).

A primeira etapa tem como objectivo a escolha de uma fatia. Comece-se por compreender as implicações de introduzir um gradiente de campo ao longo de z . Ao campo magnético estático a que se sujeita o individuo, soma-se pequenos campos de diferentes intensidades segundo o eixo z . Suponha-se, então, que na origem do eixo do z se encontra aplicado um campo B_0 , Δz

adiante encontrar-se um campo de intensidade $B_0 + \delta B$ na ordenada $2\Delta z$ estará um campo de intensidade $B_0 + 2\delta B_0$ e assim sucessivamente, até cobrir todo o corpo segundo essa direção. Certamente se compreende que, nestas condições, os prótons dos átomos de Hidrogênio adquirem, em cada plano perpendicular ao eixo z , uma determinada frequência. Este fato implica que, quando se aplica um campo RF com uma frequência específica, este atua apenas sobre os *spins* que precessionam com essa frequência. Ou seja, apenas os *spins* pertencentes a uma determinada fatia são responsáveis pela magnetização transversal medida. Tendo em conta este mecanismo, a espessura de cada fatia seja estabelecida a partir da faixa de frequências contidas no pulso RF. A este respeito deve referir-se que, uma vez que existe uma dificuldade prática em estabelecer limites muito rígidos na faixa de frequências presentes no pulso RF, se fossem consideradas fatias adjacentes, os sinais medidos não corresponderiam apenas a uma única fatia. Por este motivo, opta-se por deixar uma espessura neutra entre duas fatias. I.e. uma porção de tecido sobre a qual não se obtêm informação. Uma vez selecionada a fatia, o passo seguinte passa por escolher uma linha dessa fatia.

Após a aplicação do pulso de RF, aplica-se um gradiente ao longo de y . A aplicação desse gradiente implica que os *spins* do plano escolhido, que anteriormente se encontravam em fase, adquiram frequências diferentes. Se o gradiente segundo y estiver apenas ativo por breves instantes, o resultado da sua aplicação é o aparecimento de uma diferença de fase em cada linha da fatia considerada. Deste modo, quando o gradiente segundo y cessa, os *spins* dessa fatia rodam todos com a mesma velocidade, mas, em cada linha, encontram-se numa fase distinta. Resta escolher um ponto em cada linha, para se obter informação tridimensional. Para tanto, é utilizado um gradiente de campo segundo x . Neste caso impõe-se diferentes frequências a cada ponto de cada linha. Assim, cada linha corresponde uma fase (codificação em fase) e cada coluna corresponde uma frequência (codificação em frequência).

Após a aplicação do pulso de RF, apenas uma fatia é responsável pelo sinal. Além disso, o sinal medido contém informação sobre a fase e a frequência, as quais podem ser obtidas através da análise de Fourier. Dessa forma, o sinal é decomposto em várias componentes, cada uma correspondendo a uma determinada frequência e fase; ou seja, cada uma contendo informação sobre um determinado elemento de volume (*voxel*).

Referências

- [1] R. Lufkin, W. B. Jr., and M. Brant-Zawadzki.
- [2] W. Penfield and T. C. Erickson, “Epilepsy and cerebral localization: A study of the mechanism, treatment and prevention of epileptic seizures.,” *Charles C Thomas*, 1941.
- [3] S. Ogawa, T.-M. Lee, A. S. Nayak, and P. Glynn, “Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields,” *Magnetic Resonance in Medicine*, vol. 14, pp. 68–78, 1990.
- [4] N. K. Logothetis, “The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal,” *Philosophical Transactions: Biological Sciences*, vol. 357, no. 1424, pp. 1003–1037, 2002.
- [5] A. Capurro, L. Diambra, D. Lorenzo, O. Macadar, M. T. Martin, C. Mostaccio, A. Plastino, J. Perez, E. Rofman, M. E. Torres, and J. Velluti, “Human brain dynamics: the analysis of EEG signals with Tsallis information measure,” *Physica A*, vol. 256, no. 1, pp. 235–254, 1999.
- [6] T. Yamano, “Information theory based on nonadditive information content.,” *Physical Review E*, vol. 63, pp. 46–105, 2001.
- [7] D. B. de Araujo, W. Tedeschi, A. C. Santos, J. E. Jr., U. P. C. Neves, and O. Baffa, “Shannon entropy applied to the analysis of event-related fmri time series,” *Neuroimage*, vol. 20, pp. 311–317, 2003.
- [8] W. Tedeschi, H.-P. Müller, D. de Araujo, A. Santos, U. Neves, S. Ernè, and O. Baffa, “Generalized mutual information fMRI analysis: a study of the Tsallis q parameter,” *Physica A - Statistical and Theoretical Physics*, vol. 344, pp. 705–711, 2004.
- [9] W. Tedeschi, H.-P. Müller, D. de Araujo, A. Santos, U. Neves, S. Ernè, and O. Baffa, “Generalized mutual information fMRI analysis: a study of the Tsallis q parameter,” *Physica A - Statistical and Theoretical Physics*, vol. 352, pp. 629–644, 2005.
- [10] S. R. Amaral, *Dissertação de Mestrado: Métodos Bayesianos e a priori multiescala em fMRI*. Universidade de São Paulo - Instituto de Física, 2003.
- [11] S. Amaral, S. Rabbani, and N. Caticha, “Multigrid priors for a Bayesian approach to fMRI,” *Neuroimage*, vol. 23, no. 2, pp. 654–662, 2004.
- [12] B. R. B., W. E. C., and F. L. R., “Dynamics of blood flow and oxygenation changes during brain activation: The balloon model,” *Magnetic Resonance in Medicine*, vol. 39, no. 6, pp. 855–864, 1998.

- [13] F. K. J., J. P., and T. R., “Analysis of functional mri time series,” *Human Brain Mapping*, vol. 1, no. 2, pp. 153–171, 1994.
- [14] L. N. and Z. S. L., “Non-linear fourier time series analysis for human brain mapping by functional magnetic resonance imaging,” *Journal of the Royal Statistical Society - Series C: Applied Statistics*, vol. 46, no. 1, pp. 1–29, 1997.
- [15] F. K. J., F. P., J. O., H. A. P., R. M. D., and T. R., “Event-related fmri: characterizing differential responses,” *NeuroImage*, vol. 7, no. 1, pp. 30–40, 1998.
- [16] G. G. H., “Deconvolution of impulse response in event-related bold fmri,” *NeuroImage*, vol. 9, no. 4, p. 416–429, 1999.
- [17] M. P., “Multi-planar image formation using nmr spin echoes,” *Journal of Physics C: Solid State Physics*, vol. 10, no. 3, pp. L55–L58, 1977.
- [18] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional Magnetic Resonance Imaging*. Sinauer Associates, 1 ed., 2004.
- [19] G. K. Aguirre, E. Zarahn, and M. Mark DEsposito, “A critique of the use of the kolmogorov-smirnov (ks) statistic for the analysis of bold fmri data,” *Magnetic Resonance in Medicine*, vol. 39, no. 3, pp. 500–505, 2005.
- [20] K. K., L. W., and H. E.M., “Statistical assessment of cross-correlation and variance methods and the importance of electrocardiogram gating in functional magnetic resonance imaging,” *Magnetic Resonance Imaging*, vol. 15, no. 2, pp. 169–181, 1997.
- [21] Z. H. Xu and A. K. Chan, “Encoding with frames in mri and analysis of the signal-tonoise ratio,” *Ieee Transactions on Medical Imaging*, vol. 21, no. 4, pp. 332–342, 2002.
- [22] B. MJ, “Multidimensional wavelet analysis of functional magnetic resonance images,” *Human Brian Mapping*, vol. 6, no. 5-6, pp. 378–382, 1998.
- [23] K. J. Friston, A. P. Holmes, K. J. Worsley, J.-P. Poline, C. D. Frith, and R. S. J. Frackowiak, “Statistical parametric maps in functional imaging: A general linear approach,” *Human Brain Mapping*, vol. 2, no. 4, pp. 189–210, 1995.
- [24] D. Goodenough, K. Rossmann, and L. Lusted, “Radiographic applications of receiver operating characteristic (roc) curves.,” *Radiology*, vol. 110, pp. 89–95, 1974.
- [25] J. A. Hanley and B. McNeil, “A method of comparing the areas under receiver operating characteristic curves derived from the same cases.,” *Radiology*, vol. 148, pp. 839–843, 1983.
- [26] N. J. Salkind, *Encyclopedia of Measurement and Statistics*. Sage Publications, Inc, 1 ed., 2006.
- [27] B. P. A., J. A., W. E. C., and H. J. S., “Processing strategies for time-course data sets in functional mri of the human brain,” *Magnetic Resonance in Medicine*, vol. 30, no. 2, pp. 161–173, 1993.
- [28] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, 2nd ed., 1992.

- [29] H. D.A., O. J.M., and D. M., “Variation of bold hemodynamic responses across brain regions and subjects and their effects on statistical analyses.,” *Neuroimage*, vol. 21 (4), 1639-1651, no. 4, pp. 1639–1651, 2004.
- [30] S. D. A., P. K. K., W. K. D., C. B., and B. R. W., “Comparison of hemodynamic response nonlinearity across primary cortical areas,” *Neuroimage*, vol. 22, no. 3, pp. 1117–1127, 2004.
- [31] M. J. Sturzbecher, *Dissertação de Mestrado: Detecção e caracterização da resposta hemodinâmica pelo desenvolvimento de novos métodos de processamento de imagens funcionais por ressonância magnética*. Universidade de São Paulo - FFCLRP, 2006.
- [32] C. E. Shannon, “A mathematical theory of communication,” *The Bell System Technical Journal*, vol. 27, p. 379–423;623–656, 1948.
- [33] C. Tsallis, “Possible generalization of Boltzmann-Gibbs statistics,” *Journal of Statistical Physics*, vol. 52, no. 1-2, pp. 479–487, 1988.
- [34] T. Cover and J. Thomas, *Elements of information theory*. Wiley-Interscience, 1991.
- [35] C. Tsallis, “What are the numbers that experiments provide?,” *Química Nova*, vol. 17, no. 6, pp. 468–471, 1994.
- [36] E. P. Borges, “On a q-generalization of circular and hyperbolic functions,” *Journal of Physics A: Mathematical and General*, vol. 31, pp. 5281–5288, 1998.
- [37] H. Jeffreys, *Theory of Probability*. Oxford University Press, USA, 3 ed., 1983.
- [38] J. A. Detre, “fmri: Applications in epilepsy,” *Epilepsia*, vol. 45, no. s4, pp. 26–31, 2004.
- [39] C.-G. Bénar, D. W. Gross, Y. Wang, V. Petre, B. Pike, F. Dubeau, and J. Gotman, “The bold response to interictal epileptiform discharges,” *NeuroImage*, vol. 17, no. 3, pp. 1182–1192, 2002.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)