

2016

Instituto Politécnico de Coimbra

INSTITUTO SUPERIOR DE ENGENHARIA DE COIMBRA

# Segmentação Automática de Tecidos Cerebrais

**MESTRADO EM INSTRUMENTAÇÃO BIOMÉDICA**

AUTOR | João Nuno Loução Mestre

ORIENTADORES |

Prof. Doutora Deolinda Maria Lopes Dias Rasteiro

Prof. Doutor Carlos Alberto Batista da Silva

Coimbra, Novembro 2016





Departamento de Engenharia Eletrotécnica  
Departamento de Física e Matemática

---

## **Segmentação Automática de Tecidos Cerebrais**

Dissertação apresentada para a obtenção do grau de Mestre em  
Instrumentação Biomédica

**Autor**

**João Nuno Loução Mestre**

**Orientador**

**Prof. Doutora Deolinda Maria Lopes Dias Rasteiro**

Instituto Superior de Engenharia de Coimbra

**Prof. Doutor Carlos Alberto Batista da Silva**

Universidade do Minho

**Coimbra, Novembro, 2016**



## **AGRADECIMENTOS**

Gostaria de agradecer aos meus orientadores, à professora Deolinda Rasteiro e ao professor Carlos Silva, pela oportunidade que me deram para explorar este tema, por todo o apoio prestado, dedicação, sugestões ao longo de todo este trabalho.

E agradeço à minha família, amigos e colegas todo o apoio e motivação dado ao longo do trabalho.



## RESUMO

No âmbito da cadeira de Projeto do Mestrado em Instrumentação Biomédica foi proposto o trabalho de segmentação automática de tecidos cerebrais de imagens de ressonância magnética para ser um classificador num sistema de segmentação já existente.

O trabalho teve como objetivo principal obter um modelo estatístico de previsão das intensidades dos pixels para os diferentes tipos de tecidos cerebrais, através de modelos de mistura Gaussianas por processos de *Dirichlet*, das imagens da ressonância magnética. Com isso, foi feito um levantamento dos modelos estatísticos de previsão, quer paramétricos quer não-paramétricos, onde o modelo de mistura Gaussianas por processos de *Dirichlet* se insere. Também foi estudado o modelo de mistura Gaussianas por processos de *Dirichlet* apresentando a sua estrutura e o seu procedimento para calcular a distribuição de previsão dos dados observados.

O trabalho experimental realizado apresenta a capacidade do modelo de mistura Gaussianas por processos de *Dirichlet* fazer a previsão de dados amostrados a partir de distribuições conhecidas, onde revela que o modelo consegue estimar corretamente a distribuição preditiva dos dados e até para amostras de misturas de distribuições conhecidas. Uma vez conhecida a aptidão do modelo de mistura Gaussianas por processos de *Dirichlet* para estimar a distribuição preditiva dos dados, foi realizada a sua implementação na segmentação de tecidos cerebrais. O resultado obtido da segmentação feita por modelos de mistura Gaussianas por processos de *Dirichlet* para os tecidos cerebrais, em comparação com um método de previsão paramétrico, indica que a segmentação por modelos de mistura Gaussianas por processos de *Dirichlet* obtém melhores resultados embora com limitações. Verificando essas limitações, descobrimos que a sua causa se deve à sobreposição de intensidades dos tecidos cerebrais. Com isso, alterámos o modelo para que este se ajustasse às intensidades dos tecidos cerebrais a fim de diminuir as limitações. No caso do tecido líquido cefalorraquidiano, modificámos o modelo para trabalhar com a distribuição Rician. Observou-se que com essa alteração que a segmentação proposta melhorou com uma pequena diferença. No caso do tecido substância branca, modificámos o modelo para trabalhar com novas distribuições: *Genlogistic*, *Powernorm*, *Log Gama*, *Gumbel* enviesada à esquerda e *Johnson Su*. Apenas o modelo que trabalha com a distribuição *Johnson Su* é que consegue melhorar a segmentação dos tecidos cerebrais embora com uma diferença mínima.

Palavras-chave: processos de *Dirichlet*; modelo de mistura Gaussianas; segmentação; ressonância magnética; tecidos cerebrais



## ABSTRACT

Under the chair of Biomedical Instrumentation Master's Project was proposed to a project automatic segmentation of brain tissue of images of magnetic resonance to be a classifier on an existing segmentation system.

The work aimed to obtain a statistical model for predicting the intensities of the pixels for the different types of brain tissue, using Gaussian mixture models by Dirichlet processes, magnetic resonance images. Thus, a survey of statistical forecasting models was done either parametric or non-parametric, where the Gaussian mixture model by Dirichlet process is inserted. It also studied the Gaussian mixture model by Dirichlet processes having its structure and its procedure to calculate the expected distribution of the observed data.

The experimental work shows the ability of the Gaussian mixture model by Dirichlet process to forecasting data sampled from known distributions, which shows that the model can correctly estimate the predictive data distribution and even for known distributions mixtures samples. Once known the ability of the Gaussian mixture model by Dirichlet processes to estimate the predictive distribution of the data was carried out its implementation in the segmentation of brain tissue. The result of the segmentation performed by Gaussian mixture models by Dirichlet processes for the brain tissues, compared to a parametric prediction method, indicates that targeting Gaussian mixture models by Dirichlet process achieves better results albeit with limitations.

Checking these limitations, we found that the cause of these limitations is due to the overlap intensities of brain tissue. Therefore, we modify the model to fit the intensities of the brain tissue to decrease these limitations. In the case of cerebrospinal fluid tissue, we modified the model to work with Rice distribution. It was observed that with this change, the proposed segmentation improved with a small difference. And in the case of white matter tissue, we modified the model to work with new distributions: Genlogistic, Powernorm, Log Gamma, Gumbel skewed to the left and Johnson Su. Only the model that works with the Johnson Su distribution is that it can improve the segmentation of brain tissues although with a minimum difference.

Keywords: Dirichlet process; Gaussian mixture model; segmentation; magnetic resonance; brain tissues



## ÍNDICE

<b>AGRADECIMENTOS</b> .....	<b>i</b>
<b>RESUMO</b> .....	<b>iii</b>
<b>ABSTRACT</b> .....	<b>v</b>
<b>ÍNDICE</b> .....	<b>vii</b>
<b>ÍNDICE DE FIGURAS</b> .....	<b>x</b>
<b>ÍNDICE DE TABELAS</b> .....	<b>xii</b>
<b>SIMBOLOGIA</b> .....	<b>xiv</b>
<b>ABREVIATURAS</b> .....	<b>xviii</b>
<b>1. INTRODUÇÃO</b> .....	<b>1</b>
1.1. Enquadramento .....	1
1.2. Objetivo.....	2
1.3. Organização da dissertação .....	2
<b>2. CONCEITOS FUNDAMENTAIS</b> .....	<b>3</b>
2.1. Estatística Bayesiana.....	3
2.2 Distribuição de Dirichlet.....	4
2.3 Distribuição Riciana.....	5
2.4 Distribuição Log Gama.....	5
2.5 Distribuição Johnson Su .....	6
2.6 Distribuição Normal ou Gaussiana .....	6
2.7 Distribuição T-student multivariável.....	7
2.8 Distribuição Gumbel enviesada à esquerda .....	7
2.9 Distribuição Genlogistic .....	7
2.10 Distribuição PowerNorm .....	8
2.11 Distribuição Multinomial .....	8
2.12 Função Delta Dirac .....	8
2.13 Processos de Dirichlet.....	9
2.13.1 Esquema Urna de <i>Pólya</i> .....	9
2.13.2 Definição construtiva dos processos de <i>Dirichlet</i> .....	10
<b>3. ESTADO DA ARTE</b> .....	<b>11</b>
3.1 Métodos Paramétricos.....	11
3.2 Métodos Não Paramétricos .....	12
3.3 Sumário .....	14

---

<b>4. MODELAÇÃO ATRAVÉS DE MODELOS DE MISTURA POR PROCESSOS DE DIRICHLET.....</b>	<b>15</b>
4.1 Modelo de mistura por processos de Dirichlet .....	15
4.2 Inferência variacional para misturas .....	17
4.3 Mean-field - Inferência variacional na família das exponenciais .....	19
4.4 Misturas por processos de Dirichlet.....	19
4.4.1 Algoritmo de coordenadas ascendente.....	20
4.5 Limitação do modelo de mistura por processos de Dirichlet.....	22
4.6 Sumário .....	23
<b>5. TRABALHO EXPERIMENTAL .....</b>	<b>24</b>
5.1 Exemplificação dos modelos de mistura Gaussianas por processos de Dirichlet....	24
5.2 Procedimento do cálculo do modelo de mistura Gaussianas por processos de Dirichlet .....	26
5.3 Aplicação do modelo de mistura Gaussianas em imagens cerebrais de ressonância magnética .....	30
5.3.1 Base de Dados <i>MRBrainS</i> .....	30
5.3.2 Pré-processamento .....	31
5.3.3 Histogramas das intensidades dos tecidos cerebrais .....	35
5.3.4 Modelação dos dados através de modelos de mistura Gaussianas por processos de Dirichlet.....	37
5.4 Modificação dos modelos de mistura Gaussianas por processos de Dirichlet.....	39
5.4.1 Modificação do modelo para o líquido cerebrospinal.....	41
5.4.2 Modificação do modelo para a substância branca .....	44
5.5 Resultados .....	48
5.5.1 Métricas de Avaliação .....	48
5.5.2 Matriz de Confusão.....	49
5.5.3 Método de Avaliação.....	49
5.5.4 Avaliação do método <i>Atropos</i> .....	50
5.5.5 Avaliação da segmentação dos tecidos cerebrais por modelos de mistura Gaussianas por processos de <i>Dirichlet</i> .....	51
5.5.6 Avaliação da segmentação dos tecidos cerebrais por combinação de diferentes modelos de mistura por processos de <i>Dirichlet</i> para o caso do tecido líquido cerebrospinal.....	55
5.5.7 Avaliação da segmentação dos tecidos cerebrais por combinação de diferentes modelos de mistura por processos de <i>Dirichlet</i> para o caso do tecido substância branca.....	60
5.6 Sumário .....	80
<b>6. CONCLUSÃO E TRABALHO FUTURO.....</b>	<b>81</b>

---

---

6.1 Conclusões .....	81
6.2 Trabalho Futuro.....	81
<b>7. REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>82</b>

## ÍNDICE DE FIGURAS

Figura 1 – Funções densidade da probabilidade da distribuição <i>Dirichlet</i> .....	4
Figura 2 – Função densidade da probabilidade da distribuição Riciana.....	5
Figura 3 – Função densidade de probabilidade da distribuição <i>Log Gama</i> .....	5
Figura 4 – Função densidade de probabilidade da distribuição <i>Johnson Su</i> .....	6
Figura 5 – Função densidade de probabilidade da distribuição Normal.....	6
Figura 6 – Função densidade de probabilidade da distribuição <i>Gumbel</i> enviesada à esquerda .....	7
Figura 7 – Função densidade de probabilidade da distribuição <i>GenLogistic</i> .....	7
Figura 8 – Função densidade de probabilidade da distribuição <i>Powernorm</i> .....	8
Figura 9 – Representação da função <i>Delta Dirac</i> .....	8
Figura 10 – Representação da divisão do segmento da definição construtiva dos processos de <i>Dirichlet</i> .....	10
Figura 11 - Representação gráfica do modelo de mistura por processos de <i>Dirichlet</i> numa família exponencial.....	17
Figura 12 - Evolução do modelo de mistura por processos de <i>Dirichlet</i> .....	24
Figura 13 - Evolução do modelo de mistura por processos de <i>Dirichlet</i> de uma mistura de 2 distribuições.....	25
Figura 14 - Evolução do modelo de mistura por processos de <i>Dirichlet</i> de uma mistura de 3 distribuições.....	25
Figura 15 - Exemplo de imagens da base de dados <i>MRBrainS</i> .....	30
Figura 16 - Normalização das intensidades.....	33
Figura 17 – Comparação da mesma imagem de ressonância magnética antes e depois do pré-processamento.....	35
Figura 18 - Histograma das intensidades de cada tecido cerebral.....	35
Figura 19 - Teste de adaptação da distribuição Gaussiana e <i>T-student</i> .....	36
Figura 20 - Exemplo para cada tipo de tecido cerebral a probabilidade preditiva dos dados.....	37
Figura 21 - Exemplo de mapa de probabilidade de um sujeito.....	37
Figura 22 – Exemplo dos valores obtidos numa secção do mapa de probabilidades do tecido líquido cerebrospinal.....	38
Figura 23 - Exemplo da segmentação dos tecidos cerebrais por modelos de mistura Gaussianos por processos de <i>Dirichlet</i> .....	39
Figura 24 - Demonstração da sobreposição das intensidades para os diferentes tecidos cerebrais.....	40

Figura 25 - Demonstração da sobreposição das probabilidades preditivas dos tecidos cerebrais .....	40
Figura 26 - Comparação do teste de adaptação da distribuição gaussiana e <i>t-student</i> com os histogramas do líquido cefalorraquidiano e da substância cinzenta .....	41
Figura 27 - Comparação entre o histograma do líquido cefalorraquidiano com um exemplo de uma distribuição Riciana .....	41
Figura 28 - Comparação do teste de adaptação das distribuições às intensidades do líquido cefalorraquidiano .....	42
Figura 29 - Exemplo das probabilidades preditivas dos tecidos cerebrais .....	42
Figura 30 – Comparação entre segmentações de tecidos cerebrais .....	43
Figura 31 - Comparação do teste de adaptação da distribuição gaussiana e <i>t-student</i> com os histogramas da substância cinzenta e da substância branca .....	44
Figura 32 - Comparação do histograma da substância branca .....	45
Figura 33 - Testes de adaptação das distribuições às intensidades correspondentes à substância branca .....	45
Figura 34 - Comparação entre as probabilidades preditivas dos diferentes modelos .....	46
Figura 35 - Segmentação por diferentes modelos .....	47
Figura 36 – Segmentação pelo método <i>Atropos</i> .....	51
Figura 37 - Exemplos dos defeitos da segmentação por modelos de mistura Gaussiana com a comparação das imagens T1 e a marcação do médico .....	54
Figura 38 – Comparação da segmentação por modelos de mistura Gaussiana por processos de <i>Dirichlet</i> com a segmentação pela abordagem Rician, onde ambos têm marcações das zonas mal classificadas, e com a marcação do médico .....	59
Figura 39 - Comparação da segmentação por modelos de mistura Gaussiana por processos de <i>Dirichlet</i> com a segmentação pela abordagem <i>Johnson Su</i> , onde ambos têm marcações das zonas mal classificadas, e com a marcação do médico .....	79

## ÍNDICE DE TABELAS

Tabela 1 – Estado da arte de algoritmos paramétricos e não paramétricos.....	14
Tabela 2 – Exemplo duma matriz de classificação com quatro classes.....	49
Tabela 3 - Resultados das métricas de avaliação para a segmentação pelo método <i>Atropos</i> .....	50
Tabela 4 - Resultados das métricas de avaliação para a segmentação por modelos de mistura Gaussianas por processos de <i>Dirichlet</i> .....	52
Tabela 5 – Resultados das métricas de avaliação para a segmentação pela abordagem Riciana.....	55
Tabela 6 – Diferença das médias das métricas de avaliação entre as segmentações por modelos de mistura Gaussianas e da abordagem Riciana.....	56
Tabela 7 – Matriz de classificação da segmentação por modelos de mistura Gaussianas por processos de <i>Dirichlet</i> .....	57
Tabela 8 - Matriz de classificação da segmentação pela abordagem Riciana.....	57
Tabela 9 - Resultados das métricas de avaliação para a segmentação pela abordagem <i>Genlogistic</i> .....	60
Tabela 10 - Diferença das médias das métricas de avaliação entre as segmentações por modelos de mistura Gaussianas e da abordagem <i>Genlogistic</i> .....	61
Tabela 11 - Matriz de classificação da segmentação pela abordagem <i>Genlogistic</i> .....	62
Tabela 12 - Resultados das métricas de avaliação para a segmentação pela abordagem <i>Gumbel</i> enviesada à esquerda.....	64
Tabela 13 - Diferença das médias das métricas de avaliação entre a segmentação por modelos de mistura Gaussianas e da abordagem <i>Gumbel</i> enviesada à esquerda.....	64
Tabela 14 - Matriz de classificação da segmentação pela abordagem <i>Gumbel</i> enviesada à esquerda.....	66
Tabela 15 - Resultados das métricas de avaliação para a segmentação pela abordagem <i>Log Gama</i> .....	68
Tabela 16 - Diferença das médias das métricas de avaliação entre a segmentação por modelos de mistura Gaussianas e da abordagem <i>Log Gama</i> .....	68
Tabela 17 - Matriz de classificação da segmentação pela abordagem <i>Log Gama</i> .....	69
Tabela 18 - Resultados das métricas de avaliação para a segmentação pela abordagem <i>Powernorm</i> .....	71
Tabela 19 - Diferença das médias das métricas de avaliação entre a segmentação por modelos de mistura Gaussianas e pela abordagem <i>Powernorm</i> .....	72
Tabela 20 - Matriz de classificação da segmentação pela abordagem <i>Powernorm</i> .....	73

---

Tabela 21 - Resultados das métricas de avaliação para a segmentação pela abordagem <i>Johnson Su</i> .....	75
Tabela 22 - Diferença das médias das métricas de avaliação entre a segmentação por modelos de mistura Gaussianas e pela abordagem <i>Johnson Su</i> .....	75
Tabela 23 - Matriz de classificação da segmentação pela abordagem <i>Johnson Su</i> .....	77

## SIMBOLOGIA

$\theta$  – parâmetros desconhecidos  
 $p(\dots)$  – função de densidade de probabilidade  
 $x$  – variável dos dados  
 $p(\dots|\dots)$  – função de verossimilhança  
 $y$  - variável dos dados  
 $Beta(\cdot)$  – distribuição Beta  
 $v_1$  – parâmetro de definição da distribuição Beta  
 $v_2$  - parâmetro de definição da distribuição Beta  
 $\Gamma(\cdot)$  – função gama  
 $\alpha$  – parâmetro positivo da distribuição Dirichlet  
 $Dir(\cdot)$  – distribuição Dirichlet  
 $\sigma$  – parâmetro de definição da distribuição Riciana  
 $v$  – parâmetro de não centralidade da distribuição Riciana  
 $I_0(\cdot)$  – função de Bessel modificada do 1º tipo, de ordem zero  
 $c$  – parâmetro de definição da distribuição Log Gama  
 $a$  – parâmetro alfa da definição da distribuição Johnson Su  
 $b$  – parâmetro beta da definição da distribuição Johnson Su  
 $\varphi(\cdot)$  – função de densidade de probabilidade da distribuição Gaussiana ou Normal  
 $\mu$  - média  
 $\sigma$  – desvio-padrão  
 $\sigma^2$  – variância  
 $v$  - grau de liberdade  
 $p$  – dimensão do vetor  
 $\Sigma$  – matriz covariância  
 $c$  – parâmetro de definição da distribuição Genlogistic  
 $c$  – parâmetro de definição da distribuição Powernorm  
 $\varphi(\cdot)$  – função de densidade de probabilidade da distribuição Gaussiana  
 $\Phi(\cdot)$  - função de distribuição acumulada da distribuição Gaussiana  
 $Mult(\cdot)$  – Distribuição Multinomial  
 $n$  – número de repetições da experiência  
 $x$  – número de vezes que um dado acontecimento se repetiu  
 $p$  – probabilidade de realização do acontecimento  
 $\delta$  – função Delta Dirac  
 $PD(\cdot)$  – Processo de Dirichlet  
 $G_0$  – Distribuição base  
 $\alpha$  – parâmetro de escala positivo  
 $n$  – número de amostras  
 $\delta_{X_i}$  – função Delta Dirac na posição  $X_i$   
 $X_n$  – distribuição a posteriori dos processos de Dirichlet  
 $G$  – distribuição desconhecida  
 $\beta_i$  – comprimento da divisão do segmento  
 $\beta_k$  – resto do comprimento da divisão do segmento  
 $\pi_k$  – peso da distribuição  $G$

$x_1, \dots, x_n$  - amostras  
 $p(x|x_1, \dots, x_n, \alpha, G_0)$  – função de probabilidade preditiva  
 $\eta$  – parâmetros desconhecidos  
 $B_1, \dots, B_k$  – partições do números naturais  
 $k$  – índice do número  
 $\delta\eta_i$  – função Delta Dirac na posição  $\eta_i$   
 $X_1^*, \dots, X_{|c|}^*$  - valores distintos das amostras  
 $c$  – conjunto das variáveis de atribuição  
 $\mu_n$  – família de medidas de probabilidade  
 $V_i$  – variável aleatória que segue uma distribuição Beta(1,  $\alpha$ )  
 $\eta_i^*$  - variável aleatória que segue a distribuição base  $G_0$   
 $\pi_i(v)$  – proporções de mistura  
 $Z_n$  – variável de atribuição do componente de mistura com o ponto  $x_n$  está associado  
 $\lambda$  – hiperparâmetro  
 $a(\eta_i^*)$  – função de distribuição apropriada  
 $\theta$  - hiperparâmetro  
 $W$  – variável latente  
 $x$  - observações  
 $p(w|x, \theta)$  – distribuição posterior das variáveis latentes  
 $p(x|\theta)$  – probabilidade marginal das observações  
 $p(w_i | w_{-i}, x, \theta)$  – distribuição das amostras iterativas de cada variável latente condicionada nos valores anteriormente amostrados das outras variáveis latentes  
 $D(p||q)$  – divergência Kullback-Leibler  
 $p$  – distribuição de probabilidade discreta  
 $q$  - distribuição de probabilidade discreta  
 $q_v(w)$  – família de distribuições indexadas por um parâmetro de variação  $v$   
 $g_i(w_{-i}, x, \theta)$  – parâmetro natural para  $w_i$  quando condicionada nas restantes variáveis latentes e observações  
 $v$  – parâmetros de variação  
 $W$  – variável latente  
 $V$  – comprimento do segmento  
 $\eta^*$  - funções Delta Dirac  
 $Z$  - atribuições do aglomerado  
 $\Theta$  - hiperparâmetros  
 $\alpha$  – parâmetro de escala  
 $\lambda$  – parâmetro da distribuição base conjugada  
 $T$  - limite

$q(v, \eta^*, z)$  – família fatorizada de distribuições de variação para inferência de variação mean-field

$q_{\gamma t}(v_t)$  – distribuições Beta

$q_{\tau t}(\eta_t^*)$  – distribuições da família de exponencias com parâmetros naturais  $\tau_t$

$q_{\phi n}(Z_n)$  – distribuições multinomiais

$v$  – parâmetros de variação livres

$\Psi$  – função digamma

$p_{\text{normal}}$  – função de densidade de probabilidade da distribuição normal ou Gaussiana

$p_{\text{T-student}}$  – função de densidade de probabilidade da distribuição T-student

$\eta_{\text{normal}}$  – vetor dos componentes da mistura para cada segmento da definição construtiva dos processos de Dirichlet para a distribuição Gaussiana

$\eta_{\text{T-student}}$  – vetor dos componentes da mistura para cada segmento da definição construtiva dos processos de Dirichlet para a distribuição T-student

$Z_{n,T}$  – matriz multinomial dos segmentos da definição construtiva dos processos de Dirichlet para cada amostra

$\alpha_1$  – parâmetro de concentração

$\alpha_2$  – parâmetro de concentração

$v_i$  – pesos da divisão do segmento daa definição construtiva dos processos de Dirichlet para cada componente de mistura

$w_i$  – segmento da definição construtiva dos processos de Dirichlet segundo a distribuição a priori

$v_{i,1}$  – peso da divisão do segmento

$v_{i,2}$  – peso da divisão do segmento

$\mathcal{L}(\theta | x_n)$  – função de verosimilhança

$\gamma_i$  – soma do logaritmo das probabilidades dos dados segundo as distribuições T-student de cada componente de mistura do modelo e o logaritmo dos pesos associados a cada distribuição T-student de cada componente de mistura do modelo

$p(x|x_n, Z_{n,T})$  – função de densidade de probabilidade preditiva do modelo

$v$  – resultado da imagem verdadeira

$u$  – campo de polarização

$b$  – falta de homogeneidade multiplicativa

$n$  – ruído do scanner

$m_1$  – intensidade mínima

$m_2$  – intensidade máxima

$pc_1$  – primeiro percentil

$pc_2$  – último percentil

$\mu_p$  – marcas do histograma

$x'$  – intensidade mapeada

$p_{\text{Riciana}}$  – função de densidade de probabilidade da distribuição Ricianiana

$\eta_{\text{Riciana}}$  – vetor dos componentes da mistura para cada segmento da definição construtiva dos processos de Dirichlet para a distribuição Riciana

$p_{\text{Pownorm}}$  – função de densidade de probabilidade da distribuição Pownorm

$\eta_{\text{Pownorm}}$  – vetor dos componentes da mistura para cada segmento da definição construtiva dos processos de Dirichlet para a distribuição Pownorm

$p_{\text{Johnson Su}}$  – função de densidade de probabilidade da distribuição Johnson Su

$\eta_{\text{Johnson Su}}$  – vetor dos componentes da mistura para cada segmento da definição construtiva dos processos de Dirichlet para a distribuição Gaussiana Johnson Su

S – segmentação obtida

R – marcação manual

## **ABREVIATURAS**

HMRF – Hidden Markov Random Field

EM – Esperança maximizada

MRF – Markov Random Field

HMC - Cadeia Hidden Markov

MAP – maximização a posteriori

3D – Tridimensional

MICCAI – Medical Image Computing and Computer Assisted Intervention

KL – Kullback-Leibler

TPR – True Positive Rate

PPV – Positive Predictive Value

VP – verdadeiros positivos

FN – falsos negativos

FP – falsos positivos

VN – verdadeiros negativos

# 1. INTRODUÇÃO

## 1.1. Enquadramento

Com a introdução da ressonância magnética como ferramenta de diagnóstico para as diversas áreas da medicina houve um grande contributo para uma maior eficácia e qualidade no tratamento e diagnóstico de doenças. Um dos procedimentos para se realizar a avaliação e diagnóstico de doenças consiste em fazer um estudo das imagens obtidas por ressonância magnética através da sua segmentação.

A segmentação de imagens da ressonância magnética do cérebro em diferentes tipos de tecidos é muitas vezes uma tarefa importante e essencial no estudo de muitas doenças. Este processo possibilita a quantificação de mudanças neuroanatômicas em diferentes tecidos cerebrais, que pode servir de informação para o diagnóstico e tratamento de acompanhamento para pacientes de tumores cerebrais. Para a esclerose múltipla, por exemplo, requer a quantificação exata das lesões na substância branca, parte do cérebro para a avaliação do tratamento com recurso a fármaco [1]. No caso de acidentes vasculares cerebrais, a segmentação é usada na avaliação da extensão dos tecidos cerebrais afetados, no prognóstico da recuperação do paciente e na verificação da evolução dos danos provocados [50,51]. É também útil na doença de Alzheimer e défice cognitivo ligeiro, pois os tecidos cerebrais sofrem alterações numa fase inicial, mais precisamente, no hipocampo e no tálamo, onde a segmentação destas zonas podem servir de detetor precoce da doença [52,53,54]. Apesar da ressonância magnética produzir imagens volumétricas de alta qualidade, a deterioração da qualidade das imagens (ruído, intensidades não homogéneas e efeitos volume parcial) torna a segmentação das imagens cerebrais da ressonância magnética complexa, difícil e desafiante. Devido a estas razões, a segmentação cerebral da ressonância magnética tem vindo a ser uma área muito procurada para investigação, com uma variedade de métodos. Métodos tradicionais como *threshold* da imagem e crescimento de região geralmente falham devido a distribuições complexas de intensidades em imagens médicas. [1]

Contudo, recentemente, neste tipo de investigação, os métodos de segmentação baseados em métodos estatísticos têm sido estudados. Na segmentação de imagens cerebrais da ressonância magnética consideram-se três tipos de tecido: substância cinzenta, substância branca e líquido cerebrospinal. Das imagens obtidas da ressonância magnética é possível extrair um conjunto de características: como por exemplo, a intensidade de cada píxel. Essa característica pode ser utilizada para treinar um classificador, baseado em métodos estatísticos. Posteriormente, este pode ser usado para calcular a probabilidade da intensidade do píxel pertencer a cada tipo de tecido. Existem duas maneiras de se obter essas probabilidades: usando métodos paramétricos ou não paramétricos. Em ambas as abordagens, é necessário definir um conjunto de parâmetros, que são desconhecidos, para se obter a função de densidade de probabilidade. No caso da abordagem paramétrica, define-se um limite para o número de parâmetros e a função de densidade de probabilidade

é conhecida. Enquanto na abordagem não paramétrica, não existe limite para o conjunto de parâmetros e não se faz suposições sobre os parâmetros da função de densidade de probabilidade a partir dos dados. [2, 3]

O algoritmo Esperança Maximizada, o algoritmo de agrupamento de modelos de mistura frequencista e modelos finitos de mistura Gaussianos são alguns dos exemplos de algoritmos que usam a abordagem paramétrica para classificar e estimar as intensidades dos pixels de imagens da ressonância magnética [2, 3]. Estes métodos têm mais obstáculos relativamente aos que foram acima mencionados, como por exemplo, na limitação do número de componentes estimados poder não ser suficiente para modelar corretamente as intensidades e de intensidades de cérebros criarem grandes desvios dos histogramas de intensidades de cérebros da população normal [2].

Exemplos de algoritmos que usam a abordagem não paramétrica são: o algoritmo do desvio da média, classificador k-vizinho mais próximo e mais recentemente os modelos de mistura Gaussianos por processos de *Dirichlet*. Nesta dissertação são apresentados resultados que se inserem no último tipo de algoritmos referido, i.e., mistura de modelos Gaussianos por processos de *Dirichlet*.

## 1.2. Objetivo

O objetivo principal desta dissertação foi obter um modelo estatístico de previsão das intensidades dos pixels para os diferentes tipos de tecidos das imagens da ressonância magnética e verificar a sua contribuição para um sistema de segmentação automático. Para isso, foi estudado o modelo de mistura Gaussiano por processos de *Dirichlet* como o modelo estatístico de previsão. E também foi estudado se as distribuições utilizadas pelo modelo pudessem ser modificadas, segundo a observação dos histogramas das intensidades das imagens da ressonância magnética para melhorar o modelo estatístico de previsão.

## 1.3. Organização da dissertação

A dissertação está organizada em 6 capítulos, da seguinte forma: no capítulo 2 apresenta-se um resumo dos conceitos fundamentais utilizados na dissertação, como a estatística Bayesiana, as diferentes distribuições utilizadas na dissertação e os processos de *Dirichlet* e as suas representações. O capítulo 3 consiste no estado da arte da segmentação de tecidos cerebrais em imagens de ressonância magnética por metodologia paramétrica e não paramétrica. O capítulo 4 apresenta os conceitos teóricos da modelação por modelos de mistura por processos de *Dirichlet*. No capítulo 5 é feita a descrição do trabalho experimental realizado e apresentados os resultados obtidos. Por fim, no capítulo 6 são apresentadas as conclusões obtidas deste trabalho que nos apontam também o possível trabalho futuro.

## 2. CONCEITOS FUNDAMENTAIS

Neste capítulo são introduzidos conceitos que são utilizados ao longo da dissertação. Começamos com uma breve descrição da estatística Bayesiana. De seguida são descritas várias distribuições de probabilidades e é apresentada uma explicação para o processo de *Dirichlet* e suas representações.

### 2.1. Estatística Bayesiana

A estatística Bayesiana é um sistema para descrever incertezas epistemológicas utilizando a linguagem matemática de probabilidades [4]. Segundo a metodologia Bayesiana, as hipóteses são expressas através de distribuições de probabilidade para dados observáveis. Essas distribuições dependem de quantidades desconhecidas que se designam como parâmetros,  $\theta$ . No paradigma de *Bayes*, o conhecimento atual sobre o modelo dos parâmetros é expresso através da atribuição duma distribuição de probabilidade sobre os parâmetros, que se chama distribuição *a priori*,  $p(\theta)$  [5], isto é, exprime de alguma forma o conhecimento prévio à observação do conjunto de dados [6].

Quando novos dados,  $x$ , se tornam disponíveis, a informação que contêm, em relação ao modelo dos parâmetros, é expressa na função de verosimilhança, que é proporcional à distribuição dos dados observados, dado o modelo dos parâmetros, que se escreve como  $p(x|\theta)$  [4].

A informação é então combinada com a distribuição *a priori* para produzir uma distribuição de probabilidades atualizada que é denominada como distribuição *a posteriori*,  $p(\theta|x)$ , em que a inferência Bayesiana se baseia. O teorema de Bayes é a regra de atualização para ir atualizando a distribuição *a posteriori* [6], uma identidade elementar na teoria de probabilidade, afirma que a distribuição *a posteriori*, é proporcional à distribuição *a priori* vezes a função de verosimilhança, ou seja:

$$p(\theta|x) = \frac{p(x \cap \theta)}{p(x)} = \frac{p(\theta)p(x|\theta)}{\int_{\Theta} p(\theta)p(x|\theta) d\theta} \propto p(x|\theta)p(\theta), \quad (1)$$

onde  $p(x)$  define a distribuição marginal  $X$ , também designada por distribuição preditiva *a priori*, uma vez que sumaria a informação relativa a  $x$  antes de este ter sido observado. A equação foi substituída por uma proporcionalidade, que será útil em problemas que envolvam estimação de parâmetros, uma vez que o denominador é uma constante normalizadora. Também pode ser usado na comparação de modelos [6].

Depois do conjunto de dados  $x$  ter sido observado, para fazer predições sobre um conjunto de dados, por exemplo  $y = (y_1, \dots, y_m)$ , desconhecido mas observável, de um modelo cuja distribuição de  $y$  pode ser igual ou diferente da distribuição de  $x$  (note-se que em ambos os modelos são partilhados os mesmos parâmetros  $\theta$ ), define-se a distribuição preditiva *a posteriori*,  $p(y|x)$ , por [6]

$$p(y|x) = \int_{\Theta} p(y, \theta|x) d\theta = \int_{\Theta} p(y|x, \theta)p(\theta|x) d\theta \quad (2)$$

Esta distribuição resume a informação relativa à plausibilidade de observar um conjunto de dados  $y$ , condicionada ao conjunto de dados observados  $x$ . Se, condicionalmente a  $\theta$ , as observações a predizer forem independentes de  $x$ , então  $p(y|x, \theta) = p(y|\theta)$  e [6]

$$p(y|x) = \int_{\Theta} p(y|\theta)p(\theta|x) d\theta \quad (3)$$

A estatística Bayesiana divide-se em duas partes: estatística Bayesiana paramétrica e estatística Bayesiana não paramétrica que se diferenciam na definição dos parâmetros. Na estatística Bayesiana paramétrica define-se que os parâmetros são desconhecidos mas com dimensão finita. Na estatística Bayesiana não paramétrica, os parâmetros também são desconhecidos mas têm dimensão infinita [6,7].

### 2.2 Distribuição de Dirichlet

A distribuição de *Dirichlet* é um caso especial da distribuição *Beta*. A distribuição *Beta* é definida por:

$$Beta(x|v_1, v_2) = \frac{\Gamma(v_1 + v_2)}{\Gamma(v_1)\Gamma(v_2)} x^{v_1-1}(1-x)^{v_2-1}, \quad v_1, v_2 > 0 \quad (4)$$

se  $0 < x < 1$  e zero, caso contrário.  $\Gamma(\cdot)$  representa a função gama,  $v_1$  e  $v_2$  são os parâmetros que definem a distribuição *Beta*.

No caso do conjunto de variáveis  $x$  terem uma dimensão  $k$ ,  $(x_1, \dots, x_k)$ , para a distribuição *Beta*, obtém-se a distribuição de *Dirichlet* com parâmetros  $\alpha = \alpha_1, \dots, \alpha_k$  e  $\alpha > 0$  que é definida por: [6,8,9,10,11]

$$Dir(x_1, \dots, x_k; \alpha) = \frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} \prod_{i=1}^k x_i^{\alpha_i-1} \quad (5)$$

A figura 1 representa exemplos da função densidade da probabilidade da distribuição de *Dirichlet*.

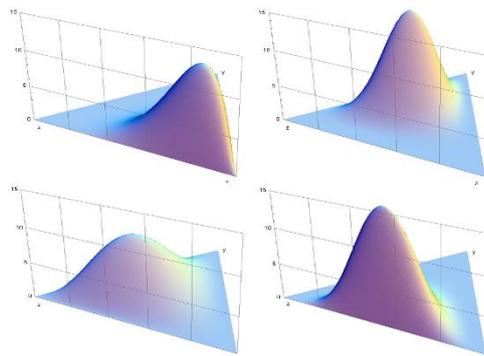


Figura 1 – Funções densidade da probabilidade da distribuição de *Dirichlet* com  $k=3$  e com  $\alpha=6,2,2$  (topo esquerdo),  $\alpha=3,7,5$  (topo direito),  $\alpha=6,2,6$  (fundo esquerdo),  $\alpha=2,3,4$  (fundo direito) [10]

### 2.3 Distribuição Riciana

A distribuição Riciana com parâmetros  $v$  e  $\sigma$ , é definida por:

$$f(x|v, \sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{(x^2 + v^2)}{2\sigma^2}\right) I_0\left(\frac{xv}{\sigma^2}\right) \quad (6)$$

onde  $I_0(\cdot)$  corresponde à função de *Bessel* modificada do 1º tipo, de ordem zero. Na figura 2 são apresentados exemplos de diferentes funções densidade de probabilidade da distribuição Riciana com diferentes parâmetros. [12,13]

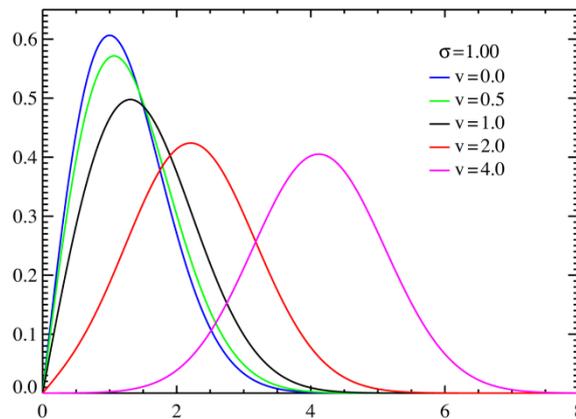


Figura 2 – Função densidade da probabilidade da distribuição Riciana com  $\sigma=1$  para diferentes  $v$  [13]

### 2.4 Distribuição Log Gama

A distribuição *Log Gama* com parâmetro  $c$  tem como função densidade de probabilidade a função  $f(x|c)$  cuja definição é:

$$f(x|c) = \frac{e^{cx-e^x}}{\Gamma(c)} \quad (7)$$

para  $x, c > 0$  [14]. A figura 3 mostra um exemplo de uma função densidade de probabilidade da distribuição *Log Gama*.

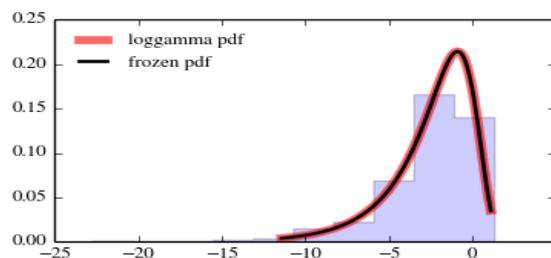


Figura 3 – Função densidade de probabilidade da distribuição *Log Gama* com  $c=0.414$  [14]

## 2.5 Distribuição Johnson Su

A função de densidade de probabilidade da distribuição *Johnson Su* com parâmetros  $a$  e  $b$ , é definida por:

$$f(x|a, b) = \frac{b}{\sqrt{x^2 + 1}} \varphi\left(a + b * \log\left(x + \sqrt{x^2 + 1}\right)\right) \quad (8)$$

Para  $x, a, b > 0$  e  $\varphi(\cdot)$  representa a função de densidade de probabilidade da distribuição Normal ou Gaussiana [15]. A figura 4 representa um exemplo duma função densidade de probabilidade de uma distribuição *Johnson Su*.

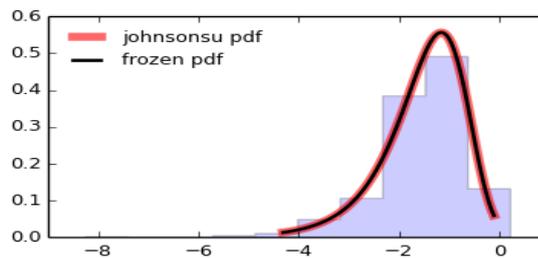


Figura 4 – Função densidade de probabilidade da distribuição *Johnson Su* com os parâmetros  $a=2.55$  e  $b=2.25$  [15]

## 2.6 Distribuição Normal ou Gaussiana

A função densidade de probabilidade da distribuição Normal ou Gaussiana com parâmetros  $\mu$  e  $\sigma$ , é definida por:

$$f(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (9)$$

Para  $\mu \in \mathbb{R}$  e  $\sigma^2 > 0$  [16]. A figura 5 mostra exemplos de funções densidade de probabilidade de uma distribuição Normal.

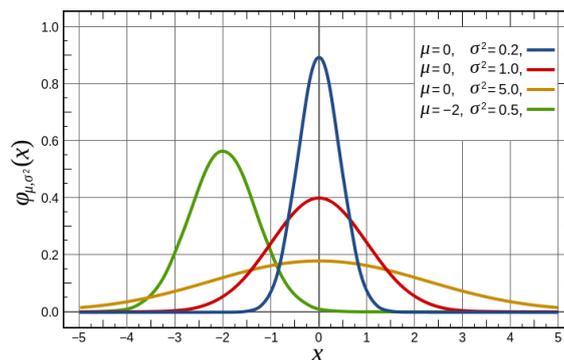


Figura 5 – Função densidade de probabilidade da distribuição Normal para diferentes  $\mu$  e  $\sigma^2$  [16]

## 2.7 Distribuição *T-student* multivariável

A função de densidade de probabilidade da distribuição *T-student* multivariável com parâmetros  $\mu$ ,  $\Sigma$  e  $\nu$ , é definida por:

$$f(x|\mu, \Sigma, \nu) = \frac{\Gamma[(\nu + p)/2]}{\Gamma\left(\frac{\nu}{2}\right) \nu^{\frac{p}{2}} \pi^{\frac{p}{2}} |\Sigma|^{1/2} \left[1 + \frac{1}{\nu} (x - \mu)^T \Sigma^{-1} (x - \mu)\right]^{(\nu+p)/2}} \quad (10)$$

Onde  $\nu$  é o grau de liberdade,  $p$  é a dimensão do vetor  $x$  e  $\Sigma$  é a matriz covariância [17].

## 2.8 Distribuição *Gumbel* enviesada à esquerda

A função densidade de probabilidade da distribuição *Gumbel* enviesada à esquerda é definida por [18]:

$$f(x) = e^{x-e^x} \quad (11)$$

A figura 6 mostra um exemplo duma função de densidade de probabilidade de uma distribuição *Gumbel* enviesada à esquerda.

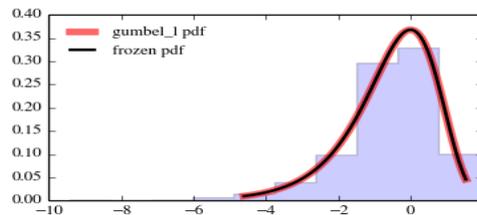


Figura 6 – Função densidade de probabilidade da distribuição *Gumbel* enviesada à esquerda [18]

## 2.9 Distribuição *Genlogistic*

A função densidade de probabilidade da distribuição *Genlogistic* com parâmetro  $c$ , é definida por [19]:

$$f(x|c) = \frac{ce^{-x}}{(1 + e^{-x})^{c+1}} \quad x, c > 0 \quad (12)$$

A figura 7 mostra um exemplo duma função densidade de probabilidade de uma distribuição *Genlogistic*.

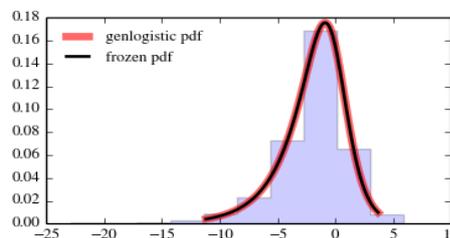


Figura 7 – Função densidade de probabilidade da distribuição *GenLogistic* com  $c = 0.412$  [19]

### 2.10 Distribuição *PowerNorm*

A função densidade de probabilidade da distribuição *PowerNorm* com parâmetro  $c$ , é definida por:

$$f(x|c) = c\varphi(x)\Phi(-x)^{c-1} \quad x, c > 0 \quad (13)$$

onde  $\varphi(\cdot)$  representa a função densidade de probabilidade da distribuição normal e  $\Phi(\cdot)$  representa a função de distribuição acumulada da distribuição normal [20]. A figura 8 apresenta um exemplo para uma distribuição *Powernorm*.

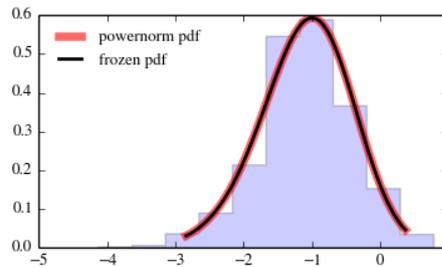


Figura 8 – Função densidade de probabilidade da distribuição *Powernorm* com  $c=4.445$  [20]

### 2.11 Distribuição Multinomial

A distribuição Multinomial é uma generalização da distribuição Binomial e a sua função densidade de probabilidade é apresentada como:

$$Mult(x_1, \dots, x_k) = \binom{n}{x_1, x_2, \dots, x_k} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k} = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k} \quad (14)$$

onde  $n$  representa o número de repetições da experiência,  $x_1, \dots, x_k$  representa o número de vezes que um dado acontecimento  $i, i = 1, \dots, k$  se repetiu,  $p_1, \dots, p_k$  representa a probabilidade de realização do acontecimento  $i$  tal que  $p_i \geq 0$ , para  $i = 1, \dots, k$  e  $\sum_{i=1}^k p_i = 1$ . [6,44]

### 2.12 Função *Delta Dirac*

A função *Delta Dirac* ou impulso unitário é definida por: [45,46]

$$\delta(x) = \begin{cases} 0, & x \neq 0 \\ \infty, & x = 0 \end{cases} \quad \text{com} \int_{-\infty}^{+\infty} \delta(x) dx = 1 \quad (15)$$

e representa-se por:

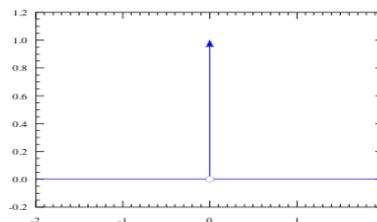


Figura 9 – Representação da função *Delta Dirac* [47]

### 2.13 Processos de *Dirichlet*

Os processos de *Dirichlet* são uma extensão da distribuição de *Dirichlet* para espaços infinitos [6]. Foi introduzido por *Ferguson (1973)* como uma medida de probabilidade aleatória sobre o espaço de medidas de probabilidade definidas sobre um espaço mensurável [6], i.e, é uma medida de probabilidade de medidas de probabilidades [6,21]. Os processos de *Dirichlet* são definidos por uma distribuição base  $G_0$  e um parâmetro escalar positivo  $\alpha$ . Para uma amostra aleatória  $X_1, X_2, \dots, X_n$ , escreve-se [6,7,8,21,22]:

$$G|G_0, \alpha \sim PD(G_0, \alpha) \quad (16)$$

$$X_1, X_2, \dots, X_n|G \sim G \quad (17)$$

$$X_n|X_1, X_2, \dots, X_{n-1} \sim \frac{1}{(\alpha + n - 1)} \left( \alpha G_0 + \sum_{i=1}^n \delta_{X_i} \right) \quad (18)$$

onde  $X_n$  representa a distribuição *a posteriori* dos processos de *Dirichlet*. Existem diferentes formas de representar os processos de *Dirichlet* como por exemplo: o esquema urna de *Pólya* apresentado por *Blackwell e MacQueen (1973)*, o processo “restaurante chinês”, o processo “restaurante indiano” e a definição construtiva dos processos de *Dirichlet* ou *stick breaking* por *Sethuraman (1994)* e *Rolin (1992)* [6,8]. Neste trabalho só serão discutidos o esquema urna de *Pólya* e a definição construtiva dos processos de *Dirichlet* porque foi nestas representações que o autor do código se baseou.

#### 2.13.1 Esquema Urna de *Pólya*

Ao integrar a medida aleatória  $G$ , a distribuição conjunta  $\{X_1, \dots, X_n\}$  segue um esquema urna de *Pólya* [21]. O esquema urna de *Pólya* consiste em ter uma urna contendo bolas que são extraídas de forma aleatória. Quando se retira uma bola, observa-se a cor dela, coloca-se novamente na urna e adiciona-se uma bola da mesma cor. Este esquema consegue criar um vetor de  $\{X_1, \dots, X_n\}$  com probabilidades condicionais  $X_n|X_1, X_2, \dots, X_{n-1} \sim \frac{\alpha G_0 + \sum_{i=1}^n \delta_{X_i}}{\alpha + n - 1}$ , ou seja, consegue calcular a distribuição *a posteriori* dos processos de *Dirichlet*. Para exemplificar, o esquema assuma que  $G_0$  é uma distribuição de cores e cada  $X_n$  representa a cor de cada bola colocada na urna. O algoritmo pode ser descrito por: iniciar o processo com a urna vazia, com a probabilidade proporcional a  $\alpha$ , extrair uma amostra  $X_n \sim G_0$  e adicionar uma bola da cor amostrada na urna, com a probabilidade proporcional a  $n - 1$ , retirar uma bola da urna, observar a cor, colocar novamente na urna juntamente com mais uma bola da mesma cor [24,25,26,27]. A probabilidade positiva é atribuída às configurações em que diferentes  $X_n$  assumem valores iguais. [21] ou seja:

$$X_n = \begin{cases} X_i^* \text{ com probabilidade } \frac{n^\circ \text{ de vezes que } X_i^* \text{ se repete}}{n - 1 + \alpha} \\ X \sim G_0 \text{ com probabilidade } \frac{\alpha}{n - 1 + \alpha} \end{cases} \quad (19)$$

**2.13.2 Definição construtiva dos processos de Dirichlet**

É uma maneira construtiva de representar a distribuição  $G$  em  $G = \prod_{k=1}^{\infty} \pi_k \delta_{X_k^*}$  usando a seguinte analogia: assume-se que temos um segmento com comprimento igual a 1, amostramos uma variável aleatória  $\beta_1 \sim Beta(1, \alpha)$ . Segundo a definição da distribuição *Beta*, o número amostrado estará entre 0 e 1, com o valor esperado  $\frac{1}{(1+\alpha)}$ . Divide-se o segmento na posição  $\beta_1$  e atribui-se a  $\pi_1$  o comprimento da parte esquerda do ponto  $\beta_1$  do segmento. Utilizando o resto do segmento, amostrar  $\beta_2 \sim Beta(1, \alpha)$ . Dividir o segmento na posição  $\beta_2$  e  $\pi_2$  é o comprimento do segmento dividido à esquerda do ponto  $\beta_2$ , i. e.,  $\pi_2 = (1 - \beta_1)\beta_2$ . Repetir este processo várias vezes para obter  $\pi_3, \pi_4, \dots$ , sendo possível continuar até infinito. Com base nisto, podemos modelar  $\pi_k = \beta_k \prod_{i=1}^{k-1} (1 - \beta_i)$ , onde  $\beta_k \sim Beta(1, \alpha)$  enquanto que nas outras representações,  $X$  é amostrado diretamente pela distribuição base  $X_k^* \sim G_0$ . Consequentemente a distribuição  $G$  pode ser escrita como uma soma de funções *Delta Dirac* com pesos  $\pi_k$  que fica  $G = \prod_{k=1}^{\infty} \pi_k \delta_{X_k^*}$ . [26,27,28,29]



*Figura 10 – Representação da divisão do segmento da definição construtiva dos processos de Dirichlet*

### 3. ESTADO DA ARTE

Neste capítulo serão apresentados trabalhos existentes que usam metodologia paramétrica e não paramétrica para classificar tecidos cerebrais em imagens de ressonância magnética.

#### 3.1 Métodos Paramétricos

A segmentação de tecidos cerebrais de imagens de ressonância magnética através de métodos paramétricos tem sido muito usada em diversos sistemas. São apresentados alguns dos métodos paramétricos mais estudados e utilizados nos últimos anos.

O artigo *Zhang et al. (2001)* propõe um modelo *Hidden Markov Random Field* (HMRF) modificado com o algoritmo de Esperança Maximizada (EM) para ultrapassar o problema do volume parcial, com o objetivo de segmentar imagens cerebrais da ressonância magnética. O algoritmo proposto segue o seguinte procedimento: no passo inicial, utiliza um estimador para obter os parâmetros dos tecidos e da classificação. É seguido pelo algoritmo de Esperança Maximizada em que atualiza as classes da classificação, os parâmetros do tecido e o *bias field* (campo de enviesamento) iterativamente. Durante as iterações, é usado o modelo *Markov Random Field* (MRF) com o método máximo *a posteriori* para estimar as classes da classificação. O método máximo *a posteriori* é aplicado para estimar o *bias field* e os parâmetros dos tecidos são estimados através da verosimilhança máxima. Os resultados mostram que o método proposto atinge o objetivo de segmentar as imagens cerebrais mas com algumas dificuldades. [34]

No artigo *Bricq et al (2008)* é apresentado um algoritmo de segmentação para imagens cerebrais da ressonância magnética com lesões de esclerose múltipla. O método classifica os tecidos cerebrais através do modelo da cadeia *Hidden Markov* (HMC) e deteta as lesões como *outliers* para o modelo. Também usa um estimador truncado de verosimilhança para extrair os *outliers* com o uso do modelo de cadeia *Hidden Markov*, o modelo proposto tem informação entre as vizinhanças dos pixels e inclui informação *a priori* através de um atlas probabilístico. Os resultados obtidos por este método mostram que é capaz de classificar os tecidos cerebrais com pouco erro e de detetar regiões do cérebro com lesões. [35]

O artigo de *Liang et al (2009)* faz uma abordagem com o algoritmo de Esperança Maximizada para o procedimento maximização *a posteriori* (MAP) para segmentar misturas de tecido cerebral em cada pixel da imagem. Considera que os tecidos cerebrais seguem uma mistura de distribuições normais. Ao modelar as distribuições com o modelo *Markov Random Field*, a esperança condicionada da distribuição *a posteriori* das misturas de tecido é determinada, tendo em conta a informação espacial, e a estimativa das intensidades das misturas de tecidos. A estimativa das intensidades das misturas de tecidos no passo seguinte é calculada através da esperança máxima condicionada. A abordagem Esperança Maximizada para o procedimento maximização *a posteriori* é alcançada por um número finito de iterações e utiliza uma estimativa inicial.

Esta estrutura fornece uma solução para o efeito do volume parcial. Os resultados demonstram que tem potencial para estimar misturas de tecidos com precisão e com eficácia. [36]

O artigo *Balafar (2011)* faz uma abordagem da segmentação de imagens cerebrais da ressonância magnética por modelos de mistura Gaussianos. Para se obter os parâmetros do modelo de mistura Gaussiano, é utilizado o algoritmo Esperança Maximizada. O objetivo deste modelo é obter o número de componentes das diferentes características obtidas, através da análise de concentrações das características. Um dos problemas deste algoritmo é a presença de ruído na imagem. Para eliminar este problema *Balafar* apresenta uma extensão deste algoritmo que inclui imagens filtradas pela média como informação de vizinhança nas concentrações de características e o histograma da imagem é utilizado nas concentrações para acelerar o processo. O resultado obtido mostra que o algoritmo proposto tem maiores níveis de desempenho comparando com os métodos apresentados. [37]

### 3.2 Métodos Não Paramétricos

Os métodos não paramétricos têm sido muito usados na segmentação e classificação de imagens mas só recentemente têm sido aplicados na segmentação de tecidos cerebrais em imagens de ressonância magnética, especialmente os modelos de mistura Gaussianos por processos de *Dirichlet*. Aqui são apresentados alguns artigos que exploraram os modelos de mistura Gaussianos por processos de *Dirichlet* na classificação de tecidos cerebrais.

No artigo *Silva (2006)* é proposto um modelo de *Bayes* não paramétrico para classificar tecidos em imagens da ressonância magnética do cérebro. O modelo conhecido por modelo de mistura por processos de *Dirichlet*, usa os processos de *Dirichlet* como distribuição *a priori* para ultrapassar as limitações dos atuais modelos paramétricos de mistura finitos. Para além de usar modelos de mistura por processos de *Dirichlet* também usa a modelação *Markov Random Field* para diminuir a influência do ruído durante o processamento da imagem e para ter informação espacial da imagem. Com estes modelos, é capaz de, através da imagem, estimar o número de misturas Gaussianas e estimar a probabilidade da densidade da intensidade dos pixels para classificar o tecido. Neste trabalho, a avaliação é feita através de imagens simuladas e reais. Com base nas imagens simuladas é apresentada uma comparação com outros algoritmos, dos valores de desempenho obtidos, onde atinge valores equivalentes aos métodos de segmentação apresentados. Faz também uma observação sobre a modelação *Markov Random Field* que tem um maior contributo para o valor de desempenho quando a imagem tem um maior nível de ruído. Já nas imagens reais, faz uma comparação com o melhor método apresentado no artigo *Marroquin et al. (2002)*, onde revela que o método baseado nos processos de *Dirichlet* apresenta melhores níveis de desempenho que o método baseado no algoritmo de Esperança Maximizada apresentado por *Marroquin et al. (2002)*. [2]

O artigo *Alemu (2007)* propõe um sistema que utiliza modelos de mistura Gaussianos e modelos de mistura por processos de *Dirichlet* para estimar a probabilidade preditiva da intensidade dos pixéis. Desenvolve um modelo mais extensivo de misturas onde inclui algumas proporções de mistura entre fontes do sinal. Coloca restrição na quantidade de componentes de mistura ao usar *Dirichlet a priori* para obter proporções esparsas de mistura e utiliza o gradiente para estimar parâmetros e fazer inferências. Também utiliza os modelos de mistura Gaussianos e heurísticas para configurar os parâmetros para posteriormente aprender as proporções de mistura dos componentes. Mas, ao configurar os parâmetros por este procedimento, o modelo proposto não consegue aprender sozinho. Mesmo assim, esta abordagem demonstra melhores resultados que sistemas que usam só modelos de mistura Gaussianos. [30]

Já o artigo *Popuri et al. (2011)* opta por criar uma estrutura para a segmentação em 3D que utiliza as concentrações das características, obtidas através da análise do conjunto multidimensional das diferentes características, construído a partir das imagens T1, T1C e T2 da ressonância magnética e as características de textura, para estimar as regiões normais do cérebro das regiões do tumor. A segmentação é controlada pela aprendizagem da estatística das regiões das imagens dentro do conjunto multidimensional das características. Utiliza o modelo de mistura por processos de *Dirichlet* para determinar se uma região de tecido do cérebro pertence a uma região considerada normal ou se pertence a uma região considerada tumor. Desta forma, este procedimento leva a uma melhor classificação das regiões com tumor a partir do tecido cerebral. Os resultados obtidos da avaliação, usando imagens reais de tumores cerebrais da ressonância magnética, mostram que o procedimento proposto foi capaz de classificar os tumores dos tecidos que o envolvem. [31]

O artigo *Schmidt et al. (2013)* apresenta duas abordagens para o estudo de hipo intensidades em imagens T2 baseadas na inferência Bayesiana: uma para o pré-processamento e outra para a análise estatística de estruturas das imagens da ressonância magnética. No pré-processamento, utilizam *software* disponível para fazer uma normalização das imagens T1C, FLAIR e T2. O algoritmo de segmentação baseado na inferência Bayesiana, utiliza os processos de *Dirichlet* para obter as distribuições *a posteriori* que indicam a probabilidade de um pixel pertencer a uma zona de hipo intensidade ou não. Para calcular essas distribuições utilizam o método *Markov Chain Monte Carlo*. Também usam um modelo *Markov Random Field* Gaussiano para atenuar a necessidade de suavizar os dados antes da análise estatística. Os resultados obtidos demonstram que métodos baseados na estatística Bayesiana são capazes de segmentar, com eficiência, imagens cerebrais da ressonância magnética e podem ser aplicados no estudo da atividade cerebral das imagens da ressonância magnética. [32]

E no artigo *Yisu Lu et al. (2014)* é apresentada a aplicação do modelo não paramétrico de mistura por processos de *Dirichlet* na segmentação de imagens cerebrais da ressonância magnética que consegue fazer a classificação dos tecidos sem especificar o número de componentes na imagem. Também utiliza um filtro anisotrópico de difusão, durante o

pré-processamento para eliminar o ruído não linear das imagens enquanto inibe a difusão nos contornos da imagem e usa o modelo *Markov Random Fields* para restringir espacialmente, ou seja, para ter informação espacial enquanto decorre a segmentação. O método é usado para encontrar onde estão as concentrações das intensidades, os *clusters*, nos pixels e associa essa concentração a um tipo de tecido. Não só utiliza as imagens T1 que a ressonância magnética obtém mas utiliza a FLAIR e a T1C para ajudar a detetar e classificar o edema nas imagens para fazer a segmentação dos tecidos do cérebro, do tumor e do edema. Revela um nível de desempenho relativamente superior aos métodos apresentados na comparação, onde se destaca na especificidade e na sensibilidade. [33]

### 3.3 Sumário

A segmentação de tecidos cerebrais de imagens de ressonância magnética por análise estatística é uma área muito estudada, onde são apresentados muitos procedimentos todos os anos. Neste trabalho foram apresentados alguns métodos de segmentação baseados em metodologia paramétrica e não paramétrica. Na Tabela 1 encontra-se o resumo das diferentes metodologias apresentadas no estado da arte.

Metodologia	Trabalhos	Abordagens
Paramétrica	<i>Zhang et al (2001)</i>	HMRF com EM
	<i>Bricq et al(2008)</i>	HMC com estimador de verosimilhança
	<i>Liang et al (2009)</i>	EM com MAP e MRF
	<i>Balafar (2011)</i>	Misturas Gaussianas com EM
Não Paramétrica	<i>Silva (2006)</i>	Modelos de mistura por processos de <i>Dirichlet</i> com MRF
	<i>Alemu (2007)</i>	Misturas Gaussianas com processos de <i>Dirichlet</i>
	<i>Popuri et al (2011)</i>	Modelos de mistura por processos de <i>Dirichlet</i>
	<i>Schmidt et al (2013)</i>	Processos de <i>Dirichlet</i> com MRF
	<i>Yisu Lu et al (2014)</i>	Modelos de mistura por processos de <i>Dirichlet</i> com MRF e filtragem anisotrópico

Tabela 1 – Estado da arte de algoritmos paramétricos e não paramétricos

## 4. MODELAÇÃO ATRAVÉS DE MODELOS DE MISTURA POR PROCESSOS DE DIRICHLET

Neste capítulo é exposta a teoria por detrás do código utilizado, em que o autor *Tom Haines* se baseou em dois trabalhos. Baseou-se no artigo [21] para o método dos modelos de mistura Gaussianos por processos de *Dirichlet* e no artigo [38] para o método de limitação dos processos de *Dirichlet*. O nosso trabalho tem como base o trabalho de *Tom Haines* e com isso estendemos esse modelo com novas funcionalidades, que são explicadas no capítulo 5.

### 4.1 Modelo de mistura por processos de Dirichlet

O nosso objetivo é através de modelos de mistura por processos de *Dirichlet* calcular a função densidade de probabilidade preditiva, que é apresentada na forma:

$$p(x|x_1, \dots, x_N, \alpha, G_0) = \int p(x|\eta)p(\eta|x_1, \dots, x_N, \alpha, G_0) \quad (20)$$

dada uma amostra  $(x_1, \dots, x_N)$ , [21].

Seja  $X$  uma variável aleatória contínua,  $G_0$  uma distribuição de probabilidade para  $\eta$  e  $\alpha$  um valor escalar real positivo. Uma medida aleatória  $G$  é distribuída de acordo com os processos de *Dirichlet*, com o parâmetro de escala  $\alpha$  e a distribuição base  $G_0$  se, para todos os números naturais  $k$  e as  $k$ -partições  $\{B_1, \dots, B_k\}$ , [21]

$$(G(B_1), G(B_2), \dots, G(B_k)) \sim \text{Dir}(\alpha G_0(B_1), \alpha G_0(B_2), \dots, \alpha G_0(B_k)) \quad (21)$$

Integrando  $G$ , a distribuição conjunta das variáveis  $\{X_1, \dots, X_n\}$  apresenta um efeito de aglomeração, condicionado em  $n - 1$  extrações, o valor  $n$ -ésimo é, com probabilidade positiva, exatamente igual a uma das extrações: [21]

$$p(\cdot | X_1, \dots, X_{n-1}) \propto \alpha G_0(\cdot) + \sum_{i=1}^{n-1} \delta_{\eta_i}(\cdot) \quad (22)$$

Assim, as variáveis  $\{X_1, \dots, X_{n-1}\}$  são divididas aleatoriamente de acordo com quais as variáveis que são iguais para o mesmo valor, com a distribuição da partição obtida a partir dum esquema de urna de *Pólya*. O conjunto de valores  $\{X_1^*, \dots, X_{|c|}^*\}$  indica os valores distintos de  $\{X_1, \dots, X_{n-1}\}$ ,  $c = \{c_1, \dots, c_{n-1}\}$  representa o conjunto das variáveis de atribuição de tal forma que  $X_i = X_{c_i}^*$  e  $|c|$  é o número de células na partição. A distribuição de  $X_n$  segue a distribuição de urna: [21]

$$X_n = \begin{cases} X_i^* \text{ com probabilidade } \frac{|\{j: c_j = i\}|}{n-1+\alpha} \\ X, X \sim G_0 \text{ com probabilidade } \frac{\alpha}{n-1+\alpha} \end{cases} \quad (23)$$

onde  $|\{j: c_j = i\}|$  é o número de vezes que o valor  $X_i^*$  ocorre em  $\{X_1, \dots, X_{n-1}\}$ .

No modelo de mistura por processos de *Dirichlet*, o processo é usado como informação *a priori* não paramétrica numa especificação de *Bayes* hierárquica: [21]

$$\begin{aligned}
 G|\{\alpha, G_0\} &\sim DP(\alpha, G_0) \\
 \eta_n|G &\sim G \\
 X_n|\eta_n &\sim p(x_n|\eta_n)
 \end{aligned} \tag{24}$$

Os dados criados a partir deste modelo podem ser particionados de acordo com os valores distintos do parâmetro. Com isto, a mistura por processos de *Dirichlet* tem uma interpretação como um modelo de mistura flexível em que o número de componentes (isto é, o número de células na partição) é aleatório e cresce à medida que são observados novos dados. [21]

A definição dos processos de *Dirichlet* através das suas distribuições de dimensão finita na Equação (17) sustenta-se no teorema da consistência de *Kolmogorov*. O teorema da consistência de *Kolmogorov* consiste em: dada uma família de medidas de probabilidade  $\mu_n$  em  $\mathbb{R}^n$ , para  $n \geq 1$ , que satisfaz a condição de consistência natural, então estas medidas combinam-se para formar uma medida  $\mu$  no produto infinito  $\mathbb{R}^\infty$  [39]. A definição construtiva dos processos de *Dirichlet* fornece uma caracterização mais detalhada dos processos de *Dirichlet*. Considerando dois conjuntos infinitos de variáveis aleatórias independentes,  $V_i \sim Beta(1, \alpha)$  e  $\eta_i^* \sim G_0$ , para  $i = \{1, 2, \dots\}$  a representação da definição construtiva para a distribuição  $G$  é: [21]

$$\pi_i(v) = v_i \prod_{j=1}^{i-1} (1 - v_j) \tag{25}$$

$$G = \sum_{i=1}^{\infty} \pi_i(v) \delta_{\eta_i^*} \tag{26}$$

Esta representação dos processos de *Dirichlet* indica que  $G$  é discreta (com probabilidade 1). O suporte de  $G$  consiste numa soma infinita de funções *Delta Dirac*, extraídas independentemente de  $G_0$ . As proporções de mistura  $\pi_i(v)$  são dadas dividindo sucessivamente uma unidade de comprimento, num número infinito de peças. O tamanho de cada uma das peças seguintes, proporcional ao resto, é dado por uma amostra independente a partir duma distribuição *Beta* ( $1, \alpha$ ). [21]

Na mistura por processos de *Dirichlet*, o vetor  $\pi(v)$  inclui o vetor infinito das proporções de mistura e  $\{\eta_1^*, \eta_2^*, \dots\}$  cujas posições representam os componentes de mistura. Seja  $Z_n$  uma variável de atribuição do componente de mistura com qual o ponto dado,  $x_n$ , está associado. Os dados podem ser descritos pelo processo que se segue: [21]

1. Extrair  $V_i | \alpha \sim Beta(1, \alpha), i = \{1, 2, \dots\}$
2. Extrair  $\eta_i^* | G_0 \sim G_0, i = \{1, 2, \dots\}$
3. Para cada ponto dos  $n$  dados:
  - (a) Extrair  $Z_n | \{v_1, v_2, \dots\} \sim Mult(\pi(v))$
  - (b) Extrair  $X_n | Z_n \sim p(x_n | \eta_{Z_n}^*)$

Segundo este método, restringe-se as misturas por processos de *Dirichlet* aquelas para as quais os dados observáveis são extraídos de uma distribuição da família exponencial e onde a distribuição base para o processo de *Dirichlet* é o correspondente conjugado precedente. [21]

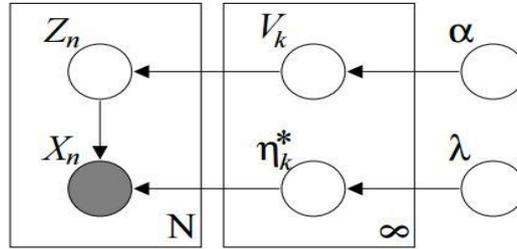


Figura 11 - Representação gráfica do modelo de mistura por processos de Dirichlet numa família exponencial [21]

Na figura 11 é representada a definição construtiva dos processos de *Dirichlet*. Os nós indicam variáveis aleatórias, os arcos orientados indicam possível dependência e as placas indicam repetição. As distribuições condicionadas de  $V_k$  e  $Z_n$  são descritas na figura acima. A distribuição de  $X_n$  condicionada a  $Z_n$  e  $\{\eta_1^*, \eta_2^*, \dots\}$  é: [21]

$$p(x_n|z_n, \eta_1^*, \eta_2^*, \dots) = \prod_{i=1}^{\infty} (h(x_n) \exp\{\eta_i^{*T} x_n - a(\eta_i^*)\})^{1_{[z_n=i]}} \quad (27)$$

onde  $a(\eta_i^*)$  é a função de distribuição apropriada e assume-se para simplificar que  $x$  é estatisticamente suficiente para estimar o parâmetro natural  $\eta$ . [21]

O vetor da estatística suficiente da família conjugada correspondente é  $(\eta^{*T}, -a(\eta^*))^T$ . A distribuição de base é: [21]

$$p(\eta^*|\lambda) = h(\eta^*) \exp\{\lambda_1^T \eta^* + \lambda_2(-a(\eta^*)) - a(\lambda)\} \quad (28)$$

onde se decompõe o hiperparâmetro  $\lambda$  tal que  $\lambda_1$  contém a primeira dimensão dos  $(\eta^*)$  componentes e  $\lambda_2$  é um escalar. [21]

## 4.2 Inferência variacional para misturas

Não há nenhuma maneira direta para calcular a distribuição posterior sob uma mistura por processos de *Dirichlet a priori*. Os métodos de inferência de aproximação são necessários para misturas por processos de *Dirichlet* e os métodos de amostragem através da cadeia de *Markov Monte Carlo* tornaram-se a melhor opção. [21]

A inferência variacional fornece uma alternativa à metodologia determinística para aproximar a verosimilhança e as distribuições posteriores. Considerando um modelo com hiperparâmetros  $\theta$ , variáveis latentes  $W = \{W_1, \dots, W_M\}$ , e observações  $x = \{x_1, \dots, x_N\}$  a distribuição posterior das variáveis latentes é dada como: [21]

$$p(w|x, \theta) = \exp\{\log p(x, w|\theta) - \log p(x|\theta)\} \quad (29)$$

Trabalhar diretamente com este posterior é normalmente difícil pela necessidade de calcular a constante de normalização. O logaritmo da probabilidade marginal das observações é: [21]

$$\log p(x|\theta) = \log \int p(w, x|\theta) dw \quad (30)$$

que pode ser difícil de calcular, uma vez que as variáveis latentes se tornam dependentes quando se condicionam os dados observados. [21]

Os algoritmos de cadeia de *Markov Monte Carlo* contornam este cálculo através da construção de uma distribuição posterior de aproximação, com base em amostras da cadeia de *Markov* cuja distribuição estacionária é o posterior de interesse. A amostragem de *Gibbs* é o algoritmo mais simples do algoritmo de cadeia *Markov Monte Carlo*, em que as amostras iterativas de cada variável latente condicionada nos valores anteriormente amostrados das outras variáveis latentes é: [21]

$$p(w_i|w_{-i}, x, \theta) = \exp\{\log p(w, x|\theta) - \log p(w_{-i}, x|\theta)\} \quad (31)$$

As constantes de normalização para estas distribuições condicionadas assumem-se disponíveis analiticamente para configurações em que a amostragem *Gibbs* é apropriada [21].

A inferência variacional é baseada na reformulação do problema de calcular a distribuição posterior como um problema de otimização e encontrar as soluções para o problema. Nesta abordagem, trabalha-se com uma classe particular de métodos de variação conhecidos como métodos de *mean-field*. Estes métodos baseiam-se na otimização da divergência *Kullback-Leibler* (KL) em relação a uma distribuição de variação. A divergência *Kullback-Leibler* é uma medida entre duas distribuições de probabilidade discretas  $q$  e  $p$ , que consiste em  $D(p||q) = \int p(x) \log \frac{p(x)}{q(x)} dx$  [43]. Em particular, seja  $q_\nu(w)$  uma família de distribuições indexadas por um parâmetro de variação  $\nu$ . O objetivo é minimizar a divergência KL entre  $q_\nu(w)$  e  $p(w|x, \theta)$ : [21]

$$D(q_\nu(w)||p(w|x, \theta)) = E_q[\log q_\nu(W)] - E_q[\log p(W, x|\theta)] + \log p(x|\theta) \quad (32)$$

onde se omitem os parâmetros de variação  $\nu$  ao usar  $q$  como uma expectativa subjacente. Note-se que a probabilidade marginal problemática não depende dos parâmetros de variação e pode ser ignorada na otimização. [21]

A minimização na equação (29) pode ser convertida alternativamente como a maximização de um limite inferior no logaritmo da verosimilhança marginal: [21]

$$\log p(x|\theta) \geq E_q[\log p(W, x|\theta)] - E_q[\log q_\nu(W)] \quad (33)$$

A diferença deste limite é a divergência entre  $q_\nu(w)$  e a distribuição posterior verdadeira. Para a estrutura de *mean-field* produzir um método de inferência computacionalmente eficiente é necessário escolher uma família de distribuições  $q_\nu(w)$  de modo que se possa controlar a otimização da equação (29). Na construção dessa família, quebram-se algumas das dependências entre as variáveis latentes que fazem o posterior verdadeiro difícil de calcular. Nas seções seguintes, consideram-se distribuições de variação totalmente fatorizadas que quebram todas as dependências. [21]

### 4.3 Mean-field - Inferência variacional na família das exponenciais

Para cada variável latente, supõe-se que a distribuição condicionada  $p(w_i|w_{-i}, x, \theta)$  é um membro da família exponencial: [21]

$$p(w_i|w_{-i}, x, \theta) = h(w_i) \exp\{g_i(w_{-i}, x, \theta)^T w_i - a(g_i(w_{-i}, x, \theta))\} \quad (34)$$

onde  $g_i(w_{-i}, x, \theta)$  é o parâmetro natural para  $w_i$  quando condicionado nas restantes variáveis latentes e observações. [21]

Nestas configurações considera-se a seguinte família de distribuições como aproximações de *mean-field* de variação: [21]

$$q_v(w) = \prod_{i=1}^M \exp\{v_i^T w_i - a(w_i)\} \quad (35)$$

onde  $v = \{v_1, v_2, \dots, v_M\}$  são os parâmetros de variação. Verifica-se que o algoritmo de variação que obtemos usando esta família totalmente fatorizada é uma reminiscência da amostragem *Gibbs*. A otimização da divergência do KL em relação a um único parâmetro de variação  $v_i$  é conseguida pelo cálculo da esperança seguinte: [21]

$$v_i = E_q[g_i(W_{-i}, x, \theta)] \quad (36)$$

Atualizando repetidamente cada parâmetro ao calcular esta esperança aumenta-se o desempenho da coordenada ascendente na divergência KL. [21]

De notar a relação interessante deste algoritmo com a amostragem de *Gibbs*. Na amostragem de *Gibbs* são extraídas de forma iterativa as variáveis latentes  $w_i$  da distribuição  $p(w_i|w_{-i}, x, \theta)$ . Na inferência variacional do *mean-field*, atualizam-se iterativamente o parâmetro de variação  $v_i$  definindo-o igual ao valor esperado de  $g_i(w_{-i}, x, \theta)$ . Esta esperança é calculada sob a distribuição variacional. [21]

### 4.4 Misturas por processos de Dirichlet

Nesta seção, desenvolve-se um algoritmo de variação de *mean-field* para a mistura por processos de *Dirichlet*. O algoritmo baseia-se na representação da definição construtiva dos processos de *Dirichlet* para misturas (Figura 10). Nesta representação, as variáveis latentes são o comprimento do segmento, as funções *Delta Dirac* e as atribuições do aglomerado:  $W = \{V, \eta^*, Z\}$ . Os hiperparâmetros são o parâmetro de escala e o parâmetro da distribuição base conjugada:  $\theta = \{\alpha, \lambda\}$ . [21]

Segundo a equação (30), podemos escrever o limite de variação no logaritmo da probabilidade marginal dos dados: [21]

$$\begin{aligned} \log p(x|\alpha, \lambda) &\geq E_q[\log p(V|\alpha)] + E_q[\log p(\eta^*|\lambda)] \\ &+ \sum_{n=1}^N (E_q[\log p(Z_n|V)] + E_q[\log p(x_n|Z_n)]) - E_q[\log q(V, \eta^*, Z)] \end{aligned} \quad (37)$$

Para explorar este limite, tem de se encontrar uma família de distribuições variacionais que se aproxime da distribuição da medida aleatória de dimensão infinita  $G$ , onde a medida aleatória é expressa em termos de conjuntos infinitos  $V = \{V_1, V_2, \dots\}$  e

$\eta^* = \{\eta_1^*, \eta_2^*, \dots\}$ . O objetivo será considerar as representações da definição construtiva dos processos de *Dirichlet* limitadas. Assim, fixando um valor  $T$  e sendo  $q(v_T = 1) = 1$ , as proporções de mistura  $\pi_t(v)$  são iguais a zero para  $t > T$  (Equação 21). [21]

As representações limitadas da definição construtiva dos processos de *Dirichlet* foram consideradas no contexto de inferência à base de amostragem para uma aproximação ao modelo de mistura por processos de *Dirichlet*. Note-se que o uso do limite é diferente, o modelo é um processo completo de *Dirichlet* e não é truncado, apenas a distribuição variacional é truncada. O nível de truncatura  $T$  é um parâmetro de variação que pode ser definido livremente, não é uma parte da especificação do modelo anterior. [21]

Propõe-se a seguinte família fatorizada de distribuições de variação para inferência de variação *mean-field*: [21]

$$q(v, \eta^*, z) = \prod_{t=1}^{T-1} q_{\gamma_t}(v_t) \prod_{t=1}^T q_{\tau_t}(\eta_t^*) \prod_{n=1}^N q_{\phi_n}(z_n) \quad (38)$$

onde  $q_{\gamma_t}(v_t)$  são distribuições *Beta*,  $q_{\tau_t}(\eta_t^*)$  são distribuições da família de exponenciais com parâmetros naturais  $\tau_t$ , e  $q_{\phi_n}(z_n)$  são distribuições multinomiais. Na seção 4.3, os parâmetros de variação livres são: [21]

$$v = \{\gamma_1, \dots, \gamma_{T-1}, \tau_1, \dots, \tau_T, \phi_1, \dots, \phi_N\} \quad (39)$$

É importante notar que há um parâmetro de variação diferente para cada variável latente sob a distribuição de variação. Por exemplo, a escolha do componente de mistura  $Z_n$  para o ponto de dados de ordem  $n$  é controlada por uma distribuição multinomial indexada por um parâmetro de variação  $\phi_n$ . Isto reflete a natureza condicionada da inferência de variação. [21]

#### 4.4.1 Algoritmo de coordenadas ascendente

Nesta seção, apresenta-se um algoritmo explícito de coordenada ascendente para otimizar o limite da equação (34) em relação aos parâmetros de variação. Todos os termos no limite envolvem cálculos normais na família exponencial, exceto o terceiro termo. Com isto, reescreve-se o terceiro termo usando a variável aleatória indicatriz: [21]

$$\begin{aligned} E_q[\log p(Z_n|V)] &= E_q \left[ \log \left( \prod_{i=1}^{\infty} (1 - V_i)^{1[Z_n > i]} V_i^{1[Z_n = i]} \right) \right] \\ &= \sum_{i=1}^{\infty} q(z_n > i) E_q[\log(1 - V_i)] + q(z_n = i) E_q[\log V_i] \end{aligned} \quad (40)$$

Recorde-se que a  $E_q[\log(1 - V_T)] = 0$  e  $q(Z_n > T) = 0$ . Assim podemos limitar esta soma em  $t = T$ : [21]

$$E_q[\log p(Z_n|V)] = \sum_{i=1}^T q(z_n > i) E_q[\log(1 - V_i)] + q(z_n = i) E_q[\log V_i] \quad (41)$$

onde

$$q(z_n = i) = \phi_{n,i} \quad (42)$$

$$q(z_n > i) = \sum_{j=i+1}^T \phi_{n,j} \quad (43)$$

$$E_q[\log V_i] = \Psi(\gamma_{i,1}) - \Psi(\gamma_{i,1} + \gamma_{i,2}) \quad (44)$$

$$E_q[\log(1 - V_i)] = \Psi(\gamma_{i,2}) - \Psi(\gamma_{i,1} + \gamma_{i,2}) \quad (45)$$

A função *digamma*, representada por  $\Psi$ , surge a partir da derivada do fator da normalização do logaritmo na distribuição *Beta*. [21]

Usando a expressão geral na equação (33) para obter o algoritmo coordenada ascendente *mean-field* obtém-se: [21]

$$\gamma_{t,1} = 1 + \sum_n \phi_{n,t} \quad (46)$$

$$\gamma_{t,2} = \alpha + \sum_n \sum_{j=t+1}^T \phi_{n,j} \quad (47)$$

$$\tau_{t,1} = \lambda_1 + \sum_n \phi_{n,t} x_n \quad (48)$$

$$\tau_{t,2} = \lambda_2 + \sum_n \phi_{n,t} \quad (49)$$

$$\phi_{n,t} \propto \exp(S_t) \quad (50)$$

para  $t \in \{1, \dots, T\}$  e  $n \in \{1, \dots, N\}$ , onde:

$$S_t = E_q[\log V_t] + \sum_{i=1}^{t-1} E_q[\log(1 - V_i)] + E_q[\eta_t^*]^T Xn - E_q[a(\eta_t^*)] \quad (51)$$

Repetindo estas atualizações, otimiza-se a equação (34) em relação aos parâmetros de variação definidos na equação (35). [21]

As implementações dos métodos de variação devem abordar a inicialização da distribuição de variação. Enquanto o algoritmo calcula um limite para qualquer valor inicial dos parâmetros de variação, más escolhas na inicialização podem levar a máximos locais que produzem limites incorretos ou fracos. A distribuição de variação pode ser iniciada incrementando na atualização os parâmetros de acordo com uma permutação aleatória dos pontos de dados, pode ser visto como uma versão de variação da amostragem de importância sequencial. Executa-se o algoritmo múltiplas vezes e escolhem-se as configurações dos parâmetros finais que dão o melhor limite na verosimilhança marginal. [21]

Para calcular a distribuição preditiva, usa-se o posterior de variação de forma análoga à forma que a aproximação empírica é usada por um algoritmo de amostragem de cadeia *Markov Monte Carlo*. A distribuição preditiva é: [21]

$$p(x_{N+1}|x, \alpha, \lambda) = \int \left( \sum_{t=1}^{\infty} \pi_t(v) p(x_{N+1}|\eta_t^*) \right) dP(v, \eta^*|x, \lambda, \alpha) \quad (52)$$

Sob a aproximação fatorizada da variação para o posterior, a distribuição das funções *Delta Dirac* e as suas proporções de mistura são dissociadas e a soma infinita é truncada. Consequentemente, a distribuição preditiva aproxima-se a um produto de esperanças que são simples de calcular sob a aproximação de variação, [21]

$$p(x_{N+1}|x, \alpha, \lambda) \approx \sum_{t=1}^T E_q[\pi_t(V)] E_q[p(x_{N+1}|\eta_t^*)] \quad (53)$$

onde  $q$  depende implicitamente de  $x$ ,  $\alpha$  e  $\lambda$ . [21]

Finalmente, são de notar duas extensões possíveis: em primeiro lugar, quando  $G_0$  não é conjugado, uma atualização simples da coordenada ascendente para  $\tau_i$  pode não ser possível, particularmente quando  $p(\eta_t^*|z, x, \lambda)$  não pertence à família exponencial. No entanto, essa atualização é possível para o caso especial de  $G_0$  ser uma mistura de distribuições conjugadas e em segundo lugar, é importante que em certas aplicações, seja possível integrar sobre uma distribuição *a priori* difusa no parâmetro de escalonamento  $\alpha$ . E ainda é possível estender o algoritmo de variação para incluir uma função *Gama a priori* sobre  $\alpha$ . [21]

Definiu-se um algoritmo de coordenada ascendente com base na equação (33) por forma a ir atualizando iterativamente  $v_i$  para  $i \in \{1, \dots, M\}$ . Tal algoritmo encontra um máximo local da Equação (30) pela proposição 2.7.1 de *Bertsekas (1999)*, sob a condição de que o lado direito da equação (30) seja estritamente convexo. [21]

Relaxando os dois pressupostos complica o algoritmo mas a ideia básica permanece igual. Se  $p(w_i|w_{-i}, x, \theta)$  não pertence à família exponencial, então pode não existir uma expressão analítica para a atualização na Equação (33). Se  $q(w)$  não for uma distribuição completamente fatorizada, então o segundo termo da equação do limite na equação (30) torna-se  $E_q[\log q(w_i|w_{-i})]$  e as simplificações podem não ser aplicáveis.[21]

#### 4.5 Limitação do modelo de mistura por processos de *Dirichlet*

A limitação da mistura variacional dos processos de *Dirichlet* tem a propriedade indesejável de que a família de variação com limite  $T$  não está contida dentro da família de variação com limite  $T + 1$ , i. e., que as famílias não estão agrupadas. O resultado é que pode haver um nível ótimo de limitação e finito  $T$  para  $q$ , que contradiz a intuição de que mais componentes vão permitir a  $q$  uma melhor aproximação, ou seja, atingir o melhor resultado quando  $T \rightarrow \infty$ . Além disso, localizar um nível de limite perto do ótimo pode ser difícil, uma vez que a equação (29) sendo uma função de  $T$  pode apresentar mínimos locais. [38] Com isto, foi utilizado um modelo variacional ligeiramente diferente para a equação (35) que permite que as famílias sobre  $T$  possam ser agrupadas. Esta configuração,  $q$  é dada por  $q(W; \varphi) = \prod_{i=1}^L [q_{v_i}(v_i; \phi_i^v) q_{\eta_i}(\eta_i; \phi_i^\eta)] \prod_{n=1}^N q_{z_n}(z_n)$ , onde deixam  $L$  ir até infinito, mas

associam os parâmetros de todos os modelos depois dum limite específico  $T$ , com  $T \ll L$ . Em particular, impõem a condição de que todos os componentes com índice  $i > T$ , as distribuições de variação  $q_{v_i}(v_i)$  e as distribuições de variação para os componentes  $q_{\eta_i}(\eta_i)$  são iguais aos seus *a priori* correspondentes, ou seja,  $q_{v_i}(v_i; \phi_i^v) = p_v(v_i|\alpha)$  e  $q_{\eta_i}(\eta_i; \phi_i^\eta) = p_v(\eta_i|\lambda)$ . [38]

Segundo o esquema,  $T$  define um nível de limite implícito da mistura variacional, já que não há parâmetros livres para otimizar além do nível de  $T$ . No entanto, ao contrário de [21], certos casos podem agora atribuir responsabilidade diferente de zero para componentes além do nível  $T$ , e, portanto, cada  $q_{z_n}(z_n)$  deve ter suporte infinito, que exige o cálculo de somas infinitas em diversas áreas de interesse. Uma implicação importante desta configuração é que as famílias variacionais que estão agora agrupadas em relação a  $T$ , para  $i > T$ ,  $q_{v_i}(v_i)$  e  $q_{\eta_i}(\eta_i)$  poderão reverter sempre para seus *a priori*, e como resultado é garantido que à medida que se aumenta  $T$ , existem soluções que diminuem a equação (29). Por isso, permite a otimização adaptativa com  $T$  a partir de  $T = 1$ , isto é, permite-nos começar a modelação sem especificar o número de componentes que os dados. Com esta alteração no modelo, a distribuição preditiva de densidade pode ser aproximada por: [38]

$$p(x_{N+1}|x, \alpha, \lambda) \approx \sum_{i=1}^T E_q[\pi_i(V)] E_q[p(x_{N+1}|\eta_i^*)] + \left[ 1 - \sum_{i=1}^T E_p[\pi_i(V)] \right] E_p[p(x_{N+1}|\eta)] \quad (54)$$

Notar que pode ser calculada analiticamente a distribuição preditiva e a convergência do modelo mesmo que envolva somas infinitas. [38]

## 4.6 Sumário

O modelo de mistura por processos de *Dirichlet* proposto consegue obter uma probabilidade *a posteriori*, com base nos dados submetidos, sem definir os parâmetros da probabilidade. Utiliza a definição construtiva dos processos de *Dirichlet* como ponto de partida para o cálculo da distribuição *a posteriori*. Para isso aplica-se um método de inferência variacional, *mean-field*, que reformula o problema de calcular uma distribuição num problema de otimização de um limite entre distribuições. Uma vez definido o método *mean-field*, utiliza-se o algoritmo de coordenada ascendente para encontrar as soluções do problema de otimização através de somas que calculam os parâmetros da distribuição, onde este processo é repetido várias vezes até que o modelo atinja convergência. Porém, este modelo limita o processo até a um valor estabelecido em vez de infinito, o que possibilita o cálculo da distribuição *a posteriori*. Para além de impor este limite, reformula novamente este limite para que os parâmetros que definem a distribuição se aproximem mais dos valores caso o processo continuasse até infinito.

## 5. TRABALHO EXPERIMENTAL

Neste capítulo são apresentados: estudos experimentais de modelos de mistura Gaussianos por processos de *Dirichlet* a estimar a densidade preditiva de uma mistura de Gaussianas, o estudo dos histogramas das intensidades dos diferentes tipos de tecidos cerebrais, a modelação das intensidades dos diferentes tipos de tecidos cerebrais através de modelos de mistura Gaussianos por processos de *Dirichlet*, a modelação por modelo de mistura por processos de *Dirichlet* modificado para trabalhar com distribuições Ricianas para um determinado tecido cerebral e também é feito para outro tipo de tecido com outras distribuições que são descritas nas próximas seções.

### 5.1 Exemplificação dos modelos de mistura Gaussianas por processos de *Dirichlet*

Para avaliarmos o modelo de mistura Gaussianos por processos de *Dirichlet* definimos uma distribuição Gaussiana com média e variância conhecida. De seguida, retiramos amostras aleatórias da distribuição Gaussiana definida para termos dados para o modelo de mistura por processos de *Dirichlet* para observar se a distribuição *a posteriori* consegue chegar, a partir das amostras aleatórias, à distribuição preditiva segundo os mesmos, que neste caso, à distribuição Gaussiana. Para se poder observar a evolução do modelo ao longo dos cálculos, os dados foram adicionados incrementalmente, isto é, se tivermos 2000 amostras começamos o nosso modelo com 8 amostras e depois de cada vez que se obtém uma aproximação à distribuição *a posteriori* adiciona-se mais amostras e repete-se novamente o mesmo processo até que tenham sido adicionadas todas as amostras disponíveis.

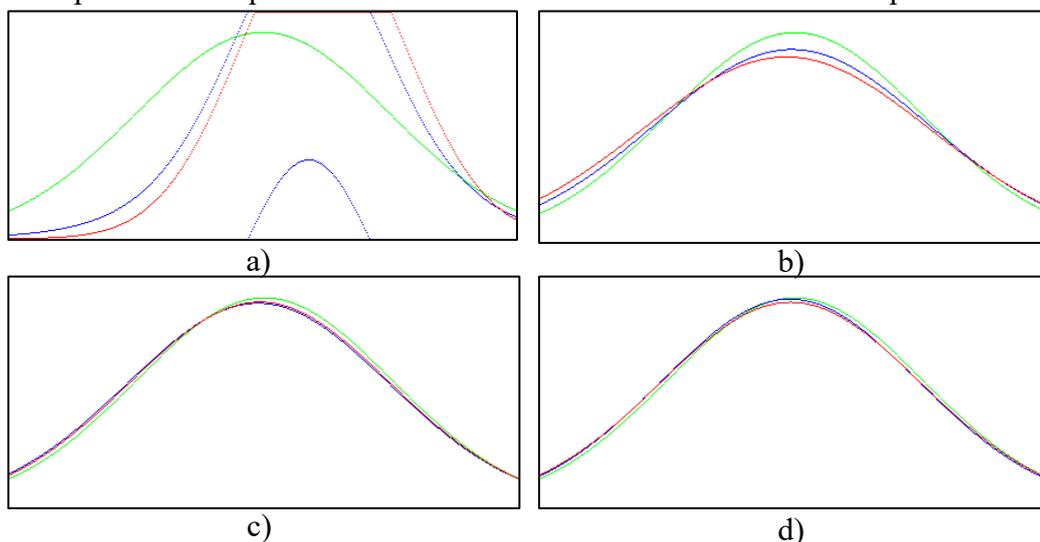


Figura 12 - Evolução do modelo de mistura por processos de *Dirichlet*. a) Estimativa com 8 amostras, b) Estimativa com 64 amostras, c) Estimativa com 1024 amostras, d) Estimativa com 2048 amostras, Linha verde – Distribuição Gaussiana, Linha azul e vermelha – probabilidade preditiva segundo o modelo por processos *Dirichlet*

A partir da figura 12 podemos observar a evolução do modelo de mistura Gaussiana por processos de *Dirichlet* com amostras de uma distribuição Gaussiana. À medida que introduzimos as nossas amostras no modelo, o modelo vai, segundo as amostras, estimar a função de densidade de probabilidade preditiva dos dados, ou seja, dá-nos uma previsão de qual é a distribuição de probabilidade, segundo as nossas amostras, de um valor futuro. Na figura 12, a linha verde representa a distribuição Gaussiana com os parâmetros conhecidos. A linha vermelha representa a probabilidade de densidade preditiva, calculada através da definição construtiva dos processos de *Dirichlet*. E a linha azul representa também a probabilidade de densidade preditiva mas calculada através da probabilidade *a posteriori* pela Equação (48). Ambos os métodos atingem o mesmo objetivo mas por definições diferentes. Também é apresentado um exemplo para uma mistura de duas distribuições Gaussianas (Figura 13), onde a linha verde corresponde à mistura de duas distribuições e a linha azul e vermelha correspondem à probabilidade preditiva segundo o modelo por processos de *Dirichlet* obtidas da mesma maneira que o primeiro exemplo e outro exemplo com uma mistura de três distribuições Gaussianas (Figura 14), onde a linha verde corresponde à mistura e a linha azul e vermelha correspondem à probabilidade preditiva segundo o modelo por processos de *Dirichlet* obtidas segundo o procedimento do primeiro exemplo.

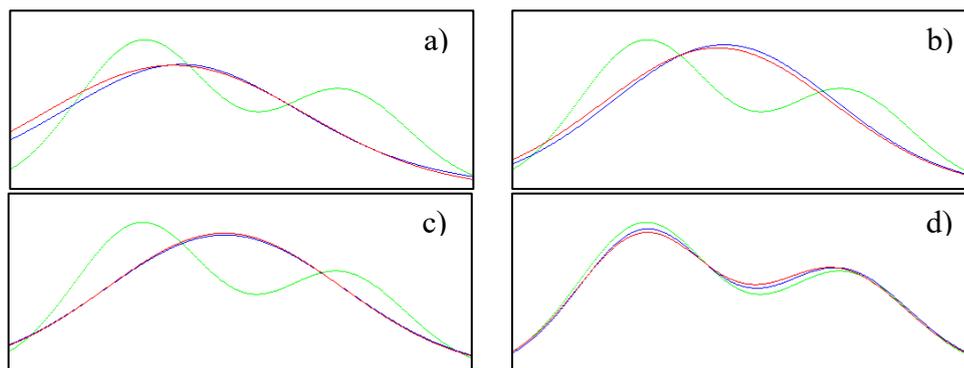


Figura 13 - Evolução do modelo de mistura por processos de *Dirichlet* de uma mistura de 2 distribuições. a) Estimativa com 8 amostras, b) Estimativa com 64 amostras, c) Estimativa com 1024 amostras, d) Estimativas com 2048 amostras

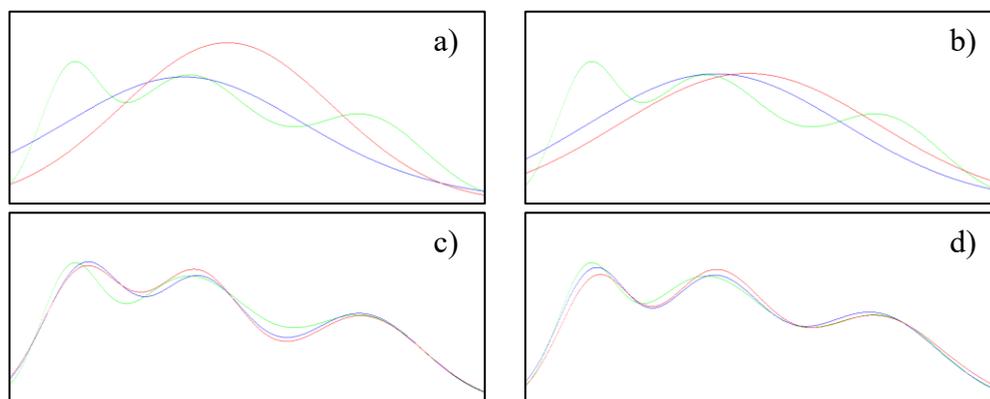


Figura 14 - Evolução do modelo de mistura por processos de *Dirichlet* de uma mistura de 3 distribuições. a) Estimativa com 8 amostras, b) Estimativa com 64 amostras, c) Estimativa com 1024 amostras, d) Estimativa com 2048 amostras

## 5.2 Procedimento do cálculo do modelo de mistura Gaussianas por processos de *Dirichlet*

Nesta secção é descrito o procedimento dos cálculos necessários para se obter o modelo de mistura Gaussiana por processos de *Dirichlet*. O primeiro passo do procedimento é a definição das variáveis necessárias para o cálculo dos parâmetros do modelo de mistura Gaussiano por processos de *Dirichlet*, como a sua dimensão,  $d$ , o limite máximo da definição construtiva dos processos de *Dirichlet*,  $T$ , as distribuições Gaussianas e *T-student*,  $p$ , a matriz multinomial dos segmentos da definição construtiva dos processos de *Dirichlet* para cada amostra,  $z$ , o limite da convergência do modelo,  $\varepsilon$ , o vetor dos componentes da mistura para cada segmento da definição construtiva dos processos de *Dirichlet*,  $\eta$ , e as variáveis auxiliares como parâmetros das distribuições usadas na definição construtiva dos processos de *Dirichlet* e nos cálculos de simplificação do modelo.

$$\begin{aligned} p_{normal} &\sim \text{Gaussiana}(\mu, \sigma) \\ p_{t\text{-student}} &\sim T - \text{Student}(\mu, \sigma, v) \\ \eta_{normal} &\sim \text{Gaussiana}(\mu_t, \sigma_t) \\ \eta_{t\text{-student}} &\sim T - \text{Student}(\mu_t, \sigma_t, v_t) \end{aligned} \quad (55)$$

onde  $\mu$  corresponde à média da distribuição,  $\sigma$  corresponde ao desvio padrão da distribuição,  $v$  corresponde ao grau de liberdade da distribuição *T-student*. O passo seguinte é a introdução dos dados no nosso modelo de mistura,  $x_n$ . Uma vez feita a introdução dos dados, o próximo passo é a obtenção da informação que se consegue retirar dos dados, isto é, uma vez que assumimos que os dados seguem uma mistura de distribuições Gaussianas, é calculado a média e o desvio padrão dos dados,  $\mu$  e  $\sigma$ , e depois é definida uma distribuição Gaussiana segundo a média e o desvio padrão calculado,  $p_{normal}$ , e denominamos esta distribuição *a priori* dos dados porque dá-nos a perspectiva de como os dados estão distribuídos antes de se calcular o modelo de mistura, segundo uma Gaussiana e também é extraída a distribuição *T-student* a partir dos parâmetros da distribuição Gaussiana,  $p_{t\text{-student}}$ .

$$\begin{aligned} \mu &= \frac{\sum_{i=1}^n x_n}{n}, \sigma = \sqrt{\frac{\sum_{i=1}^n (x_n - \mu)^2}{n - 1}}, p_{normal} = \text{Gaussiana}(\mu, \sigma) \\ p_{t\text{-student}} &= T - \text{Student}(\mu, \sigma, v) \end{aligned} \quad (56)$$

O próximo passo é o cálculo dos parâmetros do modelo de mistura Gaussiana por processos de *Dirichlet*. Visto que para o cálculo do modelo de mistura utilizamos um limite de convergência do modelo,  $\varepsilon$ , para verificar se o modelo consegue obter os parâmetros ótimos que definem o modelo, isto é, que melhor representam a distribuição dos dados introduzidos, o programa utiliza um ciclo para atingir essa convergência.

Inicialmente é definida uma distribuição de Dirichlet com o mesmo número de amostras que o número de dados introduzidos no modelo com o parâmetro de concentração  $\alpha$  igual a 32 porque se tivermos um  $\alpha$  pequeno pode levar a que o modelo não consiga atingir a convergência e com isso entraria num ciclo infinito.

$$z_{n,T} \sim Dir(\alpha = 32) \quad (57)$$

O primeiro passo do ciclo é a atualização do parâmetro de concentração do modelo definido por:

$$\begin{aligned} \alpha_1 &= \beta_1 + T \\ \alpha_2 &= \beta_2 - \sum E[\log(1 - v_i)] \end{aligned} \quad (58)$$

onde  $\alpha_1$  e  $\alpha_2$  representam o parâmetro de concentração,  $\beta_1$  e  $\beta_2$  representam a distribuição Beta, T representa o limite máximo da definição construtiva dos processos de Dirichlet e  $E[\log(1 - v_i)]$  representa a esperança do logaritmo do resto dos pesos da divisão do segmento da definição construtiva dos processos de Dirichlet para cada componente de mistura. O segundo passo é o cálculo da esperança dos valores da definição construtiva dos processos de Dirichlet segundo a distribuição a priori,  $p_{normal}$ , que é definido por:

$$\begin{aligned} E[\log(w_i)] &= \Psi(1) - \Psi\left(1 + \frac{\alpha_1}{\alpha_2}\right) \\ E[\log(1 - w_i)] &= \Psi\left(\frac{\alpha_1}{\alpha_2}\right) - \Psi\left(1 + \frac{\alpha_1}{\alpha_2}\right) \end{aligned} \quad (59)$$

onde  $\Psi$  representa a função digamma,  $E[\log(w_i)]$  corresponde à esperança do logaritmo do segmento da definição construtiva dos processos de Dirichlet segundo a distribuição a priori e  $E[\log(1 - w_i)]$  corresponde à esperança do logaritmo do resto do segmento da definição construtiva dos processos de Dirichlet segundo a distribuição a priori.

O terceiro passo do ciclo é a atualização dos pesos de cada das divisões segmento da definição construtiva dos processos de Dirichlet para os diferentes valores dos dados introduzidos no modelo, definido por:

$$\begin{aligned} v_{i,1} &= 1 + \sum z_{n,T} \\ v_{i,2} &= \frac{\alpha_1}{\alpha_2} + n - \sum_{cum} (\sum z_{n,T}) \end{aligned} \quad (60)$$

onde  $n$  corresponde ao número de amostras e  $v_{i,1}$  e  $v_{i,2}$  correspondem ao peso das divisões de cada segmento da definição construtiva dos processos de Dirichlet. O quarto passo do ciclo é calcular o logaritmo da esperança dos pesos das divisões da definição construtiva dos processos de Dirichlet do modelo definido por:

$$\begin{aligned} E[\log v_i] &= \Psi(v_{i,1}) - \Psi(v_{i,1} + v_{i,2}) \\ E[\log(1 - v_i)] &= \Psi(v_{i,2}) - \Psi(v_{i,1} + v_{i,2}) \end{aligned} \quad (61)$$

onde  $E[\log v_i]$  corresponde à esperança do logaritmo dos pesos da divisão do segmento da definição construtiva dos processos de Dirichlet para cada componente de mistura e  $\Psi$  representa a função digamma.

O quinto passo do ciclo é calcular o vetor dos componentes de mistura Gaussianos do modelo,  $\eta_{normal,T}$ , isto é, com os parâmetros definidos pela distribuição a priori Gaussiana  $p_{normal}$ ,  $\mu$  e  $\sigma$ , é calculado a distribuição Gaussiana associada a cada componente da mistura dos dados introduzidos segundo os parâmetros da distribuição a priori e como peso, a matriz multinomial  $z_{n,T}$ .

E uma vez feito esse cálculo é possível extrair as distribuições *T-student* associadas a cada componente da mistura do modelo,  $\eta_{t-student,T}$ . É definido por:

$$\begin{aligned} \eta_{normal,T} &= \text{Gaussiana}(\mu, \sigma | x_n, z_{n,T}) \\ \eta_{t-student,T} &= T - student(\mu_{\eta_{normal,T}}, \sigma_{\eta_{normal,T}} | x_n, z_{n,T}) \end{aligned} \quad (62)$$

O passo seguinte é atualizar os valores da matriz da multinomial dos segmentos da definição construtiva dos processo de Dirichlet  $z_{n,T}$ , mas, inicialmente é guardada uma versão da matriz  $z_{n,T}$ , como  $z'_{n,T}$ , antes de atualizar para compararmos com a atualização de  $z_{n,T}$  com o objetivo de verificar se foi atingida a convergência do modelo. A atualização é calculada através:

$$\begin{aligned} num &= \exp\left(E[\log(v_{i+1})] + \sum_{cum} E[\log(1 - v_{i-1})]\right) * p_{\eta_{T-student,T}}(x_n) \\ z &= \frac{num}{\sum(num) + p_{T-student}(x_n) * \frac{\exp(E[\log(w_i)] + \sum E[\log(1 - v_i)])}{1 - \exp(E[\log(1 - w_i)])}} \end{aligned} \quad (63)$$

onde  $p_{\eta_{T-student,k}}(x_n)$  corresponde à probabilidade dos valores introduzidos  $x_n$  segundo a distribuição T-student de cada componente de mistura do modelo e  $p_{t-student}(x_n)$  corresponder à probabilidade dos valores introduzidos  $x_n$  segundo a distribuição a priori T-student. E o passo final do ciclo é testar a convergência do modelo, isto é, verificar se o máximo da diferença entre a matriz multinomial antes da atualização  $z'_{n,T}$  e a matriz multinomial depois da atualização  $z_{n,T}$  é maior ou menor que  $\varepsilon$ :

$$\max\left(\left|\sum z'_{n,T} - z_{n,T}\right|\right) \quad (64)$$

se o máximo for menor que  $\varepsilon$ , o ciclo termina e significa que foi atingida a convergência do modelo mas se o máximo for maior que  $\varepsilon$  voltamos a repetir o ciclo até que seja atingida a convergência.

Este procedimento tem por base a definição de um limite máximo na definição construtiva dos processos de Dirichlet,  $T$ , escolhido no início do procedimento mas é possível fazer com que esse limite aumente à medida que procuramos os melhores parâmetros que definem o modelo de mistura Gaussiana por processos de Dirichlet.

Inicialmente fazemos o procedimento, idêntico ao em cima descrito, dos cálculos dos parâmetros do modelo com o limite máximo da definição construtiva dos processos Dirichlet  $T$  igual a 1 e depois calculamos o logaritmo negativo da verosimilhança dos dados em relação às distribuições do modelo atual. Uma vez feito esse cálculo, o limite máximo da definição construtiva dos processos de Dirichlet é aumentado em 1,  $T + 1$ , e é feito novamente o procedimento dos cálculos dos parâmetros do modelo de mistura Gaussianos e calculamos novamente o logaritmo negativo da verosimilhança dos dados. De seguida comparamos os dois valores, se o valor do logaritmo da verosimilhança dos dados do modelo com o parâmetro igual a  $T + 1$  for maior que o modelo com o parâmetro igual a  $T$  significa que o modelo com  $T + 1$  encontrou melhores parâmetros, que o modelo com o parâmetro igual a  $T$ , que definem melhor a distribuição dos dados introduzidos. E se o valor do logaritmo da verosimilhança dos dados do modelo com o parâmetro igual a  $T + 1$  for menor que o modelo com o parâmetro igual a  $T$ , este processo continua a repetir-se e a aumentar  $T$  até que esse modelo tenha um logaritmo negativo de verosimilhança dos dados maior que do modelo anterior. O cálculo do logaritmo negativo da verosimilhança dos dados em relação às distribuições do modelo atual é definido por:

$$\begin{aligned} \gamma_i &= \log(p_{\eta_{T-student,i}}(x_n)) + \log(w_i) \\ -\log(\mathcal{L}(\theta|x_n)) &= \gamma_i - \max(\gamma_i) + \sum \max(\gamma_i) + \sum \log\left(\sum e^{\gamma_i}\right) \end{aligned} \quad (65)$$

onde  $\gamma_i$  corresponde à soma do logaritmo das probabilidades dos dados segundo as distribuições T-student de cada componente de mistura do modelo,  $p_{\eta_{T-student,i}}(x_n)$ , e o logaritmo dos pesos associados a cada distribuição T-student de cada componente de mistura do modelo,  $w_i$  e  $\mathcal{L}(\theta|x_n)$  corresponde à função de verosimilhança dos dados em relação às distribuições do modelo de mistura Gaussianos por processos de Dirichlet.

Uma vez atingida a convergência do modelo é possível agora calcular a função de densidade de probabilidade preditiva de uma amostra segundo o modelo de mistura por processos de Dirichlet. É calculado através:

$$p(x|x_n, z_{n,T}) = \sum w_i \prod \left(1 - \frac{v_{i,1}}{\sum v_i}\right) p_{\eta_{T-student,k}}(x) + \prod \left(1 - \frac{v_{i,1}}{\sum v_i}\right) p_{T-student}(x) \quad (66)$$

onde  $p(x|x_n, z_{n,T})$  é a função de densidade de probabilidade preditiva do modelo,  $w_i$  corresponde o segmento da definição construtiva dos processos de Dirichlet,  $\prod \left(1 - \frac{v_{i,1}}{\sum v_i}\right)$  corresponde à divisão do segmento  $w_i$ ,  $p_{\eta_{T-student,k}}(x)$  corresponde à probabilidade da amostra segundo a distribuição T-student de cada componente de mistura do modelo e  $p_{T-student}(x)$  corresponde à probabilidade da amostra segundo a distribuição a priori T-student.

### 5.3 Aplicação do modelo de mistura Gaussianas em imagens cerebrais de ressonância magnética

Nesta secção é apresentada a aplicação do modelo de mistura Gaussianos por processos *Dirichlet* a imagens de tecidos cerebrais de ressonância magnética. Começamos por apresentar a base de dados das imagens que foi usada no trabalho experimental, o pré-processamento utilizado, os histogramas das intensidades dos tecidos cerebrais e a modelação dos dados através de modelos de mistura Gaussianas por processos de *Dirichlet*.

#### 5.3.1 Base de Dados MRBrainS

O *MICCAI Grand Challenge on MR Brain Image Segmentation 2013* [57] é um desafio que tem como objetivo de comparar algoritmos automáticos para a segmentação da massa cinzenta, substância branca e líquido cefalorraquidiano em exames de ressonância magnética do cérebro. Para o desafio, está disponível uma base de dados com vinte scans com múltiplas sequências de imagens (*T1-weighted*, com recuperação de inversão e *FLAIR*) MRI do cérebro de diferentes sujeitos. Os scans são divididos em cinco sujeitos de treino e quinze sujeitos de teste, com idades iguais ou superiores a 50 anos com diabetes, elevado risco cardiovascular e com diferentes graus de atrofia e lesões na substância branca. Para cada sujeito, há uma sequência de imagens *T1-weighted* (1mm x 1mm x 1mm), *T1-weighted* com registo na *T2-FLAIR* (0.958mm x 0.958mm x 3.0mm), *T1-weighted* com recuperação de inversão e com registo na *T2-FLAIR* (0.958mm x 0.958mm x 3.0mm) e *FLAIR* múltipla fatia (0.958mm x 0.958mm x 3.0mm). Também é fornecido marcações manuais feitas pelo médico, onde apresentam os diferentes tipos de tecidos cerebrais, como o líquido cerebrospinal, a substância cinzenta e substância branca. Estas marcações são o *ground truth* para comparar com a nossa segmentação [48]. Para este trabalho só foi utilizada a sequência de imagens *T1-weighted*.

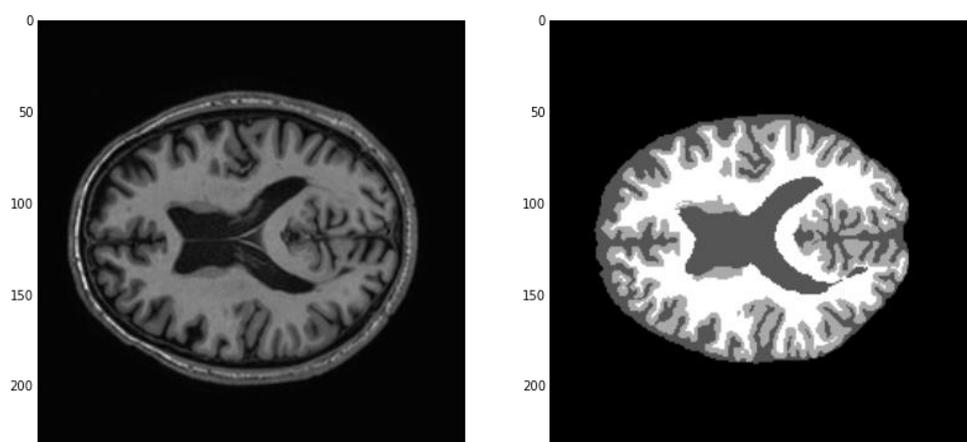


Figura 15 - Exemplo de imagens da base de dados MRBrainS. Esquerda: imagem *T1-weighted* 3D do MRBrainS, Direita: Marcação do médico (*Ground Truth*)

### 5.3.2 Pré-processamento

Antes de utilizarmos as imagens cerebrais de ressonância magnética foi feito um pré-processamento das mesmas. No pré-processamento foram usadas as mesmas etapas descritas no trabalho *Deteção automática de Espaços de Virchow-Robin em imagens de ressonância magnética* de Sérgio Pereira [49], que serviu de base para esta dissertação. O pré-processamento é dividido em quatro partes: na correção da falta de homogeneidade, na remoção do crânio, na normalização das intensidades e na remoção do ruído [49].

#### 5.3.2.1 Correção da falta de homogeneidade

A falta de homogeneidade nas imagens de ressonância magnética caracteriza-se por ser uma variação suave das intensidades, muitas vezes imperceptíveis a observadores, causada por imperfeições durante a aquisição. As possíveis causas são imperfeições nas bobinas de recepção de radiofrequência, correntes geradas durante a aplicação de gradientes do campo magnético, ou a anatomia dos pacientes, quer que se encontrem dentro ou fora do campo de observação [49].

A correção destes defeitos implica pressupor um modelo de formação da imagem; o mais utilizado assume que a imagem adquirida,  $v$ , é o resultado da imagem verdadeira,  $u$ , sujeita a um campo de polarização, que resulta na falta de homogeneidade multiplicativa,  $b$ , e da presença de ruído proveniente do *scanner*,  $n$ , e portanto independente da aquisição. A expressão matemática do modelo é definida por [49]:

$$v(x) = u(x)b(x) + n(x) \quad (67)$$

Foi implementado o método *Nick's Nonuniform Intensity Normalization* que é um método automático na correção da falta de homogeneidade baseado no histograma da imagem que pressupõe que a falta de homogeneidade esbate a imagem, devido à sua baixa frequência, por isso o objetivo é descobrir o campo de baixas frequências para que se maximize a frequência da imagem. Este método assume um modelo de formação da imagem que é livre de ruído e fazendo o logaritmo do modelo torna a falta de homogeneidade aditiva, transformando o modelo na seguinte equação [49]:

$$\hat{v}(x) = \hat{u}(x) + \hat{b}(x) \quad (68)$$

Em seguida, se  $V$ ,  $U$  e  $B$  forem as distribuições de probabilidades de  $\hat{v}$ ,  $\hat{u}$  e  $\hat{b}$ , respetivamente, a distribuição da imagem  $\hat{v}$  resulta da convolução de  $U$  com  $B$ . A distribuição de probabilidades  $V$  é calculada a partir do histograma da imagem,  $B$  é aproximada a partir de uma distribuição gaussiana e  $U$  deve ser estimada. Esta estimativa é calculada através da deconvolução entre  $V$  e uma gaussiana estreita, que representa  $B$ . Assim, pode-se estimar  $\hat{u}(E[\hat{u}|\hat{v}])$  que é subtraída à imagem corrompida,  $\hat{v}$ , resultando num mapeamento entre as intensidades da imagem e o campo da falta de homogeneidade,  $\hat{b}_s$ . Porém, este mapeamento deve ser suavizado com um operador  $S$ , usando *B-splines* [49].

$$\hat{b}_s(\hat{v}) = S\{v - E[\hat{u}|\hat{v}]\} \quad (69)$$

Este é um processo iterativo, que em cada iteração a imagem corrigida da iteração anterior é utilizada para estimar o campo de falta de homogeneidade  $\hat{b}_s$  a partir da imagem inicial, e estima novamente a imagem corrigida, definido por [49]:

$$\begin{aligned}\hat{u}^1 &= \hat{v} - S^*\{\hat{v} - E[\hat{u}|\hat{v}]\} \\ \hat{u}^2 &= \hat{u}^1 - S^*\{\hat{u}^1 - E[\hat{u}|\hat{u}^1]\} \\ &\vdots \\ \hat{u}^n &= \hat{u}^{n-1} - S^*\{\hat{u}^{n-1} - E[\hat{u}|\hat{u}^{n-1}]\}\end{aligned}\tag{70}$$

O processo repete as estimativas até que não se verifiquem mudanças em  $\hat{b}$  ou  $U$ , resultantes de deconvolução de gaussianas estreitas a  $V$ , ou seja, em cada iteração é estimada a falta de homogeneidade em relação à imagem corrigida na iteração anterior. [49]

### 5.3.2.2 Remoção do crânio

A remoção do crânio é importante num processo de segmentação de tecidos cerebrais. Neste sentido, pretende-se excluir da imagem os tecidos que não são relevantes, o que contribui para que a segmentação seja mais precisa, pois retiram-se possíveis falsos positivos. Neste tipo de procedimento pretende-se que, no final, apenas esteja presente o encéfalo e líquido cefalorraquidiano extracerebral [49].

O método optado é um método utilizado pelo *software Freesurfer*, *software* de visualização de imagens de ressonância magnética 3D, que segue uma abordagem híbrida com *watershed* e modelos deformáveis. Este método só é aplicado às sequências *T1-weighted* 3D. Ao remover-se o crânio de *T1-weighted* 3D, criou-se uma máscara, com base na imagem que apenas continha o encéfalo, e aplicava-se às restantes sequências [49].

### 5.3.2.3 Normalização das intensidades

As intensidades das imagens de ressonância magnética não são constantes entre aquisições, mesmo utilizando protocolos iguais, com a mesma pessoa e no mesmo *scanner*. Este facto torna-se relevante para processos de segmentação [49].

O método de normalização das intensidades de imagens de ressonância magnética é feito através da modificação do histograma, disponível no pacote de *software Insight Toolkit*. Este método utiliza um histograma de referência, e as intensidades de todas as outras imagens são modificadas para que os seus histogramas se transformem numa distribuição semelhante à referência. Desta forma, procura-se que as regiões do histograma que representem os mesmos tecidos fiquem alinhadas, com o objetivo de tornar as intensidades e o contraste mais consistente entre os sujeitos. As sequências são normalizadas independentemente, isto é, para normalizar *T1-weighted* 3D deve-se escolher uma referência *T1-weighted* 3D, enquanto para FLAIR deve ser escolhida uma referência FLAIR [49].

A forma de transformação do histograma pressupõe a existência de um histograma de referência, e um conjunto de marcas que servirão para fazer o emparelhamento entre as distribuições. Começa-se, então, por calcular a intensidade mínima e máxima do histograma,  $m_1$  e  $m_2$ , respetivamente, e por se excluir as caudas do mesmo, uma vez que geralmente correspondem a *outliers*. As intensidades a excluir são aquelas inferiores ao primeiro percentil,  $pc_1$ , e superiores ao último percentil,  $pc_2$ , e o intervalo entre estes dois percentis é designado por intensidade de interesse. Em seguida deve-se seleccionar as marcas do histograma,  $\mu_p$ , onde  $p$  representa o percentil. Depois de determinados os parâmetros anteriores para o histograma de referência obtém-se uma escala de referência, delimitada pelo intervalo  $[s_1 = pc_1, s_2 = pc_2]$ . Então, para normalizar as intensidades de uma nova imagem, deve-se determinar as marcas no seu histograma, e mapeá-las linearmente para a escala de referência, no intervalo  $[s_1, s_2]$ . Em seguida, faz-se um novo mapeamento linear, desta vez entre marcas. Assim sendo, se o índice  $i$  representar marcas da imagem que se pretende normalizar, o índice  $k$  designa marcas do histograma de referência, de tal forma que  $i$  e  $k$  tenham a mesma gama de valores, ou seja, o número de marcas na referência e na imagem a normalizar deve ser o mesmo, e existirem  $l$  marcas  $\mu$ , então começa-se por mapear  $[pc_{1i}, \mu_{1i}]$  para  $[s_{1k}, \mu_{1k}]$ , depois  $[\mu_{1i}, \mu_{2i}]$  para  $[\mu_{1k}, \mu_{2k}]$ , e assim sucessivamente até  $[\mu_{li}, pc_{2i}]$  para  $[\mu_{lk}, s_{2k}]$ . A expressão para o mapeamento linear das intensidades de  $[pc_{1i}, pc_{2i}]$  para  $[s_1, s_2]$  é definido por [49]:

$$x' = s_1 + \frac{x - pc_{1i}}{pc_{2i} - pc_{1i}}(s_2 - s_1) \quad (71)$$

onde  $x$  é a intensidade original e  $x'$  é a intensidade mapeada. Esta expressão pode ser utilizada para o mapeamento entre marcas, fazendo as necessárias alterações nas variáveis, para essas marcas. Durante o processo de normalização foram excluídos os *voxels* de fundo. Na Figura 16(a) encontra-se um exemplo de colocação das marcas num histograma. e na Figura 16(b) pode-se observar, graficamente, o processo de normalização linear entre marcas [49].

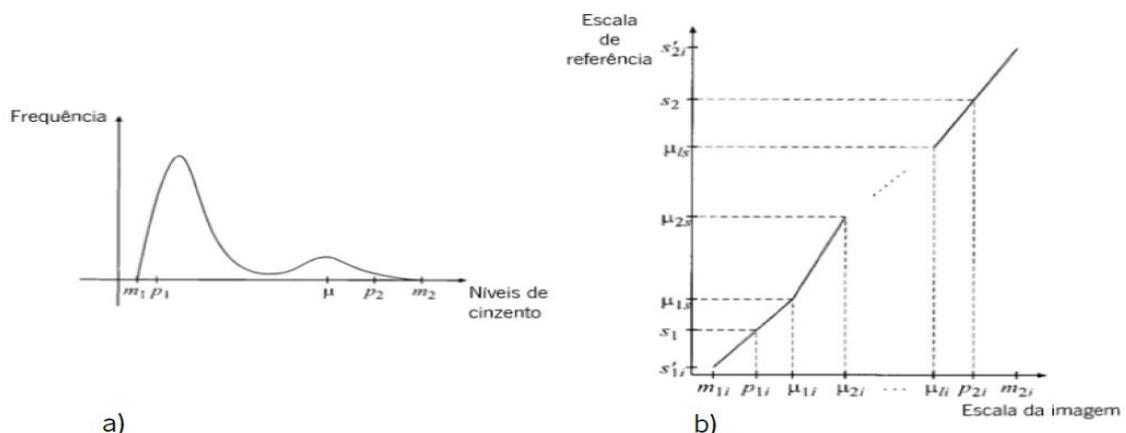


Figura 16 - Normalização das intensidades. a) Exemplo de escolha das marcas numa distribuição bimodal, na qual se optou por colocar a marca na segunda moda. b) demonstração do mapeamento linear entre o histograma da imagem que se pretende normalizar e o histograma de referência [49].

### 5.3.2.4 Remoção do ruído

Na aquisição de imagens de ressonância magnética existe um compromisso entre a resolução e a relação sinal-ruído. Na ressonância magnética, o ruído segue uma distribuição Rician. Foi utilizado um filtro denominado por *Prefiltered Rotationally Invariant Non-Local Mean Filter* devido ao facto de melhorar a qualidade da segmentação de substância branca, substância cinzenta e líquido cefalorraquidiano em imagens com elevado ruído [49].

Pode-se filtrar o ruído de uma imagem de ressonância magnética calculando a sua transformada discreta do cosseno, para cada bloco 3D (4x4x4). Em seguida, aplica-se um *threshold* igual a  $2,7\sigma$ , onde  $\sigma$  é a estimativa do desvio padrão do ruído, para se obter uma primeira estimativa da imagem filtrada. [49]

Por outro lado, a ideia do *Non-Local Mean Filter* é que numa imagem existe informação redundante, isto é, há muitas regiões semelhantes. As regiões denominam-se *patches*, sendo constituídas pelo *voxel* sob análise (central) e a sua vizinhança 3D. O processo de remoção de ruído consiste na substituição do *voxel* pela média das intensidades dos *voxels* centrais dos *patches* que são parecidos entre si. No entanto, podem existir *patches* semelhantes, mas rodados, sendo necessário utilizar medidas de rotação para os descobrir, passando a denominar-se *Rotational Invariant Non-Local Mean Filter* [49]. Este novo filtro é poderoso, pois restaura a imagem sem afetar muito a sua estrutura, mas é sensível à presença de muito ruído para estimar os *patches* mais idênticos.

O filtro *Prefiltered Rotationally Invariant Non-Local Mean* começa por filtrar a imagem utilizando o método da transformada discreta do cosseno. Posteriormente utiliza essa imagem pré-filtrada para identificar os *patches* que são semelhantes, de forma invariante à rotação, e utiliza-os para remover o ruído da imagem original com a abordagem *Non-Local Means*. Assim, consegue-se tirar partido do *Rotational Invariant Non-Local Mean Filter*, mesmo em imagens com muito ruído. A distribuição Rician para o ruído é tida em consideração quer no cálculo da semelhança dos *patches*, quer na estimativa do ruído no método da transformada discreta do cosseno [49].

Os parâmetros utilizados para o filtro foram: a dimensão do *patch* foi de 3x3x3, que o filtro deve fazer a procura de *patches* semelhantes no espaço 7x7x7 que envolve o *voxel* que está a ser analisado e que utiliza o modelo de ruído Riciano [49].

### 5.3.2.5 Resultado do pré-processamento

Uma vez feito o pré-processamento das imagens de ressonância magnética para a correção da falta de homogeneidade, a remoção do crânio, a normalização das intensidades e a remoção do ruído pode-se verificar que através da observação da Figura 17 que o resultado obtido melhora bastante a imagem de ressonância para o procedimento de segmentação.

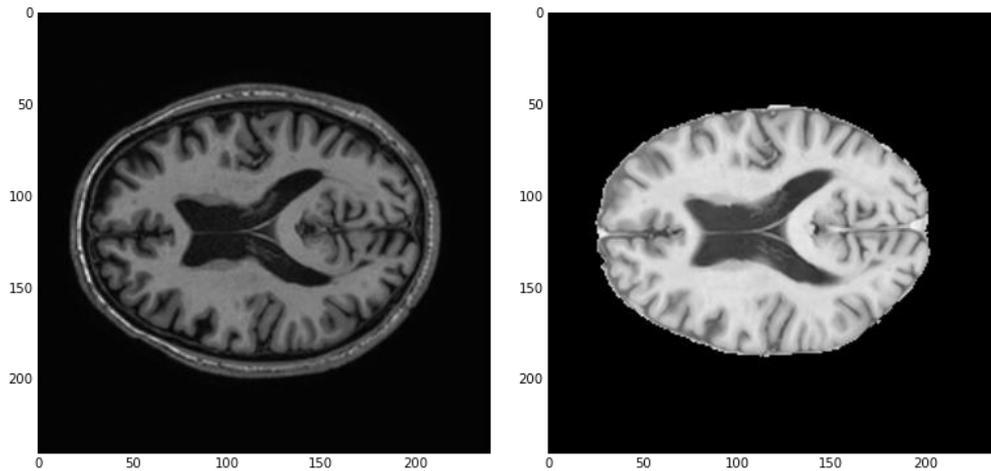


Figura 17 – Comparação da mesma imagem de ressonância magnética antes e depois do pré-processamento: Esquerda – imagem original, Direita – imagem processada

### 5.3.3 Histogramas das intensidades dos tecidos cerebrais

Através da marcação do médico, *ground truth*, é possível retirar-se as intensidades da imagem normalizada T1-weighted 3D segundo a classificação de cada tipo de tecido cerebral, ou seja, são extraídas as intensidades do líquido cefalorraquidiano, da substância cinzenta e da substância branca. E com estas intensidades construímos um histograma para ver a gama de intensidades de cada tecido. A Figura 18 apresenta o histograma das intensidades de cada tecido cerebral.

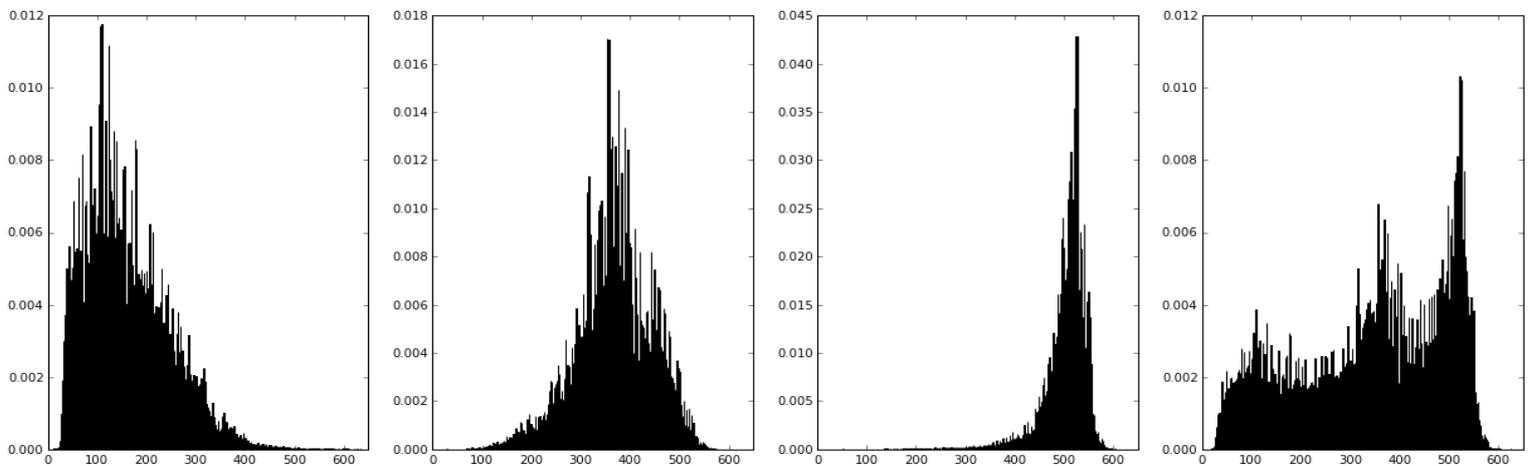


Figura 18 - Histograma das intensidades de cada tecido cerebral. 1º: Líquido Cefalorraquidiano, 2º: Substância Cinzenta, 3º: Substância Branca, 4º: Todos os tecidos

Antes de introduzirmos as intensidades de cada tecido cerebral no nosso modelo de mistura, primeiro foi feita a verificação se os dados de cada tecido seguem uma distribuição Gaussiana ou normal porque o nosso modelo trabalha com essa distribuição e se é possível reproduzir os dados segundo a distribuição. Foram feitos dois testes diferentes: o primeiro foi amostrar aleatoriamente os dados de cada tecido e depois faz-se um teste da combinação da inclinação e da curtose das amostras para verificar se as amostras se ajustam a uma distribuição Gaussiana.

Inicialmente foram extraídas aleatoriamente amostras pequenas e depois foi-se aumentando o número de amostras, por exemplo, 30, 100, 500, 1000, 5000, 10000, ..., porque segundo o Teorema de Limite Central, à medida que aumentamos o número de amostras aleatórias, a distribuição das médias amostrais tende para uma distribuição normal. Foi verificado que para amostras pequenas, entre 30 a 100 amostras, não se rejeitava a hipótese de ajuste uma vez que o p-valor era superior a 0.05, se assumirmos um intervalo de confiança a 95% mas para amostras superiores, amostras entre 500 a 10000, o p-valor do teste era inferior a 0.05 e com isso podemos concluir que os dados não seguem uma distribuição Gaussiana mas com o nosso modelo de mistura Gaussianas por processos de *Dirichlet* podemos colocar a hipótese de que os dados seguem uma mistura de distribuições Gaussianas. O segundo teste efetuado foi tentar adaptar uma distribuição Gaussiana e *T-student* para cada gama de intensidades de cada tecido cerebral, isto é, foi estimada a média e o desvio padrão dos dados das intensidades para cada tipo de tecido segundo as distribuições Gaussiana e *T-student* e de seguida calcular as suas funções de densidade.

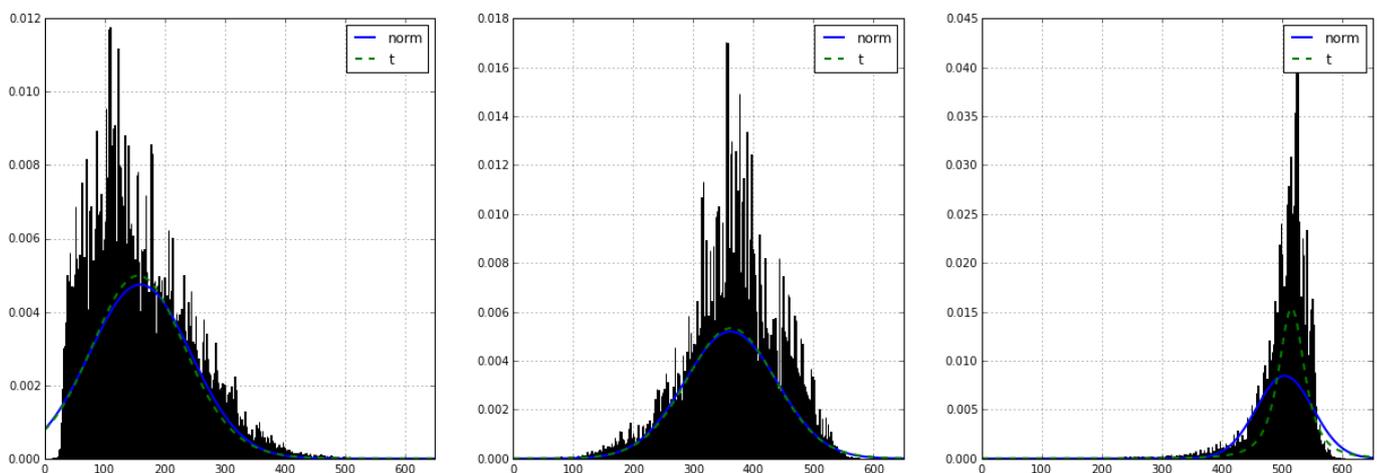


Figura 19 - Teste de adaptação da distribuição Gaussiana e *T-student*. Esquerda: Intensidades do Líquido Cerebrospinal, Meio: Intensidades da Substância Cinzenta, Direita: Intensidades da Substância Branca

A figura 19 representa o teste de adaptabilidade das intensidades dos tecidos cerebrais a uma distribuição Gaussiana e a uma distribuição *T-student*. A linha azul representa o resultado do ajuste das intensidades à distribuição Gaussiana e linha verde representa o resultado do ajuste das intensidades à distribuição *T-student*. Como podemos observar os resultados dos ajustes são aceitáveis contudo apresentam alguns defeitos. Na imagem da esquerda pode-se verificar que só a parte inicial do histograma, entre 0 e 50, não está bem representada pelas duas distribuições. Na imagem do meio, as duas distribuições representam bem a distribuição dos valores do histograma. E na imagem da direita, na gama dos 600, as duas distribuições não conseguem representar corretamente a distribuição do histograma mas a *T-student* representa melhor o histograma do que a distribuição Gaussiana. Com base nestes resultados assumimos que as intensidades têm uma maior probabilidade de seguir uma mistura de distribuições Gaussianas.

### 5.3.4 Modelação dos dados através de modelos de mistura Gaussianas por processos de Dirichlet

Nesta secção são apresentados os resultados obtidos com a modelação através de modelos de mistura Gaussianos por processos de *Dirichlet*. Foram introduzidos os dados das intensidades de cada tipo de tecido cerebral no modelo e de seguida foi calculado o modelo de mistura Gaussiano por processos de *Dirichlet* para o líquido cerebrospinal, para a substância cinzenta e para a substância branca.

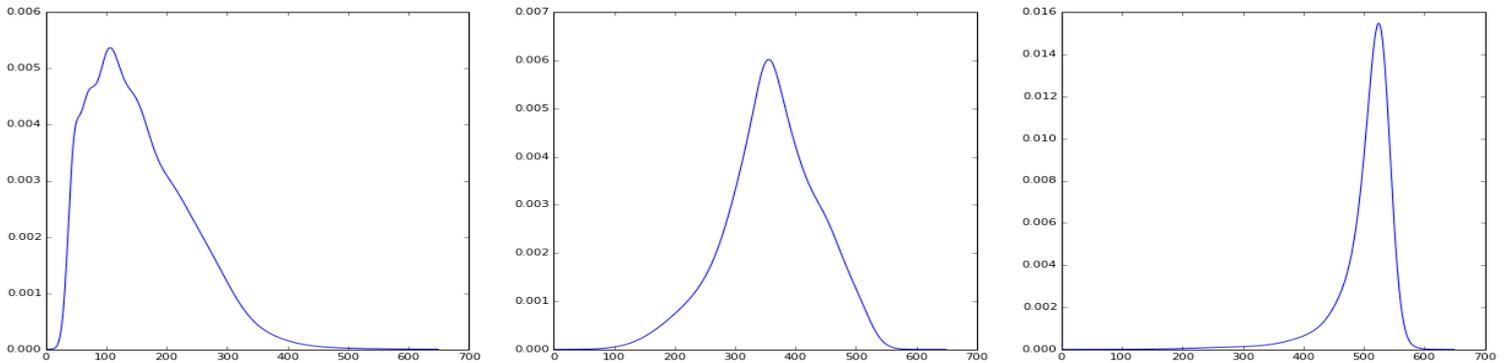


Figura 20 - Exemplo para cada tipo de tecido cerebral a probabilidade preditiva dos dados. Esquerda: Líquido Cerebrospinal. Meio: Substância Cinzenta. Direita: Substância Branca

A figura 20 representa a probabilidade preditiva de cada tipo de tecido cerebral do sujeito 1. Comparando a probabilidade preditiva com o histograma das intensidades verifica-se que o modelo de mistura Gaussianas por processos de *Dirichlet* consegue seguir a distribuição dos dados das intensidades com poucos defeitos. Com o modelo calculado são introduzidos os valores das intensidades dos diferentes sujeitos e uma vez calculada a probabilidade preditiva dos dados é construído um mapa de probabilidades para cada tipo de tecido para cada intensidade de cada píxel das sequências das imagens.

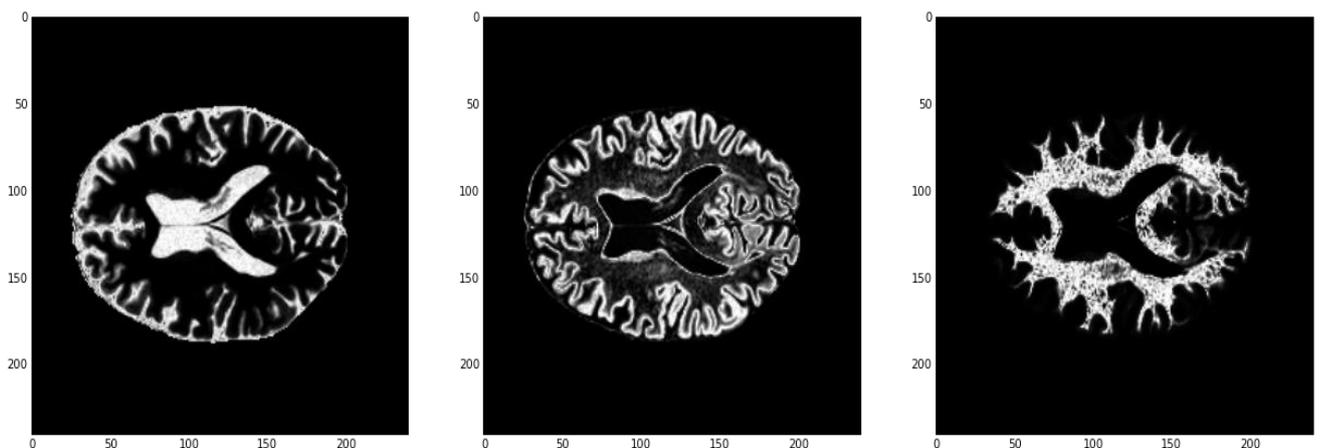


Figura 21 - Exemplo de mapa de probabilidade de um sujeito. Esquerda: Líquido Cerebrospinal, Meio: Substância Cinzenta, Direita: Substância Branca

A figura 21 apresenta os mapas de probabilidades, de um sujeito, das intensidades de cada píxel numa imagem das sequências das imagens T1-weighted 3D. Cada píxel da imagem representa a probabilidade desse píxel pertencer a um determinado tecido cerebral.

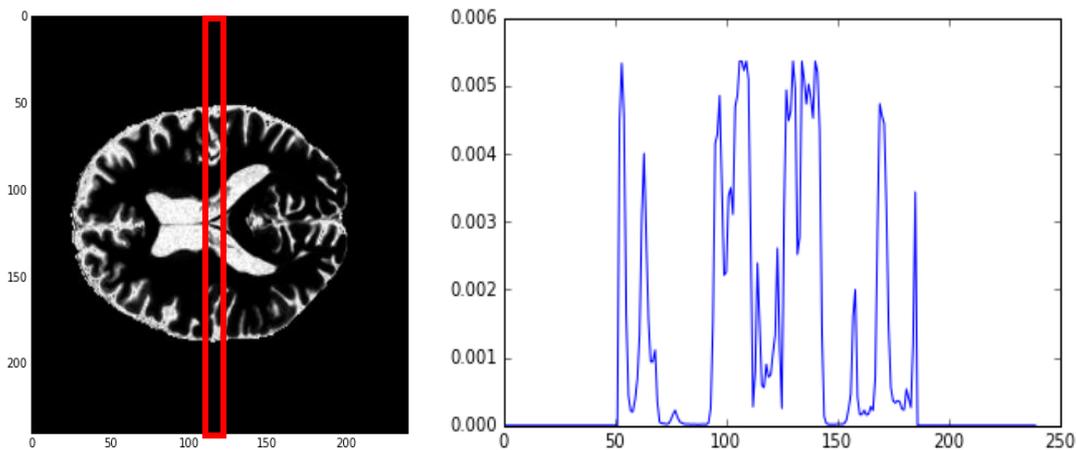


Figura 22 – Exemplo dos valores obtidos numa secção do mapa de probabilidades do tecido líquido cerebrospinal

Com base nestes mapas é feito uma comparação entre eles, ou seja, percorremos mapas de probabilidade e comparamos pixel a pixel entre cada tipo de tecido cerebral e o que tiver o valor mais alto é marcado numa nova imagem todas as imagens da sequência dos como o tecido que teve a maior probabilidade para esse pixel e é assim efetuada a segmentação dos tecidos cerebrais das imagens de ressonância magnética. A figura 22 representa uma demonstração dos valores obtidos das probabilidades das intensidades do modelo de mistura Gaussianas do tecido líquido cerebrospinal de uma secção do mapa de probabilidades.

Na figura 23 é apresentada a comparação da segmentação de um sujeito, obtida pela comparação dos mapas de probabilidades obtidos através da modelação dos dados das intensidades por modelos de mistura por processos de *Dirichlet* com a marcação do médico, com a imagem da ressonância magnética pré-processada. Com base na observação da figura 23 b) e d), verifica-se que a segmentação proposta apresenta um bom resultado mas com alguns defeitos, demonstrados na figura 23 c). A cor azul representa a marcação errada do tecido líquido cerebrospinal, a cor verde representa a marcação errada do tecido substância cinzenta e a cor vermelha representa a marcação errada do tecido substância branca. Verifica-se, através da figura 23 c), que os defeitos mais predominantes da segmentação é da marcação do tecido substância cinzenta e do tecido substância branca mas a marcação do tecido líquido cerebrospinal também tem algumas falhas. Isto deve-se ao facto de as intensidades dos tecidos cerebrais se sobreporem, isto é, os diferentes valores das intensidades de cada tecido cerebral podem estar dentro da mesma gama de valores de outro tecido cerebral e a partir dessa sobreposição, as suas probabilidades preditivas de cada tecido cerebral também vão conter essa sobreposição.

Com base nestes problemas descritos, nas próximas secções é proposto modificações aos modelos para se adaptarem às distribuições dos histogramas das intensidades dos diferentes tecidos cerebrais e os resultados desta segmentação e das modificações propostas.

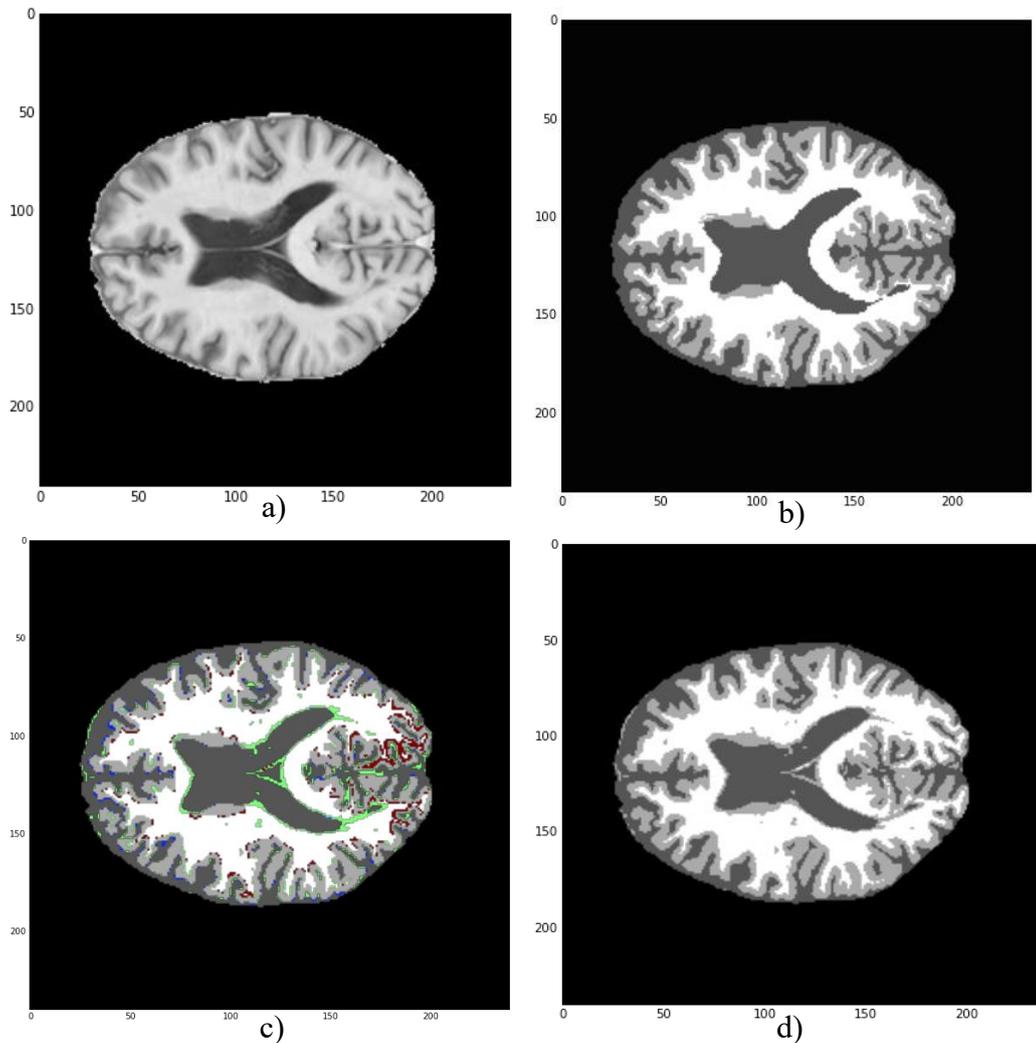


Figura 23 - Exemplo da segmentação dos tecidos cerebrais por modelos de mistura Gaussianos por processos de Dirichlet. a) Imagem da ressonância magnética b) Marcação feita pelo médico, c) Segmentação pelo modelo com marcações, d) Segmentação pelo modelo

#### 5.4 Modificação dos modelos de mistura Gaussianas por processos de Dirichlet

Nesta secção são propostas diferentes modificações aos modelos de mistura Gaussianas por processos de Dirichlet para se adaptarem melhor às distribuições dos histogramas das intensidades dos diversos tipos de tecido cerebrais. Na secção anterior foram apresentados os defeitos da segmentação dos tecidos cerebrais através da comparação dos mapas de probabilidade das intensidades segundo modelos de mistura Gaussianos por processos de Dirichlet onde a causa problemática da segmentação se deve ao facto de as distribuições das intensidades dos diferentes tecidos cerebrais terem uma gama de valores de intensidades que se sobrepõem umas com as outras o que vai afetar a probabilidade de um valor de intensidade pertencer a um tipo de tecido cerebral incorretamente.

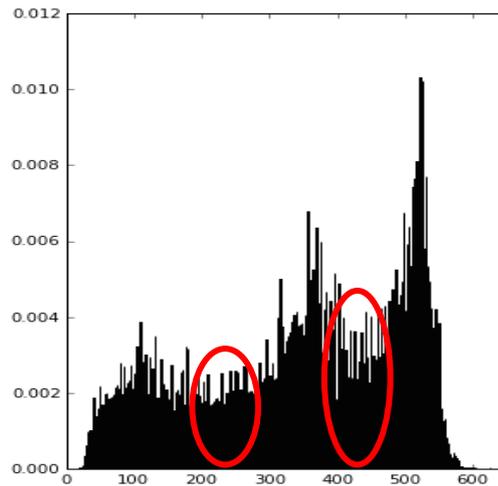


Figura 24 - Demonstração da sobreposição das intensidades para os diferentes tecidos cerebrais

Verifica-se que, com base na figura 24, existem duas sobreposições das intensidades: uma no intervalo da gama das intensidades 200 e 300, que significa uma sobreposição das intensidades do líquido cefalorraquidiano com a substância cinzenta. E outra sobreposição, no intervalo da gama das intensidades 400 e 500, que significa uma sobreposição das intensidades da substância branca com a substância cinzenta. E como estas intensidades correspondem aos dados usados para os modelos de mistura Gaussianos por processos de *Dirichlet* estes, conduzem a que as probabilidades preditivas se sobreponham também.

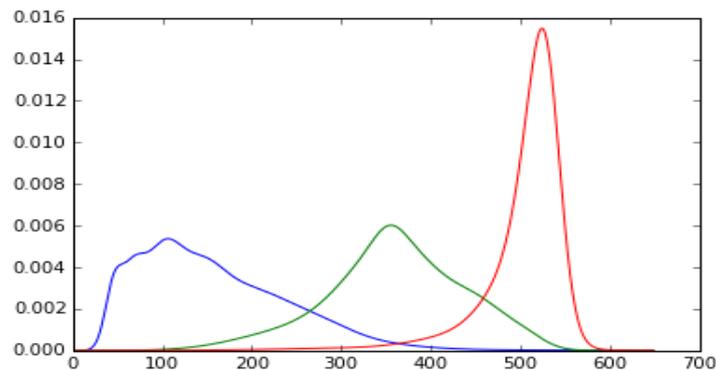


Figura 25 - Demonstração da sobreposição das probabilidades preditivas dos tecidos cerebrais. Azul: Líquido Cefalorraquidiano, Verde: Substância Cinzenta, Vermelha: Substância Branca

A figura 25 representa a sobreposição das probabilidades preditivas das intensidades segundo os modelos de mistura Gaussianos por processos de *Dirichlet*, onde se observa as sobreposições nos mesmos intervalos da gama das intensidades dos histogramas das intensidades. Como consequência nas próximas seções são propostas adaptações das distribuições usadas pelo modelo de mistura por processos de *Dirichlet*.

### 5.4.1 Modificação do modelo para o líquido cerebrospinal

Como uma das sobreposições se situa entre o líquido cerebrospinal e a substância cinzenta, fomos analisar melhor as distribuições dos histogramas do líquido cerebrospinal e da substância cinzenta.

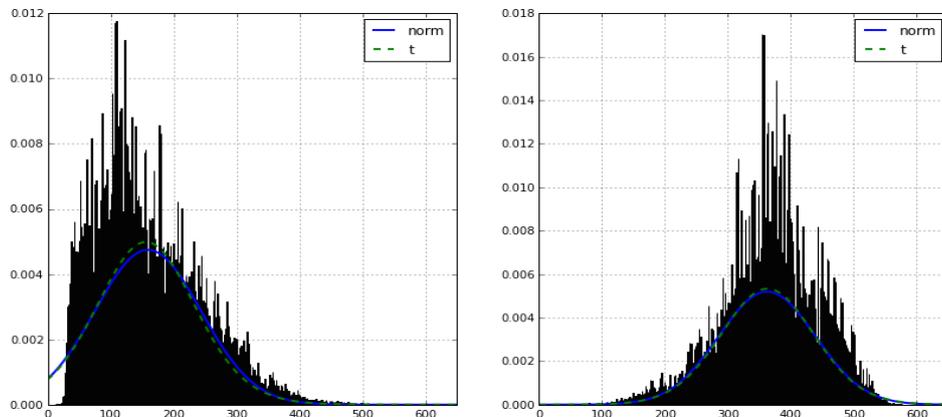


Figura 26 - Comparação do teste de adaptação da distribuição gaussiana e *t*-student com os histogramas do líquido cerebrospinal e da substância cinzenta. Esquerda: Líquido Cerebrospinal, Direita: Substância Cinzenta

Ao observar a figura 26, verificamos que as duas distribuições, quer a Gaussiana quer a *T*-student, representam melhor as intensidades da substância cinzenta do que o líquido cerebrospinal porque o histograma das intensidades da substância cinzenta aparenta ter a forma da distribuição Gaussiana e *T*-student. Podemos verificar que o histograma das intensidades do líquido cerebrospinal não tem uma forma semelhante às distribuições Gaussiana e *T*-student e isso reflete-se no teste de ajuste das distribuições, onde no intervalo da gama das intensidades 0 a 40, as distribuições não conseguem representar corretamente esse intervalo e que a curva à direita do centro das distribuições também não consegue representar os dados no mesmo local da curva. Por isso investigamos o uso de outras distribuições que o nosso modelo de mistura por processos de *Dirichlet* usa para adaptarmos melhor aos dados. Com esse objetivo, encontramos uma distribuição que segue a distribuição do histograma das intensidades do líquido cerebrospinal, a distribuição Rician. Utilizamos esta distribuição para aproximar o modelo à distribuição do histograma e a curva da Rician diminui mais rapidamente do que as distribuições Gaussianas e *T*-student para que ocorra menos sobreposições de probabilidades preditivas.

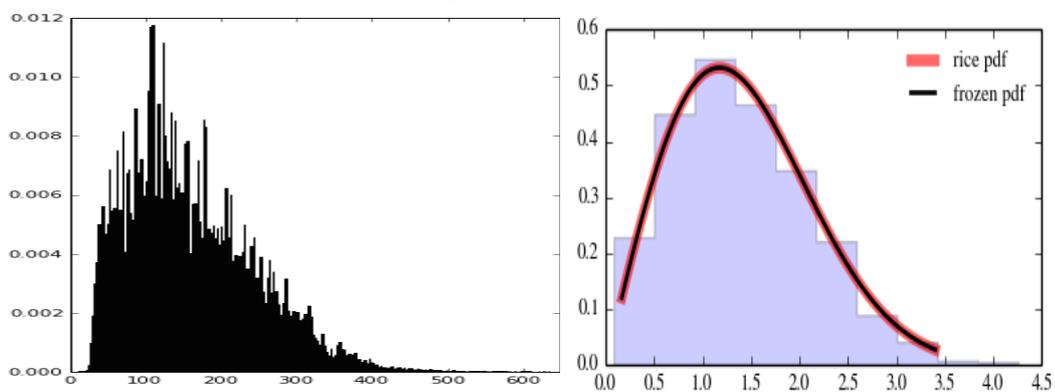


Figura 27 - Comparação entre o histograma do líquido cerebrospinal com um exemplo de uma distribuição Rician [54]

Antes de alterar o modelo para trabalhar com distribuições Ricianas foi feito um teste de ajuste da distribuição com os dados das intensidades do líquido cerebrospinal.

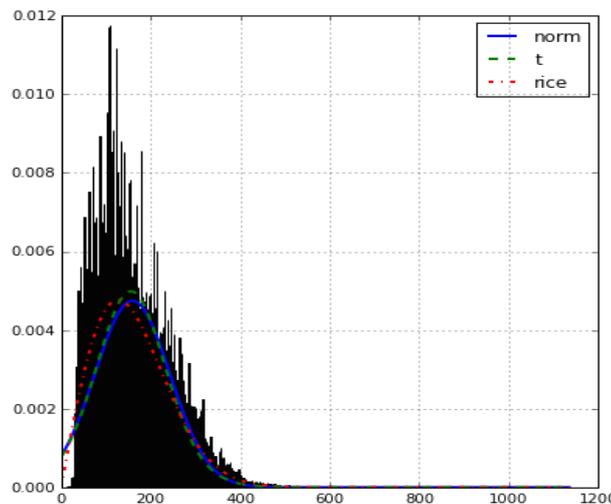


Figura 28 - Comparação do teste de adaptação das distribuições às intensidades do líquido cerebrospinal. Azul: Distribuição Gaussiana, Verde: Distribuição T-student, Vermelho: Distribuição Riciana

Com a figura 28 é possível observar que a distribuição Riciana representa um pouco melhor que as distribuições Gaussianas e T-student devido ao facto de a distribuição resultante do teste de adaptação mostrar que está mais centrada onde se concentram mais os dados das intensidades. Com a propriedade de o modelo de misturas por processos de *Dirichlet* trabalhar com distribuições de famílias exponenciais e convexas é possível alterar a distribuição com que trabalha para outras distribuições da família exponencial. A modificação do modelo é feita com a substituição da distribuição Gaussiana e T-student pela distribuição Riciana nas equações que utilizam distribuições para o cálculo do modelo de mistura por processos de *Dirichlet*. O procedimento modificado estima os parâmetros da distribuições Riciana como distribuição *a priori* e depois utilizamos os parâmetros para a distribuição Riciana em cada componente de mistura do modelo para usarmos no cálculo do modelo de mistura por processos de Dirichlet, definido por:

$$\begin{aligned} p_{Riciana} &\sim Riciana(v, \sigma) \\ \eta_{Riciana} &\sim Riciana(v_t, \sigma_t) \end{aligned} \quad (72)$$

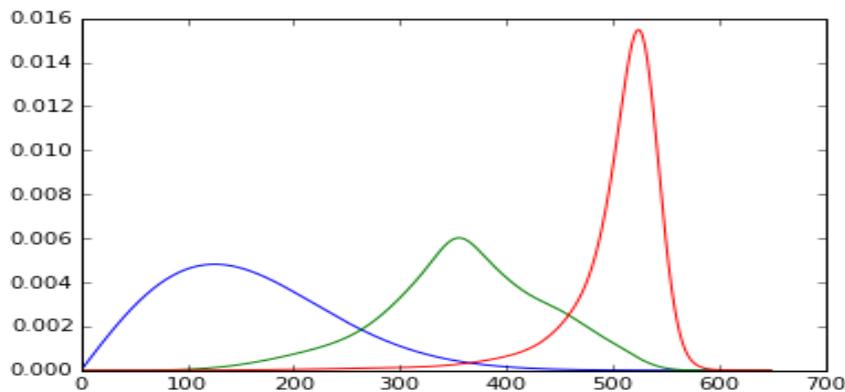


Figura 29 - Exemplo das probabilidades preditivas dos tecidos cerebrais. Azul: Líquido Cerebrospinal, Verde: Substância Cinzenta, Vermelha: Substância Branca

A figura 29 representa as probabilidades preditivas dos tecidos cerebrais de um sujeito. A linha azul corresponde à probabilidade do líquido cerebrospinal obtido através de modelos de mistura Ricianas por processos de *Dirichlet*. A linha verde corresponde à probabilidade da substância cinzenta e a linha vermelha corresponde à probabilidade da substância branca, ambas obtidas através de modelos de mistura Gaussianos por processos de *Dirichlet*. Com os modelos calculados foram feitos os mapas de probabilidade das intensidades para cada modelo de cada tecido cerebral e a sua segmentação.

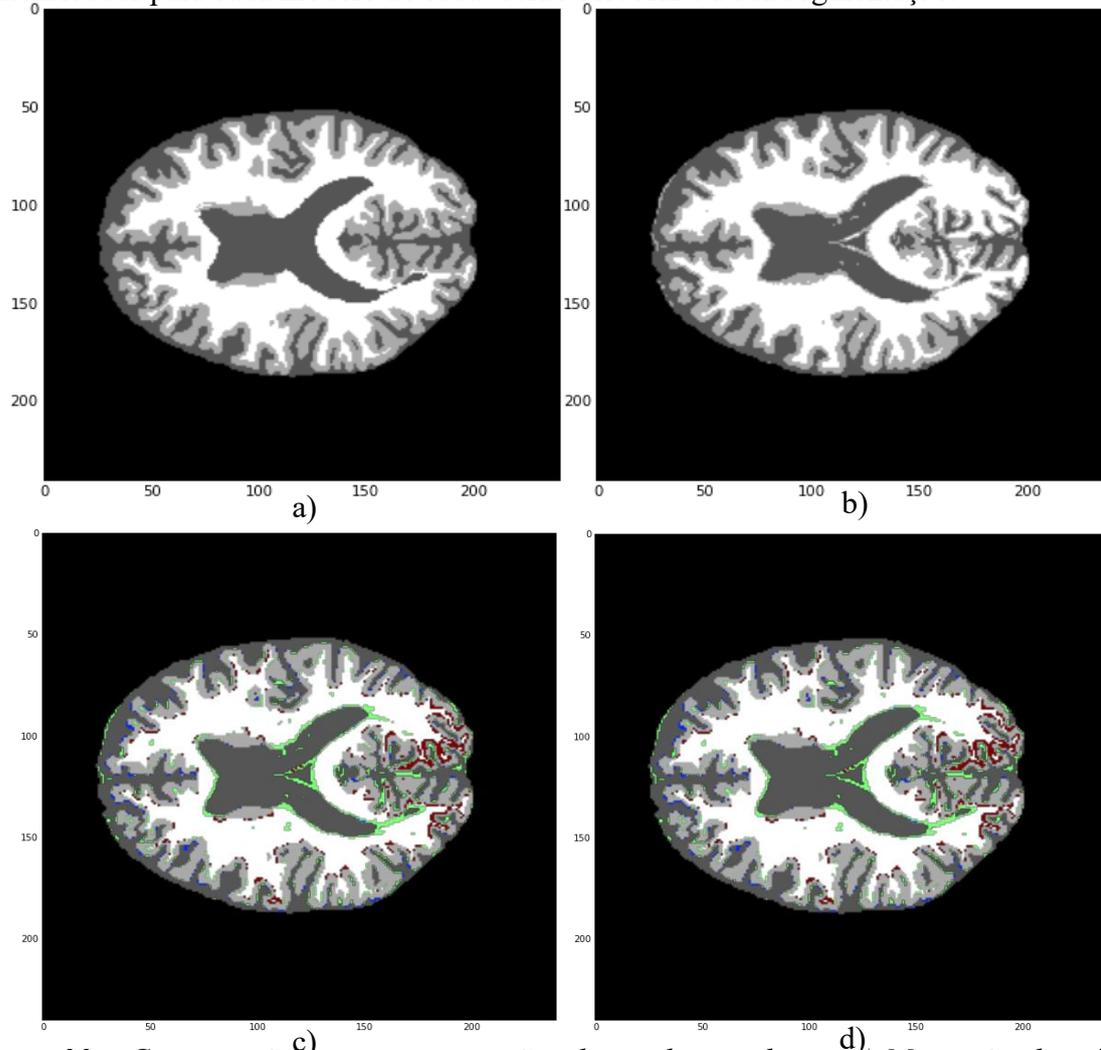


Figura 30 – Comparação entre segmentações de tecidos cerebrais. a) Marcação do médico, b) Segmentação por combinação de modelos, c) Segmentação por combinação de modelos com marcações, d) Segmentação por modelos de mistura Gaussianos com marcações

A figura 30 apresenta a comparação de segmentações dos tecidos cerebrais de imagens de ressonância magnética. A figura 30 a) representa a marcação do médico, a figura 30 b) representa a segmentação dos tecidos cerebrais com a combinação dos modelos de mistura Gaussiana por processos de Dirichlet para os tecidos substância cinzenta e branca e o modelo de mistura Ricianas por processos de Dirichlet para o tecido líquido cerebrospinal, a figura 30 c) representa a segmentação da figura 30 b) mas com marcações das zonas mal classificadas obtidas através da comparação com a marcação do médico, onde a cor azul

corresponde ao tecido líquido cerebrospinal, a cor verde corresponde ao tecido substância cinzenta e a cor vermelha corresponde ao tecido substância branca e a figura 30 d) representa segmentação por modelos de mistura Gaussianas por processo de Dirichlet com marcações das zonal mal classificadas. Com base na observação da figura 30 verifica-se que só com base na visualização não é possível chegar a uma conclusão se há melhoria na classificação do líquido cerebrospinal e nas diferenças entre segmentações. Por isso, nas próximas secções são apresentados os resultados através de métricas que avaliam as diferenças entre segmentações.

#### 5.4.2 Modificação do modelo para a substância branca

Nesta secção abordamos o problema da sobreposição entre a probabilidade preditiva da substância cinzenta e da substância branca. Foi efetuado o mesmo procedimento que na modificação para o modelo do tecido líquido cerebrospinal.

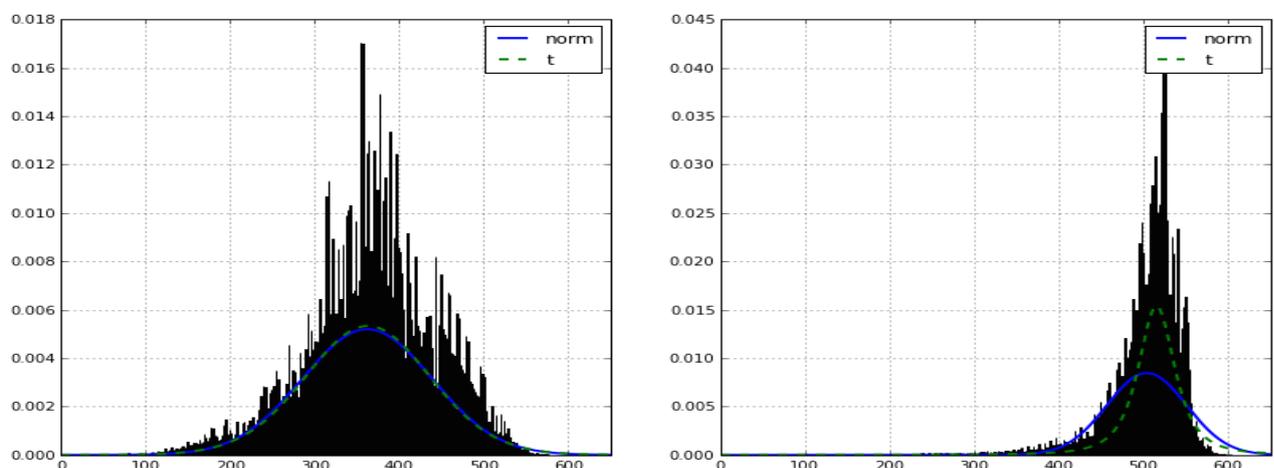


Figura 31 - Comparação do teste de adaptação da distribuição gaussiana e *t*-student com os histogramas da substância cinzenta e da substância branca. Esquerda: Substância Cinzenta, Direita: Substância Branca

Ao observar a figura 31, verificamos que a substância branca não tem propriamente a forma da distribuição Gaussiana mas a distribuição *T-student* consegue adaptar-se melhor que a distribuição Gaussiana, ainda assim tem alguns defeitos na distribuição do histograma como a curva à esquerda da maior concentração dos dados do histograma não corresponde. Devido a esse facto fomos à procura de distribuições que tenham o formato do histograma das intensidades da substância branca para minimizar a sobreposição entre intensidades e melhorar a probabilidade preditiva. Foram encontradas várias distribuições mas só são apresentadas 5 distribuições porque apresentaram uma melhor adaptação que as outras. As distribuições utilizadas são a *Genlogistic*, a *Gumbel* enviesada à esquerda, a *Johnson Su*, a *Powernorm* e a *Log Gama*. A figura 32 mostra a comparação do histograma das intensidades da substância branca com as distribuições referenciadas.

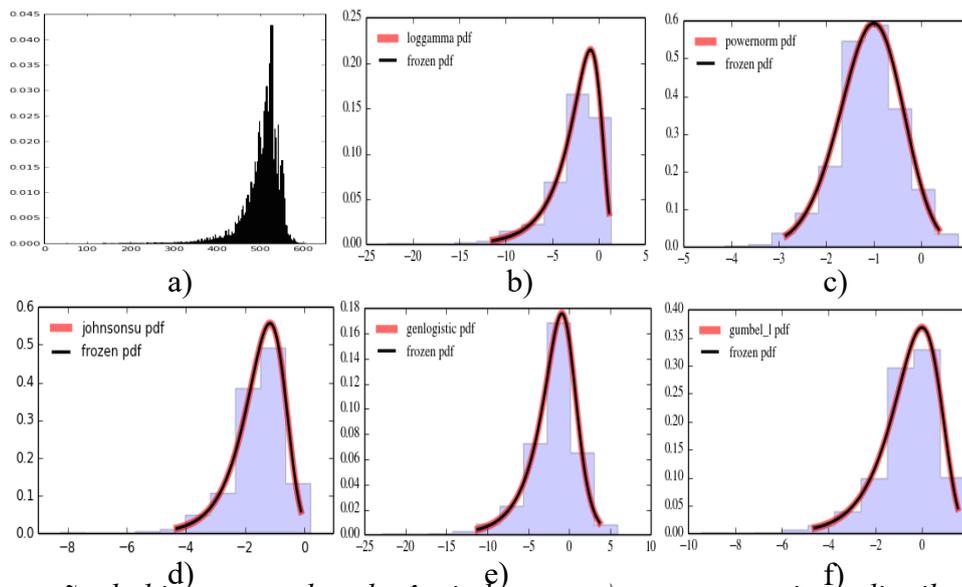


Figura 32 - Comparação do histograma da substância branca, a), com as seguintes distribuições: b) Log Gama [17], c) Powernorm [28], d) Johnson Su [19], e) Genlogistic [26], f) Gumbel enviesada à esquerda [24]

Com base na figura 32, verificamos que existem semelhanças do histograma com as distribuições apresentadas e com isso decidimos fazer testes de adaptação segundo essas distribuições com os dados para verificarmos se essas distribuições conseguem representar melhor os dados das intensidades que o modelo utilizado. A figura 33 apresenta os testes de adaptação para cada distribuição com os dados das intensidades da substância branca.

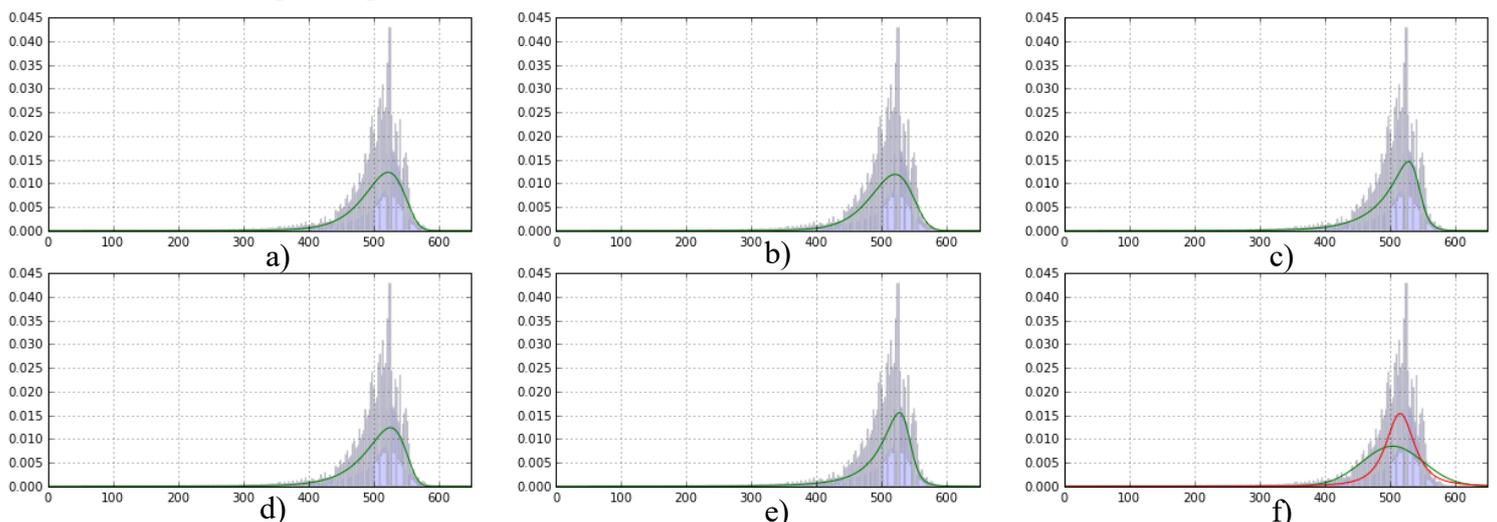


Figura 33 - Testes de adaptação das distribuições às intensidades correspondentes à substância branca. Distribuições: a) Gumbel enviesada à esquerda, b) Powernorm, c) Genlogistic, d) Log Gama, e) Johnson Su, f) Gaussiana (verde) e T-student (vermelha)

Com a comparação dos testes, baseada na figura 33, podemos chegar à conclusão que as distribuições propostas conseguem uma melhor estimação dos dados que a Gaussiana e T-student ficando mais centradas nos valores de maior probabilidade do histograma e a subida da distribuição é semelhante ao histograma das intensidades, embora a distribuição T-student consegue uma melhor representação que a distribuição Gaussiana.

Com base nestas conclusões, foi alterado o modelo de mistura para trabalhar com estas distribuições. A alteração do modelo é feita pela substituição da distribuição Gaussiana e *T-student* por cada distribuição proposta nas equações que utilizam distribuições para o cálculo do modelo de mistura por processos de *Dirichlet*, isto é possível devido à propriedade do modelo de misturas por processos de *Dirichlet* trabalhar com distribuições de famílias exponenciais e convexas. O procedimento agora estima os parâmetros das distribuições propostas como distribuições *a priori* e de seguida utilizamos os parâmetros para a nova distribuição proposta em cada componente de mistura para usarmos no cálculo do modelo de mistura por processos de Dirichlet, definido por:

$$\begin{aligned}
 p_{\text{Powernorm}} &\sim \text{Powernorm}(c) \\
 \eta_{\text{Powernorm}} &\sim \text{Powernorm}(c_t) \\
 &\vdots \\
 p_{\text{Johnsons}} &\sim \text{Johnson } Su(a, b) \\
 \eta_{\text{Johnsons}} &\sim \text{Johnson } Su(a_t, b_t)
 \end{aligned}
 \tag{73}$$

A figura 34 demonstra as probabilidades preditivas calculadas através dos modelos de mistura modificados para cada distribuição proposta. A linha azul corresponde à probabilidade preditiva das intensidades para substância cinzenta segundo modelos de mistura Gaussianos. E a linha verde corresponde à probabilidade preditiva das intensidades para substância branca segundo os modelos de mistura para cada distribuição proposta para substituir a distribuição Gaussiana.

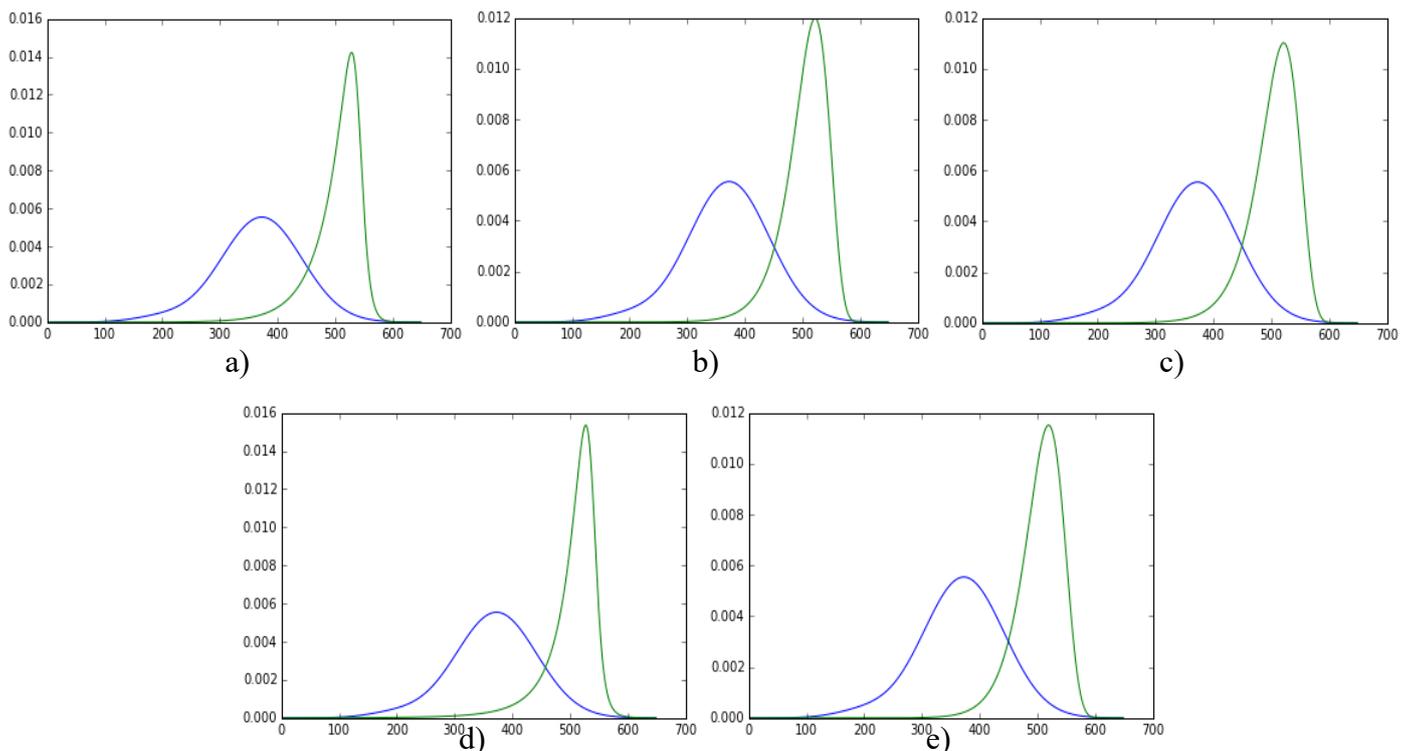


Figura 34 - Comparação entre as probabilidades preditivas dos diferentes modelos. a) Genlogistic, b) Gumbel enviesada à esquerda, c) Log Gama, d) Johnson Su, e) Powernorm

Observando a figura 34 é possível verificar que as diferenças entre as probabilidades preditivas dos modelos são difíceis de se observar. A única diferença visível é que os máximos para cada distribuição são diferentes, o que significa que quando o modelo estava a calcular os parâmetros que definem a probabilidade preditiva com diferentes distribuições deram diferentes pesos para as diferentes intensidades.

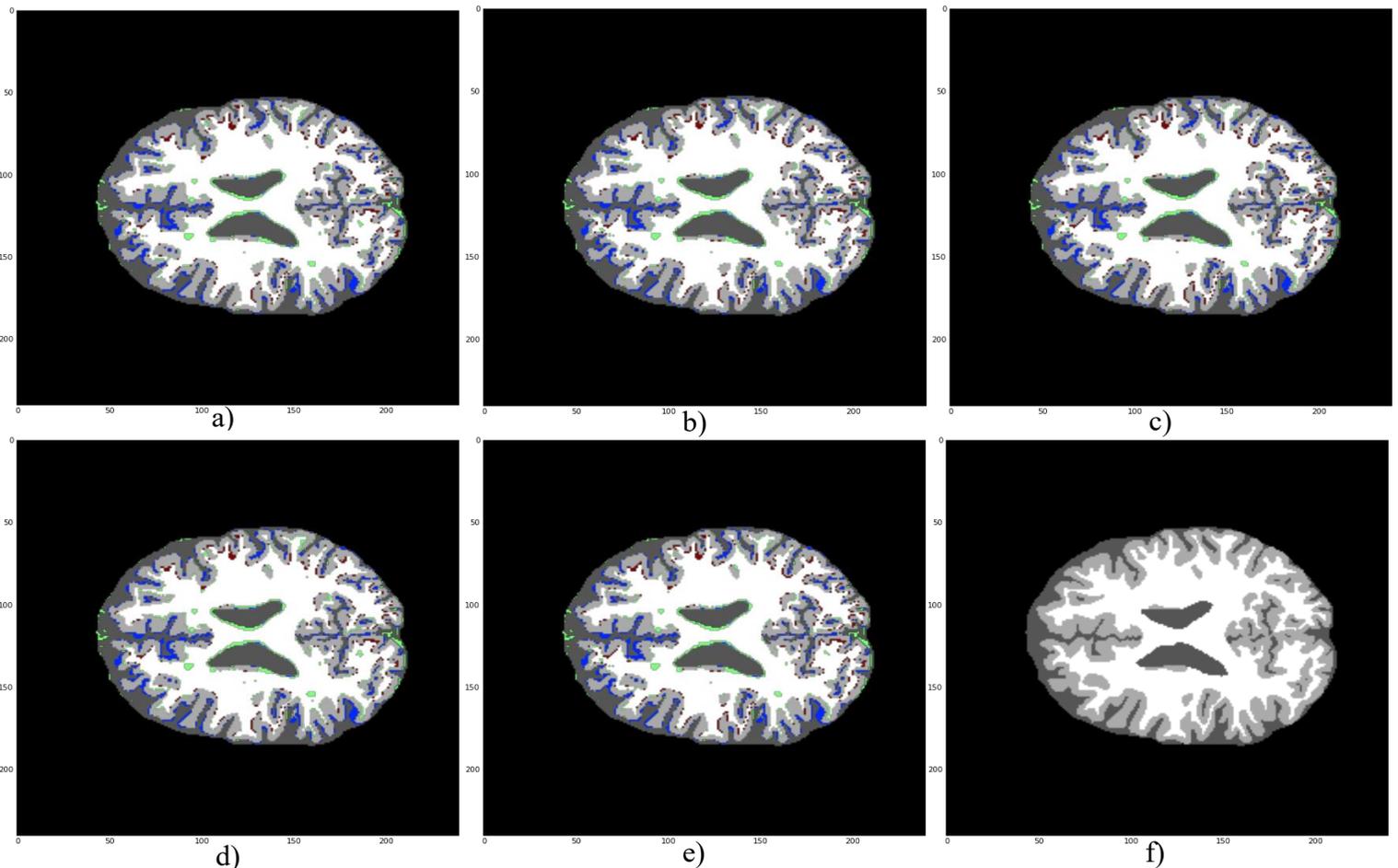


Figura 35 - Segmentação por diferentes modelos. a) Genlogistic, b) Gumbel enviesado à esquerda, c) Log Gama, d) Johnson Su, e) PowerNorm, f) Marcação pelo médico

Devido a esse efeito dos modelos alterados na probabilidade preditiva conduz também a que os mapas de probabilidades e a segmentação tenham poucas diferenças entre as distribuições utilizadas. A figura 35 representa as diferentes segmentações de tecidos cerebrais das combinações de modelos de mistura por processos de *Dirichlet* com marcações das zonas mal classificadas dos tipos de tecidos, onde a cor azul corresponde ao tecido líquido cerebrospinal, a cor verde corresponde ao tecido substância cinzenta e a cor vermelha corresponde ao tecido substância branca. A partir da figura 35, verifica-se que as diferenças na classificação de tecidos entre modelos é difícil de se detetar só por observação. Devido a esse problema, a análise da eficiência e o contributo das diferentes distribuições que introduzem aos modelos de mistura é feito através de métricas de avaliação que serão apresentadas nas próximas secções.

## 5.5 Resultados

Nesta secção são apresentados os resultados obtidos das métricas de avaliação e da matriz de confusão das diferentes segmentações dos tecidos cerebrais obtidos por modelos de mistura por processos de *Dirichlet*.

### 5.5.1 Métricas de Avaliação

A avaliação das segmentações dos tecidos é feita através do cálculo das seguintes métricas: *Dice*, *Jaccard*, TPR ou Sensibilidade, PPV e Especificidade.

A métrica *Dice* mede a sobreposição entre a mesma classe segmentada pelo algoritmo implementado e pela marcação manual. Considerando que *S* representa a segmentação obtida, *R* a marcação manual, *VP* (verdadeiros positivos) o número de pixéis corretamente classificados, *FP* (falsos positivos) a quantidade de pixéis detetados como pertencentes à classe, mas que na verdade não são, e *FN* (falsos negativos) o número daqueles que não foram incluídos na classe pela segmentação, mas que deveriam ter sido, então pode-se calcular o coeficiente *Dice* utilizando a seguinte equação: [49]

$$Dice = \frac{2|S \cap R|}{|S| + |R|} = \frac{2VP}{FN + 2VP + FP} \quad (74)$$

Esta métrica permite concluir acerca dos *VP*, *FN* e *FP*. Além disso, é uma das métricas mais populares na segmentação de ressonância magnética do cérebro, em conjunto com o *Jaccard*, podendo ser convertidas uma na outra. A métrica *Jaccard* pode ser calculada através da seguinte equação [49]:

$$Jaccard = \frac{|S \cap R|}{|S \cup R|} = \frac{VP}{FN + VP + FP} \quad (75)$$

A métrica *True Positive Rate* (TPR) representa a razão dos casos positivos que foram classificados corretamente, ou seja, considerados verdadeiros, sendo também denominada sensibilidade. Esta pode ser calculada segundo a seguinte equação:

$$TPR = Sensibilidade = \frac{VP}{VP + FN} \quad (76)$$

A métrica *Positive Predictive Value* (PPV), sendo também denominado por precisão, representa a razão dos casos preditivos positivos que foram classificados corretamente, ou seja, considerados verdadeiros e calcula-se pela seguinte equação:

$$PPV = \frac{VP}{VP + FP} \quad (77)$$

A métrica Especificidade significa a razão dos casos negativos que foram classificados corretamente, ou seja, considerados verdadeiros (verdadeiros negativos - *VN*) e calcula-se pela seguinte equação:

$$Especificidade = \frac{VN}{VN + FP} \quad (78)$$

### 5.5.2 Matriz de Confusão

A matriz de confusão é uma matriz que descreve o desempenho de um modelo de classificação ou classificador para um conjunto de dados de teste em que os valores verdadeiros são conhecidos. A matriz mostra o número de previsões ou casos preditivos corretos ou verdadeiros e incorretos ou falsos feitos pelo modelo de classificação em comparação com os resultados conhecidos dos dados. A matriz tem o formato  $N \times N$ , onde  $N$  é o número de classes. O desempenho do modelo é avaliado utilizando os dados da matriz. As linhas da matriz correspondem aos casos verdadeiros e as colunas correspondem aos casos preditivos [55]. Para o nosso caso a tabela é  $4 \times 4$  porque o resultado da segmentação devolve-nos quatro valores: 0 que corresponde aos pontos que não interessam para a segmentação, 1 que corresponde ao tecido líquido cerebrospinal, 2 que corresponde ao tecido substância cinzenta e 3 que corresponde ao tecido substância branca. A tabela 2 apresenta a matriz de confusão de quatro classes.

	<b>Previsão 0</b>	<b>Previsão 1</b>	<b>Previsão 2</b>	<b>Previsão 3</b>
<b>Verdadeiro 0</b>	$a_{ij}$			
<b>Verdadeiro 1</b>				
<b>Verdadeiro 2</b>				
<b>Verdadeiro 3</b>				

Tabela 2 – Exemplo duma matriz de confusão com quatro classes

Na diagonal principal da matriz,  $a_{ii}$ ,  $i = 1, \dots, 4$ , encontram-se os valores verdadeiros positivos (VP) para cada uma das classes do modelo de classificação. Selecionando um dos VP os valores situados na sua linha,  $a_{j}$ ,  $j = 1, \dots, 4$ , são os valores falsos positivos (FP) e os valores situados na sua coluna,  $a_{i}$ ,  $i = 1, \dots, 4$ , são os valores falsos negativos (FN) e os restantes valores que não se encontram nem na linha nem na coluna do VP selecionado são os valores verdadeiros negativos (VN) [56].

### 5.5.3 Método de Avaliação

São usadas as cinco sequências de imagens dos sujeitos de treino fornecidos pela base de dados *MRBrainS* para avaliar a segmentação através de modelos de mistura por processos de *Dirichlet*. O procedimento de avaliação é realizada por avaliação cruzada, ou seja utilizam-se as intensidades dos 4 sujeitos para treinar um modelo para cada tipo de tecido cerebral e depois usa-se o sujeito que ficou de fora do treino para testar a segmentação. Fazemos este procedimento para cada sujeito, no que resulta em cinco testes, um para cada sujeito.

### 5.5.4 Avaliação do método *Atropos*

No trabalho de *Sérgio Pereira* [49] é apresentado o método *Atropos* para extração das probabilidades *a posteriori*, que vêm de uma segmentação prévia da imagem segundo o mesmo método e que de seguida as probabilidades são usadas como característica para o seu sistema de segmentação [49].

O método *Atropos* é fornecido em pacotes de *software* de segmentação de imagens através de estatística, que utilizam misturas Gaussianas e *Markov Random Field* para segmentar os tecidos cerebrais. O método aplica a estatística Bayesiana, ou seja, que depende de um modelo de verosimilhança e de uma probabilidade *a priori*, modelizada por uma estrutura *Markov Random Field* e que usa uma mistura de Gaussianas, isto é, uma distribuição Gaussiana para cada classe [49]. Foi realizado uma avaliação ao método *Atropos* com os cinco sujeitos da base de dados *MRBrainS*, só com as sequências de imagens *T1-weighted* 3D com diferentes parâmetros para o *Markov Random Field* para se otimizar o parâmetro de suavização do *Markov Random Field* e para estabelecer uma comparação com o seu método proposto por *Sérgio Pereira* [49]. Uma vez encontrado o parâmetro ótimo de suavização foram obtidos os seguintes resultados da segmentação pelo método *Atropos*: [49]

	Sujeito 1	Sujeito 2	Sujeito 3	Sujeito 4	Sujeito 5	Média ± Desvio Padrão
<b>Líquido Cerebrospinal</b>						
<b>Jaccard</b>	0,6793	0,6861	0,6602	0,6749	0,6584	0,6718 ± 0,0108
<b>Dice</b>	0,8090	0,8138	0,7953	0,8059	0,7940	0,8036 ± 0,0077
<b>PPV</b>	0,8949	0,8012	0,8016	0,8749	0,7606	0,8267 ± 0,0503
<b>Sensibilidade</b>	0,7381	0,8268	0,7892	0,7470	0,8305	0,7863 ± 0,0386
<b>Especificidade</b>	0,9525	0,9118	0,9090	0,9445	0,9044	0,9244 ± 0,0199
<b>Substância Cinzenta</b>						
<b>Jaccard</b>	0,6231	0,5800	0,6162	0,6216	0,6746	0,6231 ± 0,0302
<b>Dice</b>	0,7678	0,7342	0,7625	0,7667	0,8057	0,7674 ± 0,0228
<b>PPV</b>	0,7821	0,8122	0,7610	0,8104	0,8799	0,8091 ± 0,0402
<b>Sensibilidade</b>	0,7541	0,6699	0,7641	0,7274	0,7430	0,7317 ± 0,0332
<b>Especificidade</b>	0,8720	0,8790	0,8618	0,8789	0,9169	0,8817 ± 0,0187
<b>Substância Branca</b>						
<b>Jaccard</b>	0,7437	0,6358	0,7685	0,7033	0,7721	0,7247 ± 0,0508
<b>Dice</b>	0,8530	0,7774	0,8691	0,8258	0,8714	0,8393 ± 0,0350
<b>PPV</b>	0,7735	0,6891	0,8223	0,7199	0,7924	0,7594 ± 0,0485
<b>Sensibilidade</b>	0,9508	0,8916	0,9215	0,9683	0,9679	0,9400 ± 0,0296
<b>Especificidade</b>	0,8902	0,8554	0,9102	0,8714	0,8991	0,8853 ± 0,0196

Tabela 3 - Resultados das métricas de avaliação para a segmentação pelo método *Atropos*

Chegou-se à conclusão que na otimização do parâmetro de suavização do *Markov Random Field* há um compromisso na segmentação de detalhes na imagem de ressonância magnética. Para diferentes valores do parâmetro de suavização, na presença de lesões na substância branca, o método classifica essas zonas como tecido substância cinzenta e que à medida que se aumenta o parâmetro de suavização as zonas de substância branca estreitas

são substituídas por substância cinzenta e com isso remove detalhes mais finos [49]. A figura 36 apresenta os defeitos descritos pela segmentação de tecidos cerebrais pelo método *Atropos*.

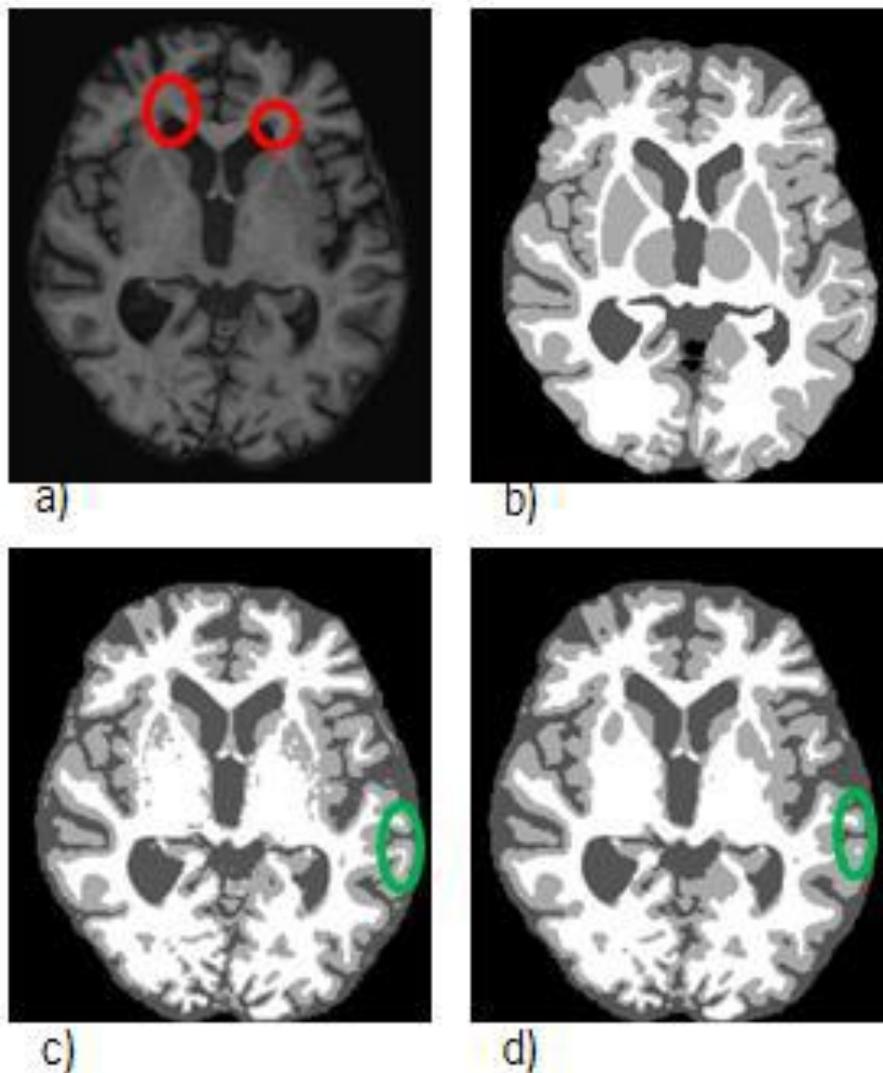


Figura 36 – Segmentação pelo método *Atropos*: a) Sequência T1, marcada com círculos vermelhos em locais de lesão da substância cinzenta, b) Marcação do médico, c) e d) Segmentação da mesma imagem, mas com o parâmetro de suavização de 0,05 e 0,8, respectivamente.[49]

### 5.5.5 Avaliação da segmentação dos tecidos cerebrais por modelos de mistura Gaussianas por processos de *Dirichlet*

Avaliou-se a segmentação de tecidos cerebrais por modelos de mistura Gaussianas por processos de *Dirichlet* através das métricas de avaliação. A tabela 4 e a figura 37 apresentam os resultados obtidos das métricas de avaliação e exemplos da segmentação de tecidos cerebrais por modelos de mistura Gaussianas por processos de *Dirichlet*, respectivamente.

	Sujeito 1	Sujeito 2	Sujeito 3	Sujeito 4	Sujeito 5	Média ± Desvio Padrão
<b>Líquido Cerebrospinal</b>						
<b>Jaccard</b>	0,8077	0,7600	0,7728	0,8079	0,7271	0,7751 ± 0,0306
<b>Dice</b>	0,8936	0,8636	0,8718	0,8938	0,8420	0,8730 ± 0,0195
<b>PPV</b>	0,9289	0,8251	0,8797	0,9188	0,7761	0,8657 ± 0,0577
<b>Sensibilidade</b>	0,8609	0,9059	0,8640	0,8701	0,9202	0,8842 ± 0,0241
<b>Especificidade</b>	0,9630	0,9153	0,9420	0,9588	0,8997	0,9358 ± 0,0246
<b>Substância Cinzenta</b>						
<b>Jaccard</b>	0,6826	0,6300	0,6595	0,6935	0,7014	0,6734 ± 0,0259
<b>Dice</b>	0,8114	0,7730	0,7948	0,8190	0,8245	0,8046 ± 0,0187
<b>PPV</b>	0,7972	0,8285	0,7740	0,8427	0,8884	0,8262 ± 0,0393
<b>Sensibilidade</b>	0,8261	0,7245	0,8168	0,7966	0,7692	0,7866 ± 0,0367
<b>Especificidade</b>	0,8778	0,8848	0,8634	0,8988	0,9220	0,8894 ± 0,0199
<b>Substância Branca</b>						
<b>Jaccard</b>	0,7775	0,6690	0,7779	0,7625	0,8241	0,7622 ± 0,0510
<b>Dice</b>	0,8748	0,8016	0,8751	0,8653	0,9035	0,8641 ± 0,0337
<b>PPV</b>	0,8550	0,7669	0,8965	0,8042	0,8894	0,8424 ± 0,0499
<b>Sensibilidade</b>	0,8956	0,8397	0,8547	0,9364	0,9182	0,8889 ± 0,0367
<b>Especificidade</b>	0,9430	0,9107	0,9564	0,9253	0,9558	0,9382 ± 0,0178

Tabela 4 - Resultados das métricas de avaliação para a segmentação por modelos de mistura Gaussianas por processos de Dirichlet

Com base nos resultados apresentados na Tabela 3 e Tabela 4, podemos comparar as métricas de avaliação da segmentação dos tecidos cerebrais por modelos de mistura Gaussianas por processos de *Dirichlet* e da segmentação de tecidos cerebrais pelo método *Atropos*. Observando os resultados, verifica-se que para todas as métricas de avaliação da segmentação por modelos de mistura Gaussianas por processo de *Dirichlet* são superiores às métricas de avaliação da segmentação com a exceção da métrica Sensibilidade para o caso do tecido substância branca que é superior para a segmentação pelo método *Atropos* do que a pelo método por modelos de mistura Gaussianas por processos de *Dirichlet*, (0,9400 para o método *Atropos* e 0,8889 para o método dos modelos de mistura), o que indica que o método *Atropos* consegue classificar melhor, em média, os pontos considerados substância branca embora que em geral o método por modelos de mistura Gaussianas por processos de *Dirichlet* tem uma métrica *Dice* superior (0,8641 para o método por modelos de mistura e 0,8393 para o método *Atropos*). A diferença da métrica *Dice* entre os dois para o caso do líquido cerebrospinal é de 0.0692, ou seja, houve uma melhoria na classificação de 6.92%. Para a substância cinzenta, tem uma diferença de 0.0372 no que traduz numa melhoria na classificação da substância cinzenta de 3.72%. E para a substância branca tem uma diferença de 0.0248 que representa uma melhoria de 2.48% para a classificação da substância branca. Com isto podemos concluir que a segmentação por modelos de mistura Gaussianas por processos de *Dirichlet* consegue classificar melhor que as abordagens que usam a metodologia paramétrica que é o caso do método *Atropos*. Uma vez feita a análise das tabelas, fomos comprovar a análise dos resultados com um teste estatístico. O teste escolhido foi um teste *Wilcoxon* que testa a

diferença das médias de duas amostras relacionadas para determinar se a diferença das médias é igual a zero. O teste comparou as médias de cada métrica da segmentação por modelos de mistura Gaussianas por processos de *Dirichlet* com as correspondentes do método *Atropos* para cada tipo de tecido cerebral. A hipótese nula do teste é se a diferença das médias das amostras é igual a zero, ou seja, se a diferença é nula. O resultado obtido do teste *Wilcoxon* foi de  $p - \text{valor} = 0,043$ , para os casos das médias das métricas do tecido líquido cerebrospinal e do tecido substância cinzenta, o que indica que a diferença é estatisticamente significativa entre os métodos de segmentação, porque o p-valor é menor que 0.05, e com isso a segmentação por modelos de mistura Gaussianas por processos de *Dirichlet* classifica melhor que o método *Atropos* para estes dois tipos de tecido. E para o caso do tecido substância branca, o resultado foi de  $p - \text{valor} = 0,225$ , o que indica que a diferença não é estatisticamente significativa entre os métodos de segmentação, porque o p-valor é maior que 0.05, revelando que a segmentação por modelos de mistura Gaussianas por processos de *Dirichlet* não consegue classificar melhor o tecido substância branca que o método *Atropos*. Mas mesmo com estas melhorias, a segmentação por processos de *Dirichlet* também tem as suas limitações.

Existem zonas na imagem de ressonância magnética em que as intensidades de diferentes tecidos têm intensidades muito semelhantes e leva a que os modelos não sejam capazes de diferenciar corretamente os tecidos e classificam incorretamente. A figura 37 apresenta diversos defeitos da segmentação usando os processos de *Dirichlet*. Como mostra a figura 37 c) e f) existem zonas de pontos isolados de substância cinzenta em sítios classificados como substância branca e verificando as suas imagens T1, a) e d), observa-se que nessa zonas a diferença de intensidades entre tecidos é muito pequena. Mas também ao comparar a imagem g), h) e i) verifica-se que a segmentação classifica tecidos onde a marcação do médico não classifica, o que leva a uma classificação considerada errada mas podem ter sido detetadas lesões no cérebro.

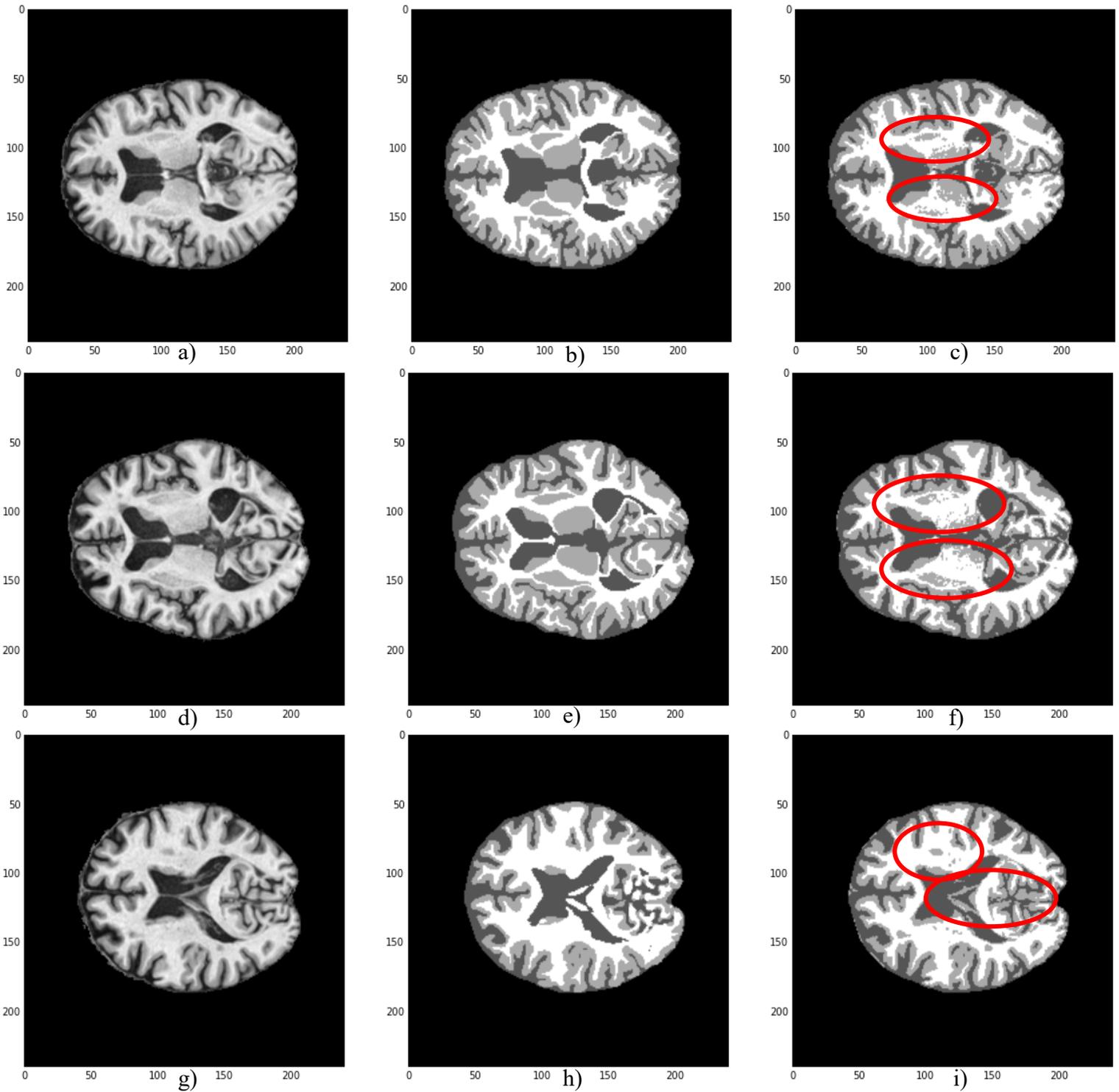


Figura 37 - Exemplos dos defeitos da segmentação por modelos de mistura Gaussianas com a comparação das imagens T1 e a marcação do médico. a),d) e g) Imagens T1; b),e) e h) Marcação do médico; c),f) e i) Segmentação por modelos de mistura Gaussianas com marcações das zonas com defeitos

Com base nesses defeitos foi proposto uma alteração das distribuições usadas no modelo para se verificar se ocorreriam melhorias na segmentação de tecidos por modelos de mistura por processos de *Dirichlet*. Os resultados dessas modificações são apresentados nas secções seguintes.

### 5.5.6 Avaliação da segmentação dos tecidos cerebrais por combinação de diferentes modelos de mistura por processos de *Dirichlet* para o caso do tecido líquido cerebrospinal

Nesta secção é feita a análise da segmentação dos tecidos cerebrais por combinação do modelo de mistura Ricianas por processos de *Dirichlet* para o tecido líquido cerebrospinal com os modelos de mistura Gaussianas por processos de *Dirichlet* para os tecidos substância cinzenta e branca, que se irá denominar por abordagem Riciana, pelos resultados obtidos das métricas de avaliação e da matriz de confusão para cada tipo de tecido cerebral da segmentação. Na Tabela 5 são apresentados os resultados das métricas de avaliação.

	Sujeito 1	Sujeito 2	Sujeito 3	Sujeito 4	Sujeito 5	Média±Desvio Padrão
<b>Líquido Cerebrospinal</b>						
<b>Jaccard</b>	0,8051	0,7606	0,7714	0,8079	0,7307	0,7752±0,0289
<b>Dice</b>	0,8920	0,8641	0,8710	0,8938	0,8444	0,8730±0,0184
<b>PPV</b>	0,9322	0,8298	0,8832	0,9188	0,7826	0,8693±0,0560
<b>Sensibilidade</b>	0,8552	0,9012	0,8591	0,8701	0,9168	0,8805±0,0243
<b>Especificidade</b>	0,9651	0,9184	0,9442	0,9589	0,9038	0,9381±0,0235
<b>Substância Cinzenta</b>						
<b>Jaccard</b>	0,6826	0,6321	0,6603	0,6935	0,7061	0,6749±0,0261
<b>Dice</b>	0,8114	0,7746	0,7954	0,8190	0,8277	0,8056±0,0188
<b>PPV</b>	0,7939	0,8261	0,7716	0,8427	0,8871	0,8243±0,0400
<b>Sensibilidade</b>	0,8296	0,7291	0,8208	0,7966	0,7758	0,7904±0,0360
<b>Especificidade</b>	0,8748	0,8821	0,8608	0,8988	0,9204	0,8874±0,0205
<b>Substância Branca</b>						
<b>Jaccard</b>	0,7772	0,6688	0,7779	0,7625	0,8240	0,7621±0,0510
<b>Dice</b>	0,8746	0,8015	0,8751	0,8653	0,9035	0,8640±0,0338
<b>PPV</b>	0,8545	0,7667	0,8963	0,8041	0,8893	0,8422±0,0499
<b>Sensibilidade</b>	0,8957	0,8397	0,8547	0,9365	0,9182	0,8890±0,0368
<b>Especificidade</b>	0,9428	0,9106	0,9563	0,9253	0,9557	0,9381±0,0178

Tabela 5 – Resultados das métricas de avaliação para a segmentação pela abordagem Riciana

Comparando as Tabelas 4 e 5, verifica-se que os resultados obtidos são bastante semelhantes. Mas através da diferença entre as métricas de avaliação dos diferentes tecidos cerebrais dos dois métodos poderemos analisar se houve melhorias por parte da abordagem Riciana em relação à segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*. A Tabela 6 apresenta as diferenças das médias das métricas de avaliação entre a distribuição Gaussiana com a distribuição Riciana.

	<b>Líquido Cerebrospinal</b>	<b>Substância Cinzenta</b>	<b>Substância Branca</b>
<b>Jaccard</b>	-0,0001	-0,0015	0,0001
<b>Dice</b>	0,0000	-0,0010	0,0001
<b>PPV</b>	-0,0036	0,0019	0,0002
<b>Sensibilidade</b>	0,0037	-0,0038	-0,0001
<b>Especificidade</b>	-0,0023	0,0020	0,0001

Tabela 6 – Diferença das médias das métricas de avaliação entre as segmentações por modelos de mistura Gaussianas e da abordagem Riciana

Segundo a Tabela 6, as diferenças das médias das métricas para o tecido líquido cerebrospinal mostra que a segmentação, em geral, dos dois métodos são idênticas porque o valor de *Dice* é 0 e *Jaccard* tem uma diferença de 0,0001 a favor da abordagem Riciana. E a abordagem Riciana tem um PPV e Especificidade melhor que a segmentação por modelos de mistura Gaussianas embora pior Sensibilidade, o que indica que a abordagem Riciana tem uma melhor precisão a classificar as intensidades que pertencem à classe e consegue identificar melhor as intensidades que não pertencem à classe mas perde na classificação de intensidades que realmente pertencem à classe mas são classificados como outro tecido. Para o caso da substância cinzenta, a avaliação da segmentação, no geral, mostra que a abordagem Riciana classifica melhor que a segmentação por modelos de mistura Gaussianas porque o valor da diferença da métrica *Dice* e *Jaccard* é 0,0010 e 0,0015, respetivamente a favor da abordagem Riciana. Mas a segmentação por modelos de mistura Gaussianas tem um PPV e Especificidade melhor que a abordagem Riciana embora pior Sensibilidade, o que significa que a abordagem Riciana tem uma pior precisão e classifica mais intensidades que não pertencem ao tipo de tecido mas melhora a classificação de intensidades que realmente pertencem à classe mas são classificados como outro tecido. Em relação à substância branca, os resultados de todas as diferenças das médias das métricas revelam que a diferença é mínima, chega a um valor máximo de 0,0002 no que podemos dizer que as segmentações são quase idênticas.

Realizado a análise das tabelas, fomos comprovar a análise dos resultados das métricas dos dois métodos com um teste estatístico. O teste escolhido foi um teste *Wilcoxon* que comparou as médias de cada métrica da segmentação por modelos de mistura Gaussianas por processos de *Dirichlet* com as correspondentes da abordagem Riciana para cada tipo de tecido cerebral. A hipótese nula do teste é se a diferença das médias das amostras é igual a zero, ou seja, se a diferença é nula. O resultado obtido do teste *Wilcoxon* foi de  $p - valor = 0,715$ , para o caso das médias das métricas do tecido líquido cerebrospinal, para o caso do tecido substância cinzenta, foi de  $p - valor = 0,893$  e para o caso do tecido substância branca, foi de  $p - valor = 0,157$ , o que indica que a diferença para esta amostra não é estatisticamente significativa entre os métodos de segmentação para cada tipo de tecido, porque o p-valor de cada tecido é maior que 0.05. Com o resultado deste teste não podemos concluir que a segmentação pela abordagem Riciana seja melhor que a segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*. Com isso, fomos comparar as matrizes de classificação de ambas as segmentações dos dois métodos para termos uma análise mais detalhada do desempenho de cada método, representadas

pelas tabelas 7 e 8, onde o valor 0 corresponde aos pontos que não interessam ser classificados, o valor 1 corresponde ao tecido líquido cerebrospinal, o valor 2 corresponde ao tecido substância cinzenta e o valor 3 corresponde ao tecido substância branca.

<b>Gaussiana</b>	<b>Previsão 0</b>	<b>Previsão 1</b>	<b>Previsão 2</b>	<b>Previsão 3</b>
<b>Verdadeiro 0</b>	11519575	0	0	0
<b>Verdadeiro 1</b>	0	656285	85161	1832
<b>Verdadeiro 2</b>	0	99974	729411	102878
<b>Verdadeiro 3</b>	0	3500	67807	557577

*Tabela 7 – Matriz de classificação da segmentação por modelos de mistura Gaussianas por processos de Dirichlet*

<b>Ricana</b>	<b>Previsão 0</b>	<b>Previsão 1</b>	<b>Previsão 2</b>	<b>Previsão 3</b>
<b>Verdadeiro 0</b>	11519575	0	0	0
<b>Verdadeiro 1</b>	0	653444	87868	1966
<b>Verdadeiro 2</b>	0	96329	733049	102885
<b>Verdadeiro 3</b>	0	3361	67927	557596

*Tabela 8 - Matriz de classificação da segmentação pela abordagem Riciana*

Comparando as tabelas 7 e 8, verifica-se que para o caso do tecido líquido cerebrospinal (valor 1), a segmentação pela abordagem Riciana classifica menos pontos verdadeiros como líquido (653444 para a abordagem Riciana e 656285 para a segmentação por modelos de mistura Gaussianas) e deteta mais falsos positivos que deviam ser tecido substância cinzenta e branca que a segmentação por modelos de mistura Gaussianas (87868 que deviam ser substância cinzenta e 1966 substância branca, ambos da abordagem Riciana em relação aos 85161 que deviam ser substância cinzenta e 1832 para a substância branca, ambos da segmentação por modelos de mistura Gaussianas) mas classifica menos pontos que não pertencem à classe mas que deviam pertencer do que a segmentação por modelos de mistura Gaussianas (96329 como substância cinzenta e 3361 como substância branca da abordagem Riciana para 99974 como substância cinzenta e 3500 como substância branca da segmentação por modelos de mistura Gaussianas). O que significa que a abordagem Riciana piora um pouco a segmentação do tecido líquido cerebrospinal, com o compromisso de perder alguns pontos como líquido cerebrospinal mas prevê melhor que a segmentação por modelos de mistura Gaussianas. No caso do tecido substância cinzenta (valor 2), a segmentação pela abordagem Riciana classifica mais pontos corretamente como tecido substância cinzenta (733049 para a abordagem Riciana e 729411 para a segmentação por modelos de mistura Gaussianas) e deteta menos falsos positivos que deviam ser tecido líquido cerebrospinal mas para o tecido substância branca classifica ligeiramente mais que a segmentação por modelos de mistura Gaussianas (96329 que deviam ser líquido cerebrospinal e 102885 substância branca, para a abordagem Riciana em relação aos 99974 que deviam ser líquido cerebrospinal e 102878 substância branca, para a segmentação por modelos de mistura Gaussianas) mas classifica mais pontos que não pertencem à classe e que deviam pertencer, no caso do tecido substância branca é uma pequena diferença, do que a segmentação por modelos de mistura de Gaussianas (87868 como líquido

cerebrospinal e 67927 como substância branca para a abordagem Riciana para 85161 como líquido cerebrospinal e 67807 como substância branca para a segmentação por modelos de mistura Gaussianas). E com isso, a abordagem Riciana melhora um pouco a segmentação do tecido substância cinzenta mas com uma perda de detalhes na segmentação que compensa a melhoria da segmentação comparando com a segmentação por modelos de mistura Gaussianas. E finalmente para o caso do tecido substância branca (valor 3), a segmentação pela abordagem Riciana classifica ligeiramente mais pontos corretamente como substância branca (557596 para a abordagem Riciana e 557577 para a segmentação por modelos de mistura Gaussianas) e deteta menos falsos positivos que deviam ser tecido líquido cerebrospinal mas para os que deviam ser tecido substância cinzenta é ligeiramente superior que a segmentação por modelos de mistura Gaussianas (3361 que deviam ser líquido cerebrospinal e 67927 substância cinzenta, ambos da abordagem Riciana em relação aos 3500 que deviam ser líquido cerebrospinal e 67807 substância cinzenta, ambos da segmentação por modelos de mistura Gaussianas) e identifica ligeiramente mais pontos que não pertencem ao tecido mas que deviam pertencer, quer para o tecido líquido cerebrospinal quer para o tecido substância cinzenta, em relação à segmentação por modelos de mistura de Gaussianas (1966 como líquido cerebrospinal e 102885 como substância cinzenta da abordagem Riciana para 1832 como líquido cerebrospinal e 102878 como substância cinzenta do modelo de mistura Gaussianas). O que significa que a abordagem Riciana melhora ligeiramente a segmentação do tecido substância branca que a segmentação por modelos de mistura Gaussianas, com o compromisso de marcar mais pontos falsos, exceto no tecido líquido cerebrospinal, mas é uma diferença pequena.

Com isto podemos concluir que a utilização da segmentação de tecidos cerebrais por combinação do modelo de mistura Ricianas por processos de *Dirichlet* para o tecido líquido cerebrospinal com os modelos de mistura Gaussianas por processos de *Dirichlet* para os tecidos substância cinzenta e branca tem as suas vantagens e desvantagens, em que as vantagens são ligeiramente superiores às desvantagens visto que consegue classificar melhor o tecido substância cinzenta e branca perdendo alguns detalhes para o tecido líquido cerebrospinal devido à alteração da forma da distribuição para Riciana e por isso esta abordagem contribui uma melhoria para a segmentação em relação à segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*. Mas mesmo com esta melhoria, este novo modelo apresenta os mesmos problemas que foram apresentados para a segmentação por modelos de mistura Gaussianas. A figura 38 compara a segmentação pela abordagem Riciana com a segmentação por modelos de mistura Gaussianas, com ambos os métodos, assinalados com marcações das zonas mal classificadas, onde a cor azul corresponde ao tecido líquido cerebrospinal, a cor verde corresponde ao tecido substância cinzenta e a cor vermelha corresponde ao tecido substância branca, e com as marcações do médico, onde se verifica que as diferenças entre segmentações são mínimas.

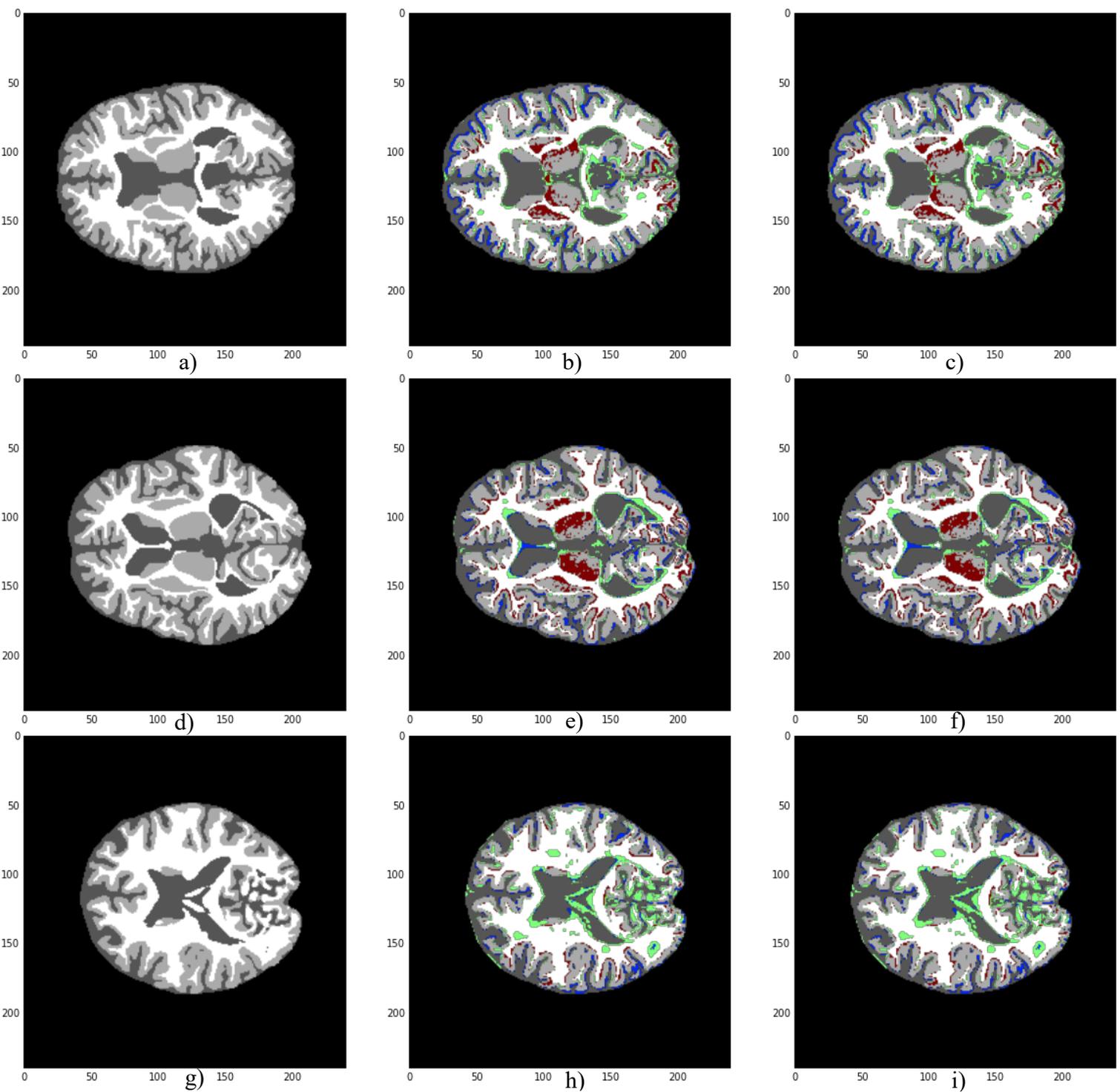


Figura 38 – Comparação da segmentação por modelos de mistura Gaussianas por processos de Dirichlet com a segmentação pela abordagem Riciana, onde ambos têm marcações das zonas mal classificadas, e com a marcação do médico. a), d) e g) Marcação do médico; b), e) e h) Segmentação por modelos de mistura Gaussianas; c), f) e i) Segmentação pela abordagem Riciana

### 5.5.7 Avaliação da segmentação dos tecidos cerebrais por combinação de diferentes modelos de mistura por processos de *Dirichlet* para o caso do tecido substância branca

Nesta secção é feita a análise da segmentação dos tecidos cerebrais por combinação do modelo de mistura Ricianas por processos de *Dirichlet* para o tecido líquido cerebrospinal com o modelo de mistura Gaussianas por processos de *Dirichlet* para o tecido substância cinzenta e para o tecido substância branca foi usado modelo de mistura por processos de *Dirichlet* com as seguintes distribuições: *Genlogistic*, *Gumbel* enviesada à esquerda, *Log Gama*, *Powernorm* e *Johnson Su*, em que para cada segmentação se irá denominar por abordagem consoante a distribuição utilizada para estimar as intensidades do tecido substância branca, ou seja, abordagem *Genlogistic*, abordagem *Gumbel* enviesada à esquerda, abordagem *Log Gama*, abordagem *Powernorm* e abordagem *Johnson Su*. Nesta análise foi usado para as intensidades do líquido cerebrospinal o modelo de mistura Ricianas por processos de *Dirichlet* porque houve uma pequena melhoria na segmentação dos tecidos cerebrais. É feito a comparação dos resultados obtidos nesta secção com os resultados obtidos na segmentação por modelos de mistura Gaussianas por processos de *Dirichlet* para verificar se a alteração das distribuições nos dados das intensidades do tecido substância branca melhoram a segmentação. Nas Tabelas 9, 12, 15, 18, 21 são apresentados os resultados das métricas de avaliação para abordagem.

#### 5.5.7.1 Abordagem *Genlogistic*

Comparando as Tabelas 4 e 9, tal como acontece na secção anterior, verifica-se que os resultados obtidos são bastante parecidos. Mas através da diferença entre as médias das métricas de avaliação dos diferentes tecidos cerebrais dos dois métodos pode-se analisar se houve melhorias por parte da abordagem *Genlogistic* em relação à segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*.

	Sujeito 1	Sujeito 2	Sujeito 3	Sujeito 4	Sujeito 5	Média±Desvio Padrão
<b>Líquido Cerebrospinal</b>						
<b>Jaccard</b>	0,8053	0,7608	0,7714	0,8079	0,7307	0,7752±0,0289
<b>Dice</b>	0,8922	0,8641	0,8710	0,8938	0,8444	0,8731±0,0184
<b>PPV</b>	0,9322	0,8299	0,8832	0,9188	0,7826	0,8693±0,0560
<b>Sensibilidade</b>	0,8554	0,9014	0,8591	0,8701	0,9168	0,8805±0,0243
<b>Especificidade</b>	0,9651	0,9184	0,9442	0,9589	0,9038	0,9381±0,0235
<b>Substância Cinzenta</b>						
<b>Jaccard</b>	0,6744	0,6273	0,6576	0,6856	0,7010	0,6692±0,0253
<b>Dice</b>	0,8055	0,7710	0,7935	0,8134	0,8242	0,8015±0,0183
<b>PPV</b>	0,7998	0,8282	0,7769	0,8445	0,8922	0,8283±0,0395
<b>Sensibilidade</b>	0,8113	0,7211	0,8108	0,7846	0,7658	0,7787±0,0335
<b>Especificidade</b>	0,8819	0,8851	0,8666	0,9017	0,9253	0,8921±0,0200
<b>Substância Branca</b>						
<b>Jaccard</b>	0,7748	0,6667	0,7803	0,7558	0,8210	0,7597±0,0511
<b>Dice</b>	0,8731	0,8001	0,8766	0,8609	0,9017	0,8625±0,0339
<b>PPV</b>	0,8374	0,7588	0,8666	0,7918	0,8769	0,8303±0,0490
<b>Sensibilidade</b>	0,9120	0,8461	0,8867	0,9432	0,9279	0,8992±0,0369
<b>Especificidade</b>	0,9335	0,9058	0,9510	0,9188	0,9496	0,9318±0,0175

Tabela 9 - Resultados das métricas de avaliação para a segmentação pela abordagem *Genlogistic*

A Tabela 10 apresenta as diferenças das médias das métricas de avaliação entre a segmentação por modelos de mistura Gaussianas com a abordagem *Genlogistic*.

	<b>Líquido Cerebrospinal</b>	<b>Substância Cinzenta</b>	<b>Substância Branca</b>
<b>Jaccard</b>	-0,0001	0,0042	0,0025
<b>Dice</b>	-0,0001	0,0031	0,0016
<b>PPV</b>	-0,0036	-0,0021	0,0121
<b>Sensibilidade</b>	0,0037	0,0079	-0,0103
<b>Especificidade</b>	-0,0023	-0,0027	0,0064

Tabela 10 - Diferença das médias das métricas de avaliação entre as segmentações por modelos de mistura Gaussianas e da abordagem *Genlogistic*

Segundo a Tabela 10, as diferenças das médias das métricas para o tecido líquido cerebrospinal indica que a segmentação, em geral, dos dois métodos são idênticas porque o valor de *Dice* tem uma diferença de 0,0001 e *Jaccard* igualmente, a favor da abordagem *Genlogistic*. E a abordagem *Genlogistic* tem um PPV e Especificidade melhor que a segmentação por modelos de mistura Gaussianas embora pior Sensibilidade, o que indica que a abordagem *Genlogistic* tem uma melhor precisão a classificar as intensidades que pertencem à classe e consegue identificar melhor as intensidades que não pertencem à classe mas perde na classificação de intensidades que realmente pertencem à classe mas são classificados como outro tecido. Para o caso da substância cinzenta, a avaliação da segmentação, no geral, mostra que a abordagem *Genlogistic* não consegue classificar melhor que a segmentação por modelos de mistura Gaussianas porque o valor da diferença da métrica *Dice* e *Jaccard* é 0,0042 e 0,0031, respectivamente a favor da segmentação por modelos de mistura Gaussianas. Mas a segmentação por modelos de mistura Gaussianas tem uma Sensibilidade melhor que a abordagem *Genlogistic* mas piora na métrica Especificidade e PPV, o que indica que a abordagem *Genlogistic* tem uma melhor precisão e classifica menos intensidades que não pertencem ao tipo de tecido mas piora a classificação de intensidades que realmente pertencem à classe mas são classificados como outro tecido. Em relação à substância branca, os resultados de todas as diferenças das médias das métricas revelam que a segmentação por modelos de mistura Gaussianas é melhor que a abordagem *Genlogistic*, exceto na métrica Sensibilidade. Isto significa que a abordagem *Genlogistic* só é melhor a classificar intensidades que realmente pertencem ao tecido substância branca mas são classificados como outro tecido.

Uma vez feita a análise da tabela, fomos comprovar a análise dos resultados das métricas dos dois métodos com um teste estatístico. O teste escolhido foi um teste *Wilcoxon* que comparou as médias de cada métrica da segmentação por modelos de mistura Gaussianas por processos de *Dirichlet* com as correspondentes da abordagem *Genlogistic* para cada tipo de tecido cerebral. A hipótese nula do teste é se a diferença das médias das amostras é igual a zero, ou seja, se a diferença é nula. O resultado obtido do teste *Wilcoxon* foi de  $p - \text{valor} = 0,498$ , para o caso das médias das métricas do tecido líquido cerebrospinal, para o caso do tecido substância cinzenta, foi de  $p - \text{valor} = 0,225$  e para o caso do tecido substância branca, foi de  $p - \text{valor} = 0,345$ , o que indica que a diferença para esta

amostra não é estatisticamente significativa entre os métodos de segmentação para cada tipo de tecido, porque o p-valor de cada tecido é maior que 0.05. Com o resultado deste teste não podemos concluir que a segmentação pela abordagem *Genlogistic* seja melhor que a segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*.

Com isso, fomos comparar as matrizes de classificação de ambas as segmentações dos dois métodos para termos uma análise mais detalhada do desempenho de cada método. A Tabela 11 representa a matriz de classificação da abordagem *Genlogistic*, onde o valor 0 corresponde aos pontos que não interessam ser classificados, o valor 1 corresponde ao tecido líquido cerebrospinal, o valor 2 corresponde ao tecido substância cinzenta e o valor 3 corresponde ao tecido substância branca.

<b>Genlogistic</b>	<b>Previsão 0</b>	<b>Previsão 1</b>	<b>Previsão 2</b>	<b>Previsão 3</b>
<b>Verdadeiro 0</b>	11519575	0	0	0
<b>Verdadeiro 1</b>	0	653507	87766	2005
<b>Verdadeiro 2</b>	0	96333	722432	113498
<b>Verdadeiro 3</b>	0	3366	61424	564094

Tabela 11 - Matriz de classificação da segmentação pela abordagem *Genlogistic*

Comparando as tabelas 7 e 11, verifica-se que para o caso do tecido líquido cerebrospinal (valor 1), a segmentação pela abordagem *Genlogistic* classifica menos pontos verdadeiros como esse tecido (653507 para a abordagem *Genlogistic* e 656285 para a segmentação por modelos de mistura Gaussianas) e deteta mais falsos positivos que deviam ser tecido substância cinzenta e branca que a segmentação por modelos de mistura Gaussianas (87766 que deviam ser substância cinzenta e 2005 substância branca, ambos da abordagem *Genlogistic* em relação aos 85161 que deviam ser substância cinzenta e 1832 substância branca, ambos da segmentação por modelos de mistura Gaussianas) mas classifica menos pontos que não pertencem à classe mas que deviam pertencer do que a segmentação por modelos de mistura Gaussianas (96333 como substância cinzenta e 3366 como substância branca da abordagem *Genlogistic* para 99974 como substância cinzenta e 3500 como substância branca da segmentação por modelos de mistura Gaussianas). O que significa que a abordagem *Genlogistic* piora a segmentação do tecido líquido cerebrospinal, com o compromisso de perder alguns pontos como líquido cerebrospinal mas prevê melhor que a segmentação por modelos de mistura Gaussianas. No caso do tecido substância cinzenta (valor 2), a segmentação pela abordagem *Genlogistic* classifica menos pontos corretamente como esse tecido (722432 para a abordagem *Genlogistic* e 729411 para a segmentação por modelos de mistura Gaussianas) e deteta menos falsos positivos que deviam ser tecido líquido cerebrospinal mas para o tecido substância branca classifica mais que a segmentação por modelos de mistura Gaussianas (96333 que deviam ser líquido cerebrospinal e 113498 substância branca, ambos da abordagem *Genlogistic* em relação aos 99974 que deviam ser líquido cerebrospinal e 102878 substância branca, ambos da segmentação por modelos de mistura Gaussianas) e também classifica mais pontos que não pertencem à classe mas que deviam e que são classificados como líquido cerebrospinal e

menos pontos que não pertencem à classe mas que deviam e são marcados como substância branca do que a segmentação por modelos de mistura de Gaussianas (87766 como líquido cerebrospinal e 61424 como substância branca para a abordagem *Genlogistic* para 85161 como líquido cerebrospinal e 67807 como substância branca para a segmentação por modelos de mistura Gaussianas). E com isso, a abordagem *Genlogistic* piora a segmentação do tecido substância cinzenta com a desvantagem de perder de detalhes na segmentação. E finalmente para o caso do tecido substância branca (valor 3), a segmentação pela abordagem *Genlogistic* classifica mais pontos corretamente como esse tecido (564094 para a abordagem *Genlogistic* e 557577 para a segmentação por modelos de mistura Gaussianas) e deteta menos falsos positivos que deviam ser tecido líquido cerebrospinal e substância cinzenta que a segmentação por modelos de mistura Gaussianas (3366 que deviam ser líquido cerebrospinal e 61424 substância cinzenta, ambos da abordagem *Genlogistic* em relação aos 3500 para o que deviam ser líquido cerebrospinal e 67807 substância cinzenta, ambos da segmentação por modelos de mistura Gaussianas) e classifica mais pontos que não pertencem ao tecido mas que deviam pertencer, quer para o tecido líquido cerebrospinal quer para o tecido substância cinzenta, em relação à segmentação por modelos de mistura de Gaussianas (2005 como líquido cerebrospinal e 113498 como substância cinzenta da abordagem *Genlogistic* para 1832 como líquido cerebrospinal e 102878 como substância cinzenta da segmentação por modelos de mistura Gaussianas). Logo, a abordagem *Genlogistic* introduz uma pequena melhoria à segmentação do tecido substância branca que a segmentação por modelos de mistura Gaussianas, por classificar menos pontos que realmente não pertencem ao tecido substância branca e classificar mais pontos corretos mas com exclusão de algum tecido.

Por isso, chegamos à conclusão que a utilização da segmentação de tecidos cerebrais por combinação do modelo de mistura Ricianas por processos de *Dirichlet* para o tecido líquido cerebrospinal com o modelo de mistura Gaussianas por processos de *Dirichlet* para o tecido substância cinzenta e com o modelo de mistura *Genlogistic* por processos de *Dirichlet* para o tecido substância branca interfere negativamente no caso do tecido substância branca na classificação dos tecidos líquido cerebrospinal e substância cinzenta embora melhorasse, com uma perda de definição, a classificação do tecido substância branca, devido à alteração da distribuição para *Genlogistic* para o caso do tecido substância branca, levando a não utilizarmos esta distribuição para o tecido substância branca.

### 5.5.7.2 Abordagem *Gumbel* enviesada à esquerda

Comparando as Tabelas 4 e 12, observa-se que os resultados obtidos são bastante similares. Mas com a diferença entre as médias das métricas de avaliação dos diferentes tecidos cerebrais dos dois métodos é possível analisar se houve melhorias por parte da abordagem *Gumbel* enviesada à esquerda em relação à segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*.

	Sujeito 1	Sujeito 2	Sujeito 3	Sujeito 4	Sujeito 5	Média±Desvio Padrão
<b>Líquido Cerebrospinal</b>						
<b>Jaccard</b>	0,8053	0,7608	0,7714	0,8079	0,7307	0,7753±0,0289
<b>Dice</b>	0,8922	0,8642	0,8710	0,8938	0,8444	0,8731±0,0184
<b>PPV</b>	0,9320	0,8299	0,8832	0,9187	0,7826	0,8693±0,0559
<b>Sensibilidade</b>	0,8556	0,9015	0,8591	0,8701	0,9168	0,8806±0,0243
<b>Especificidade</b>	0,9650	0,9184	0,9442	0,9588	0,9038	0,9381±0,0235
<b>Substância Cinzenta</b>						
<b>Jaccard</b>	0,6710	0,6217	0,6557	0,6762	0,6977	0,6645±0,0253
<b>Dice</b>	0,8031	0,7667	0,7920	0,8069	0,8219	0,7981±0,0184
<b>PPV</b>	0,8014	0,8300	0,7793	0,8458	0,8942	0,8302±0,0394
<b>Sensibilidade</b>	0,8049	0,7124	0,8052	0,7713	0,7604	0,7708±0,0343
<b>Especificidade</b>	0,8840	0,8879	0,8693	0,9044	0,9274	0,8946±0,0198
<b>Substância Branca</b>						
<b>Jaccard</b>	0,7729	0,6638	0,7805	0,7476	0,8183	0,7566±0,0516
<b>Dice</b>	0,8719	0,7980	0,8767	0,8556	0,9001	0,8604±0,0343
<b>PPV</b>	0,8315	0,7503	0,8812	0,7786	0,8702	0,8224±0,0509
<b>Sensibilidade</b>	0,9164	0,8521	0,8723	0,9495	0,9320	0,9044±0,0366
<b>Especificidade</b>	0,9303	0,9007	0,9481	0,9116	0,9462	0,9274±0,0187

Tabela 12 - Resultados das métricas de avaliação para a segmentação pela abordagem *Gumbel* enviesada à esquerda

A Tabela 13 apresenta as diferenças das médias das métricas de avaliação entre a segmentação por modelos de mistura Gaussianas com a abordagem *Gumbel* enviesada à esquerda. Com base na Tabela 13, as diferenças das médias das métricas para o tecido líquido cerebrospinal revela que a segmentação, em geral, dos dois métodos são idênticas porque o valor de *Dice* tem uma diferença de 0,0001 e *Jaccard* tem uma 0,0002, a favor da abordagem *Gumbel* enviesada à esquerda. Mas a abordagem *Gumbel* enviesada à esquerda tem um PPV e Especificidade melhor que a segmentação por modelos de mistura Gaussianas embora pior Sensibilidade, o que indica que a abordagem *Gumbel* enviesada à esquerda tem uma melhor precisão a classificar as intensidades que pertencem à classe e consegue identificar melhor as intensidades que não pertencem à classe mas perde na classificação de intensidades que realmente pertencem à classe mas são classificados como outro tecido.

<b>GUMBEL</b>	<b>Líquido Cerebrospinal</b>	<b>Substância Cinzenta</b>	<b>Substância Branca</b>
<b>Jaccard</b>	-0,0002	0,0089	0,0056
<b>Dice</b>	-0,0001	0,0065	0,0037
<b>PPV</b>	-0,0036	-0,0040	0,0200
<b>Sensibilidade</b>	0,0036	0,0158	-0,0155
<b>Especificidade</b>	-0,0023	-0,0052	0,0108

Tabela 13 - Diferença das médias das métricas de avaliação entre a segmentação por modelos de mistura Gaussianas e da abordagem *Gumbel* enviesada à esquerda

Para o caso da substância cinzenta, a avaliação da segmentação, no geral, mostra que a abordagem *Gumbel* enviesada à esquerda não consegue classificar melhor que a segmentação por modelos de mistura Gaussianas porque o valor da diferença da métrica *Dice* e *Jaccard* é 0,0065 e 0,0089, respetivamente a favor da segmentação por modelos de mistura Gaussianas. Mas a segmentação por modelos de mistura Gaussianas tem uma Sensibilidade melhor que a abordagem *Gumbel* enviesada à esquerda mas perde para a abordagem *Gumbel* na métrica Especificidade e PPV, o que indica a abordagem *Gumbel* enviesada à esquerda tem uma pior classificação de intensidades que realmente pertencem à classe mas são classificados como outro tecido mas classifica menos pontos que não pertencem à classe e tem uma melhor precisão das intensidades. Em relação à substância branca, os resultados das diferenças das médias das métricas revelam que a segmentação por modelos de mistura Gaussianas é, novamente, melhor que a abordagem *Gumbel* enviesada à esquerda, exceto na métrica Sensibilidade. Isto significa que a abordagem *Gumbel* enviesada à esquerda só é melhor a classificar intensidades que realmente pertencem ao tecido substância branca mas são classificados como outro tecido.

Realizado a análise da tabela, fomos comprovar a análise dos resultados das métricas dos dois métodos com um teste estatístico. O teste escolhido foi o teste *Wilcoxon* que comparou as médias de cada métrica da segmentação por modelos de mistura Gaussianas por processos de *Dirichlet* com as correspondentes da abordagem *Gumbel* enviesada à esquerda para cada tipo de tecido cerebral. A hipótese nula do teste é se a diferença das médias das amostras é igual a zero, ou seja, se a diferença é nula. O resultado obtido do teste *Wilcoxon* foi de  $p - \text{valor} = 0,416$ , para o caso das médias das métricas do tecido líquido cerebrospinal, para o caso do tecido substância cinzenta, foi de  $p - \text{valor} = 0,225$  e para o caso do tecido substância branca, foi de  $p - \text{valor} = 0,345$ , o que indica que a diferença para esta amostra não é estatisticamente significativa entre os métodos de segmentação para cada tipo de tecido, porque o p-valor de cada tecido é maior que 0.05. Com o resultado deste teste não podemos concluir que a segmentação pela abordagem *Gumbel* enviesada à esquerda seja melhor que a segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*.

Com isso, fomos comparar as matrizes de classificação de ambas as segmentações dos dois métodos para termos uma análise mais detalhada do desempenho de cada método. A Tabela 14 representa a matriz de classificação da segmentação pela abordagem *Gumbel* enviesada à esquerda, onde o valor 0 corresponde aos pontos que não interessam ser classificados, o valor 1 corresponde ao tecido líquido cerebrospinal, o valor 2 corresponde ao tecido substância cinzenta e o valor 3 corresponde ao tecido substância branca.

<i>Gumbel</i>	<i>Previsão 0</i>	<i>Previsão 1</i>	<i>Previsão 2</i>	<i>Previsão 3</i>
<b>Verdadeiro 0</b>	11519575	0	0	0
<b>Verdadeiro 1</b>	0	653555	87686	2037
<b>Verdadeiro 2</b>	0	96335	715100	120828
<b>Verdadeiro 3</b>	0	3399	58105	567380

Tabela 14 - Matriz de classificação da segmentação pela abordagem *Gumbel* enviesada à esquerda

Analisando as tabelas 7 e 14, verifica-se que para o caso do tecido líquido cerebrospinal (valor 1), a segmentação pela abordagem *Gumbel* enviesada à esquerda classifica menos pontos corretamente como líquido (653555 para a abordagem *Gumbel* enviesada à esquerda e 656285 para a segmentação por modelos de mistura Gaussianas) e deteta mais falsos positivos que deviam ser tecido substância cinzenta e branca que a segmentação por modelos de mistura Gaussianas (87686 que deviam ser substância cinzenta e 2037 substância branca, ambos da abordagem *Gumbel* enviesada à esquerda, em relação aos 85161 que deviam ser substância cinzenta e 1832 substância branca, ambos da segmentação por modelos de mistura Gaussianas) mas classifica menos pontos que não pertencem à classe mas que deviam pertencer do que a segmentação por modelos de mistura Gaussianas (96335 como substância cinzenta e 3399 como substância branca da abordagem *Gumbel* para 99974 como substância cinzenta e 3500 como substância branca do modelo de mistura Gaussianas). O que significa que a abordagem *Gumbel* enviesada à esquerda piora a segmentação do tecido líquido cerebrospinal, com o compromisso de perder alguns pontos como líquido cerebrospinal mas prevê melhor que a segmentação por modelos de mistura Gaussianas. No caso do tecido substância cinzenta (valor 2), a segmentação pela abordagem *Gumbel* enviesada à esquerda classifica menos pontos corretamente como substância cinzenta (715100 para a abordagem *Gumbel* e 729411 para a segmentação por modelos de mistura Gaussianas) e deteta menos falsos positivos que deviam ser tecido líquido cerebrospinal mas para o tecido substância branca identifica mais que a segmentação por modelos de mistura Gaussianas (96335 que deviam ser líquido cerebrospinal e 120828 substância branca, ambos da abordagem *Gumbel* enviesada à esquerda em relação aos 99974 que deviam ser líquido cerebrospinal e 102878 substância branca, ambos da segmentação por modelos de mistura Gaussianas) e também classifica mais pontos que não pertencem à classe mas que deviam e que foram classificados como líquido cerebrospinal mas menos como substância branca do que a segmentação por modelos de mistura de Gaussianas (87686 como líquido cerebrospinal e 58105 como substância branca para a abordagem *Gumbel* em relação a 85161 como líquido cerebrospinal e 67807 como substância branca da segmentação por modelos de mistura Gaussianas). E com isso, a abordagem *Gumbel* enviesada à esquerda piora a segmentação do tecido substância cinzenta com a desvantagem de perder definição na segmentação. E para o caso do tecido substância branca (valor 3), a segmentação pela abordagem *Gumbel* enviesada à esquerda classifica mais pontos verdadeiros como esse tecido (567380 para a abordagem *Gumbel* e 557577 para a segmentação por modelos de mistura Gaussianas) e

deteta menos falsos positivos que deviam ser tecido líquido cerebrospinal e substância cinzenta que a segmentação por modelos de mistura Gaussianas (3399 que deviam ser líquido cerebrospinal e 58105 substância cinzenta, ambos da abordagem *Gumbel* em relação aos 3500 que deviam ser líquido cerebrospinal e 67807 substância cinzenta, ambos da segmentação por modelos de mistura Gaussianas) mas classifica mais pontos que não pertencem à classe e que deviam pertencer do que o modelo de mistura de Gaussianas (2037 como líquido cerebrospinal e 120828 como substância cinzenta da abordagem *Gumbel* enviesada à esquerda em relação a 1832 como líquido cerebrospinal e 102878 como substância cinzenta da segmentação por modelos de mistura Gaussianas). Logo, a abordagem *Gumbel* enviesada à esquerda introduz uma pequena melhoria à segmentação do tecido substância branca que a segmentação por modelos de mistura Gaussianas, por classificar menos pontos que realmente não pertencem ao tecido substância branca e classificar mais pontos corretos mas com exclusão de algum tecido.

E com esses factos, chegamos à conclusão que a utilização da segmentação de tecidos cerebrais por combinação do modelo de mistura Ricianas por processos de *Dirichlet* para o tecido líquido cerebrospinal com o modelo de mistura Gaussianas por processos de *Dirichlet* para o tecido substância cinzenta e com o modelo de mistura *Gumbel* enviesada à esquerda para o tecido substância branca traduz-se numa pior classificação dos tecidos líquido cerebrospinal e substância cinzenta embora melhorasse, também com uma perda de definição, a classificação do tecido substância branca, devido à alteração da distribuição para *Gumbel* enviesada à esquerda, no caso do tecido substância branca. Tendo em conta estes factos, a distribuição *Gumbel* enviesada à esquerda não contribui para a segmentação do tecido substância branca.

## 5.5.7.3 Abordagem Log Gama

	Sujeito 1	Sujeito 2	Sujeito 3	Sujeito 4	Sujeito 5	Média±Desvio Padrão
<b>Líquido Cerebrospinal</b>						
<b>Jaccard</b>	0,8053	0,7608	0,7714	0,8079	0,7307	0,7753±0,0289
<b>Dice</b>	0,8922	0,8642	0,8710	0,8938	0,8444	0,8731±0,0184
<b>PPV</b>	0,9320	0,8299	0,8832	0,9187	0,7826	0,8693±0,0559
<b>Sensibilidade</b>	0,8556	0,9015	0,8591	0,8701	0,9168	0,8806±0,0243
<b>Especificidade</b>	0,9650	0,9184	0,9442	0,9588	0,9038	0,9381±0,0235
<b>Substância Cinzenta</b>						
<b>Jaccard</b>	0,6709	0,6159	0,6538	0,6716	0,6949	0,6614±0,0263
<b>Dice</b>	0,8031	0,7623	0,7907	0,8035	0,8200	0,7959±0,0192
<b>PPV</b>	0,8013	0,8319	0,7822	0,8466	0,8969	0,8318±0,0396
<b>Sensibilidade</b>	0,8049	0,7034	0,7993	0,7647	0,7552	0,7655±0,0365
<b>Especificidade</b>	0,8839	0,8908	0,8725	0,9057	0,9300	0,8966±0,0199
<b>Substância Branca</b>						
<b>Jaccard</b>	0,7727	0,6607	0,7811	0,7436	0,8163	0,7549±0,0525
<b>Dice</b>	0,8718	0,7957	0,8771	0,8529	0,8989	0,8593±0,0350
<b>PPV</b>	0,8315	0,7417	0,8756	0,7722	0,8639	0,8170±0,0520
<b>Sensibilidade</b>	0,9162	0,8582	0,8786	0,9526	0,9368	0,9085±0,0353
<b>Especificidade</b>	0,9303	0,8954	0,9449	0,9080	0,9428	0,9243±0,0195

Tabela 15 - Resultados das métricas de avaliação para a segmentação pela abordagem Log Gama

Comparando as Tabelas 4 e 15, observa-se que os resultados obtidos são muito parecidos. Mas com a diferença entre as médias das métricas de avaliação dos diferentes tecidos cerebrais dos dois métodos é possível analisar se houve melhorias por parte da abordagem *Log Gama* em relação à segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*. A Tabela 16 apresenta as diferenças das médias das métricas de avaliação entre a segmentação por modelos de mistura Gaussianas com a abordagem *Log Gama*. Analisando a Tabela 16, as diferenças das médias das métricas para o tecido líquido cerebrospinal mostra que a segmentação, em geral, dos dois métodos são idênticas porque o valor de *Dice* tem uma diferença de 0,0001 e *Jaccard* tem uma diferença de 0,0002, a favor da abordagem *Log Gama*. Mas a abordagem *Log Gama* tem um PPV e Especificidade melhor que a segmentação por modelos de mistura Gaussianas embora pior Sensibilidade, o que indica que a abordagem *Log Gama* tem uma melhor precisão a classificar as intensidades que pertencem à classe e consegue identificar melhor as intensidades que não pertencem à classe mas perde na classificação de intensidades que realmente pertencem à classe mas são classificados como outro tecido. Para o caso da substância cinzenta, a avaliação da segmentação, no geral, mostra que a abordagem *Log Gama* não consegue classificar melhor que a segmentação por modelos de mistura Gaussianas porque o valor

LOG GAMA	Líquido Cerebrospinal	Substância Cinzenta	Substância Branca
<b>Jaccard</b>	-0,0002	0,0120	0,0073
<b>Dice</b>	-0,0001	0,0087	0,0048
<b>PPV</b>	-0,0036	-0,0056	0,0254
<b>Sensibilidade</b>	0,0036	0,0211	-0,0196
<b>Especificidade</b>	-0,0023	-0,0072	0,0139

Tabela 16 - Diferença das médias das métricas de avaliação entre a segmentação por modelos de mistura Gaussianas e da abordagem Log Gama

diferença da métrica *Dice* e *Jaccard* é 0,0087 e 0,0120, respetivamente a favor da segmentação por modelos de mistura Gaussianas. Mas a segmentação por modelos de mistura Gaussianas tem uma Sensibilidade melhor que a abordagem *Log Gama* mas perde para a abordagem *Log Gama* na métrica Especificidade e PPV, o que indica que a abordagem *Log Gama* tem uma pior classificação de intensidades que realmente pertencem à classe mas são classificados como outro tecido mas classifica menos pontos que não pertencem à classe e tem uma melhor precisão das intensidades. Em relação à substância branca, os resultados das diferenças das médias das métricas revelam que a segmentação por modelos de mistura Gaussianas é melhor que a abordagem *Log Gama*, exceto na métrica Sensibilidade. Isto significa que a abordagem *Log Gama* só é melhor a classificar intensidades que realmente pertencem ao tecido substância branca mas são classificados como outro tecido.

Uma vez feita a análise da tabela, fomos comprovar a análise dos resultados das métricas dos dois métodos com um teste estatístico. O teste escolhido foi um teste *Wilcoxon* que comparou as médias de cada métrica da segmentação por modelos de mistura Gaussianas por processos de *Dirichlet* com as correspondentes da abordagem *Log Gama* para cada tipo de tecido cerebral. A hipótese nula do teste é se a diferença das médias das amostras é igual a zero, ou seja, se a diferença é nula. O resultado obtido do teste *Wilcoxon* foi de  $p - \text{valor} = 0,416$ , para o caso das médias das métricas do tecido líquido cerebrospinal, para o caso do tecido substância cinzenta, foi de  $p - \text{valor} = 0,225$  e para o caso do tecido substância branca, foi de  $p - \text{valor} = 0,345$ , o que indica que a diferença para esta amostra não é estatisticamente significativa entre os métodos de segmentação para cada tipo de tecido, porque o p-valor de cada tecido é maior que 0.05. Com o resultado deste teste não podemos concluir que a segmentação pela abordagem *Log Gama* seja melhor que a segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*.

Com isso, fomos comparar as matrizes de classificação de ambas as segmentações dos dois métodos para termos uma análise mais detalhada do desempenho de cada método. A Tabela 17 representa a matriz de classificação da segmentação pela abordagem *Log Gama*, onde o valor 0 corresponde aos pontos que não interessam ser classificados, o valor 1 corresponde ao tecido líquido cerebrospinal, o valor 2 corresponde ao tecido substância cinzenta e o valor 3 corresponde ao tecido substância branca.

<b>Log Gama</b>	<b>Previsão 0</b>	<b>Previsão 1</b>	<b>Previsão 2</b>	<b>Previsão 3</b>
<b>Verdadeiro 0</b>	11519575	0	0	0
<b>Verdadeiro 1</b>	0	653555	87578	2145
<b>Verdadeiro 2</b>	0	96335	709933	125995
<b>Verdadeiro 3</b>	0	3399	55510	569975

Tabela 17 - Matriz de classificação da segmentação pela abordagem *Log Gama*

Comparando as tabelas 7 e 17, verifica-se que para o caso do tecido líquido cerebrospinal (valor 1), a segmentação pela abordagem *Log Gama* classifica menos pontos corretamente como líquido cerebrospinal (653555 para a abordagem *Log Gama* e 656285 para a segmentação por modelos de mistura Gaussianas) e deteta mais falsos positivos que deviam ser tecido substância cinzenta e branca que a segmentação por modelos de mistura Gaussianas (87578 que deviam ser substância cinzenta e 2145 substância branca, ambos da abordagem *Log Gama* em relação aos 85161 que deviam ser substância cinzenta e 1832 substância branca, ambos da segmentação por modelos de mistura Gaussianas) mas classifica menos pontos que não pertencem à classe mas que deviam pertencer do que a segmentação por modelos de mistura Gaussianas (96335 como substância cinzenta e 3399 como substância branca da abordagem *Log Gama* para 99974 como substância cinzenta e 3500 como substância branca do modelo de mistura Gaussianas). O que significa que a abordagem *Log Gama* piora a segmentação do tecido líquido cerebrospinal, com o compromisso de perder alguns pontos como líquido cerebrospinal mas prevê melhor que a segmentação por modelos de mistura Gaussianas. Para o caso do tecido substância cinzenta (valor 2), a segmentação pela abordagem *Log Gama* classifica menos pontos corretamente como substância cinzenta (709933 para a abordagem *Log Gama* e 729411 para a segmentação por modelos de mistura Gaussianas) e deteta menos falsos positivos que deviam ser tecido líquido cerebrospinal mas para o tecido substância branca identifica mais que a segmentação por modelos de mistura Gaussianas (96335 que deviam ser líquido cerebrospinal e 125995 substância branca, ambos da abordagem *Log Gama* em relação aos 99974 que deviam ser líquido cerebrospinal e 102878 substância branca, ambos da segmentação por modelos de mistura Gaussianas) e classifica mais pontos que não pertencem à classe mas que deviam e que foram classificados como líquido cerebrospinal mas menos pontos que não pertencem à classe mas que deviam e são marcados como substância branca do que a segmentação por modelos de mistura de Gaussianas (87578 como líquido cerebrospinal e 55510 como substância branca da abordagem *Log Gama* para 85161 como líquido cerebrospinal e 67807 como substância branca da segmentação por modelos de mistura Gaussianas). E com isso, a abordagem *Log Gama* piora a segmentação do tecido substância cinzenta com a desvantagem de perder definição na segmentação. E para o caso do tecido substância branca (valor 3), a segmentação pela abordagem *Log Gama* classifica mais pontos corretamente como substância branca (569975 para a abordagem *Log Gama* e 557577 para a segmentação por modelos de mistura Gaussianas) e deteta menos falsos positivos que deviam ser tecido líquido cerebrospinal e substância cinzenta que a segmentação por modelos de mistura Gaussianas (3399 para o líquido cerebrospinal e 55510 para a substância cinzenta, ambos da abordagem *Log Gama* em relação aos 3500 para o líquido cerebrospinal e 67807 para a substância cinzenta, ambos da segmentação por modelos de mistura Gaussianas) mas classifica mais pontos que não pertencem à classe e que deviam pertencer do que a segmentação por modelos de mistura de Gaussianas (2145 como líquido cerebrospinal e 125995 como substância cinzenta da abordagem *Log Gama*

em relação a 1832 como líquido cerebrospinal e 102878 como substância cinzenta da segmentação por modelos de mistura Gaussianas). Por isso, a abordagem *Log Gama* introduz uma pequena melhoria à segmentação do tecido substância branca que a segmentação por modelos de mistura Gaussianas, por classificar menos pontos que realmente não pertencem ao tecido substância branca e classificar mais pontos corretos mas com a exclusão de algum tecido.

É com estes argumentos, concluímos que a utilização da segmentação de tecidos cerebrais por combinação do modelo de mistura Ricianas por processos de *Dirichlet* para o tecido líquido cerebrospinal com o modelo de mistura Gaussianas por processos de *Dirichlet* para o tecido substância cinzenta e com o modelo de mistura *Log Gama* para o tecido substância branca traduz-se numa pior classificação dos tecidos líquido cerebrospinal e substância cinzenta embora melhorasse, também com uma perda de definição, a classificação do tecido substância branca, devido à alteração da distribuição para *Log Gama*, no caso do tecido substância branca. Tendo em conta estes factos, a distribuição *Log Gama* não contribui para a segmentação do tecido substância branca.

#### 5.5.7.4 Abordagem *Powernorm*

	Sujeito 1	Sujeito 2	Sujeito 3	Sujeito 4	Sujeito 5	Média±Desvio Padrão
<b>Líquido Cerebrospinal</b>						
<b>Jaccard</b>	0,8054	0,7608	0,7714	0,8079	0,7307	0,7753±0,0289
<b>Dice</b>	0,8922	0,8642	0,8710	0,8938	0,8444	0,8731±0,0184
<b>PPV</b>	0,9321	0,8299	0,8832	0,9187	0,7826	0,8693±0,0560
<b>Sensibilidade</b>	0,8556	0,9015	0,8591	0,8701	0,9168	0,8806±0,0243
<b>Especificidade</b>	0,9650	0,9184	0,9442	0,9588	0,9038	0,9381±0,0235
<b>Substância Cinzenta</b>						
<b>Jaccard</b>	0,6677	0,6218	0,6559	0,6763	0,6948	0,6633±0,0243
<b>Dice</b>	0,8007	0,7668	0,7922	0,8069	0,8199	0,7973±0,0177
<b>PPV</b>	0,8029	0,8302	0,7795	0,8459	0,8969	0,8311±0,0400
<b>Sensibilidade</b>	0,7985	0,7124	0,8052	0,7713	0,7552	0,7685±0,0334
<b>Especificidade</b>	0,8860	0,8881	0,8695	0,9044	0,9299	0,8956±0,0204
<b>Substância Branca</b>						
<b>Jaccard</b>	0,7711	0,6640	0,7808	0,7476	0,8163	0,7560±0,0510
<b>Dice</b>	0,8707	0,7981	0,8769	0,8556	0,8988	0,8600±0,0339
<b>PPV</b>	0,8257	0,7503	0,8812	0,7786	0,8639	0,8199±0,0495
<b>Sensibilidade</b>	0,9210	0,8523	0,8726	0,9495	0,9367	0,9064±0,0376
<b>Especificidade</b>	0,9270	0,9007	0,9480	0,9116	0,9428	0,9260±0,0180

Tabela 18 - Resultados das métricas de avaliação para a segmentação pela abordagem *Powernorm*

Comparando as Tabelas 4 e 18, observa-se que os resultados obtidos de cada método são quase idênticos. Mas com a diferença entre as médias das métricas de avaliação dos diferentes tecidos cerebrais dos dois métodos é possível analisar se houve melhorias por parte da abordagem *Powernorm* em relação à segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*. A Tabela 19 apresenta as diferenças das médias das métricas de avaliação entre a segmentação por modelos de Gaussiana com a abordagem *Powernorm*.

<b>POWERNORM</b>	<b>Líquido Cerebrospinal</b>	<b>Substância Cinzenta</b>	<b>Substância Branca</b>
<b>Jaccard</b>	-0,0002	0,0101	0,0062
<b>Dice</b>	-0,0001	0,0073	0,0041
<b>PPV</b>	-0,0036	-0,0049	0,0225
<b>Sensibilidade</b>	0,0036	0,0181	-0,0175
<b>Especificidade</b>	-0,0023	-0,0062	0,0122

Tabela 19 - Diferença das médias das métricas de avaliação entre a segmentação por modelos de mistura Gaussianas e pela abordagem Powernorm

Observando a Tabela 19, as diferenças das médias das métricas para o tecido líquido cerebrospinal revela que a segmentação, em geral, dos dois métodos são parecidas porque o valor de *Dice* tem uma diferença de 0,0001 e *Jaccard* tem uma diferença de 0,0002, a favor da abordagem *Powernorm*. Mas a abordagem *Powernorm* tem um PPV e Especificidade melhor que a segmentação por modelos de mistura Gaussianas embora tenha uma Sensibilidade pior, o que indica que a abordagem *Powernorm* tem uma melhor precisão a classificar as intensidades que pertencem à classe e consegue classificar melhor as intensidades que não pertencem à classe mas perde na classificação de intensidades que realmente pertencem à classe mas são classificados como outro tecido. Para o caso da substância cinzenta, a avaliação da segmentação, no geral, mostra que a abordagem *Powernorm* não consegue classificar melhor que a segmentação por modelos de mistura Gaussianas porque o valor da diferença da métrica *Dice* e *Jaccard* é 0,0073 e 0,0101, respetivamente, a favor da segmentação por modelos de mistura Gaussianas. Mas a segmentação por modelos de mistura Gaussianas tem uma Sensibilidade melhor que a abordagem *Powernorm* mas piora na métrica Especificidade e PPV, o que indica que a abordagem *Powernorm* tem uma pior classificação de intensidades que realmente pertencem à classe mas são classificados como outro tecido mas classifica menos pontos que não pertencem à classe e uma melhor precisão das intensidades. Em relação à substância branca, os resultados das diferenças das médias das métricas revelam que a segmentação por modelos de mistura Gaussianas é melhor que a abordagem *Powernorm*, exceto na métrica Sensibilidade. Isto significa que a abordagem *Powernorm* só é melhor a classificar intensidades que realmente pertencem ao tecido substância branca mas são classificados como outro tecido.

Realizado a análise da tabela, fomos comprovar a análise dos resultados das métricas dos dois métodos com um teste estatístico. O teste escolhido foi um teste *Wilcoxon* que comparou as médias de cada métrica da segmentação por modelos de mistura Gaussianas por processos de *Dirichlet* com as correspondentes da abordagem *Powernorm* para cada tipo de tecido cerebral. A hipótese nula do teste é se a diferença das médias das amostras é igual a zero, ou seja, se a diferença é nula. O resultado obtido do teste *Wilcoxon* foi de  $p - valor = 0,416$ , para o caso das médias das métricas do tecido líquido cerebrospinal, para o caso do tecido substância cinzenta, foi de  $p - valor = 0,225$  e para o caso do tecido substância branca, foi de  $p - valor = 0,345$ , o que indica que a diferença para esta

amostra não é estatisticamente significativa entre os métodos de segmentação para cada tipo de tecido, porque o p-valor de cada tecido é maior que 0.05. Com o resultado deste teste não podemos concluir que a segmentação pela abordagem *Powernorm* seja melhor que a segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*.

Com isso, fomos comparar as matrizes de classificação de ambas as segmentações dos dois métodos para termos uma análise mais detalhada do desempenho de cada método. A Tabela 20 representa a matriz de classificação da segmentação pela abordagem *Powernorm*, onde o valor 0 corresponde aos pontos que não interessam ser classificados, o valor 1 corresponde ao tecido líquido cerebrospinal, o valor 2 corresponde ao tecido substância cinzenta e o valor 3 corresponde ao tecido substância branca.

<b>Powernorm</b>	<b>Previsão 0</b>	<b>Previsão 1</b>	<b>Previsão 2</b>	<b>Previsão 3</b>
<b>Verdadeiro 0</b>	11519575	0	0	0
<b>Verdadeiro 1</b>	0	653548	87622	2108
<b>Verdadeiro 2</b>	0	96335	712938	122990
<b>Verdadeiro 3</b>	0	3387	56837	568660

*Tabela 20 - Matriz de classificação da segmentação pela abordagem Powernorm*

Comparando as tabelas 7 e 20, verifica-se que para o caso do tecido líquido cerebrospinal (valor 1), a segmentação pela abordagem *Powernorm* classifica menos pontos verdadeiros como líquido cerebrospinal (653548 para a abordagem *Powernorm* e 656285 para a segmentação por modelos de mistura Gaussianas) e deteta mais falsos positivos que deviam ser tecido substância cinzenta e branca que a segmentação por modelos de mistura Gaussianas (87622 que deviam ser substância cinzenta e 2108 substância branca, ambos da abordagem *Powernorm* em relação aos 85161 que deviam ser substância cinzenta e 1832 substância branca, ambos da segmentação por modelos de mistura Gaussianas) mas classifica menos pontos que não pertencem ao tecido líquido cerebrospinal mas que deviam pertencer, do que a segmentação por modelos de mistura de Gaussianas (96335 como substância cinzenta e 3387 como substância branca da abordagem *Powernorm* para 99974 como substância cinzenta e 3500 como substância branca da segmentação por modelos de mistura Gaussianas). Isto significa que a abordagem *Powernorm* piora a segmentação do tecido líquido cerebrospinal, com o compromisso de perder alguns pontos como líquido cerebrospinal mas prevê melhor que a segmentação por modelos de mistura Gaussianas. Para o caso do tecido substância cinzenta (valor 2), a segmentação pela abordagem *Powernorm* classifica menos pontos corretamente como esse tecido (712938 para a abordagem *Powernorm* e 729411 para a segmentação por modelos de mistura Gaussianas) e deteta menos falsos positivos que deviam ser tecido líquido cerebrospinal mas para a substância branca classifica mais que a segmentação por modelos de mistura Gaussianas (96335 que deviam ser líquido cerebrospinal e 122990 substância branca, ambos da abordagem *Powernorm* em relação aos 99974 que deviam ser líquido cerebrospinal e 102878 para a substância branca, ambos da segmentação por modelos de mistura Gaussianas) e classifica mais pontos que não pertencem ao tecido substância cinzenta mas que deviam e que são classificados como líquido cerebrospinal mas menos pontos que não

pertencem à classe mas que deviam e são identificados como substância branca do que a segmentação por modelos de mistura de Gaussianas (87622 como líquido cerebrospinal e 56837 como substância branca da abordagem *Powernorm* para 85161 como líquido cerebrospinal e 67807 como substância branca da segmentação por modelos de mistura Gaussianas). Com isso, a abordagem *Powernorm* piora a segmentação do tecido substância cinzenta com a desvantagem de perder de detalhes na segmentação. E para o caso do tecido substância branca (valor 3), a segmentação pela abordagem *Powernorm* classifica mais pontos verdadeiros como esse tecido (568660 para a abordagem *Powernorm* e 557577 para a segmentação por modelos de mistura Gaussianas) e deteta menos falsos positivos que deviam ser tecido líquido cerebrospinal e substância cinzenta que a segmentação por modelos de mistura Gaussianas (3387 que deviam ser líquido cerebrospinal e 56837 substância cinzenta, ambos da abordagem *Powernorm* em relação aos 3500 que deviam ser líquido cerebrospinal e 67807 substância cinzenta, ambos da segmentação por modelos de mistura Gaussianas) mas classifica mais pontos como que não pertencem ao tecido substância branca mas que deviam pertencer, quer para o tecido líquido cerebrospinal quer para o tecido substância cinzenta, em relação à segmentação por modelos de mistura de Gaussianas (2108 como líquido cerebrospinal e 122990 como substância cinzenta da abordagem *Powernorm* para 1832 como líquido cerebrospinal e 102878 como substância cinzenta da segmentação por modelos de mistura Gaussianas). E por isso, a abordagem *Powernorm* introduz uma pequena melhoria à segmentação do tecido substância branca que a segmentação por modelos de mistura Gaussianas, por classificar menos pontos que realmente não pertencem ao tecido substância branca e classificar mais pontos corretos mas com exclusão de algum tecido.

Concluindo, a utilização da segmentação de tecidos cerebrais por combinação do modelo de mistura Ricianas por processos de *Dirichlet* para o tecido líquido cerebrospinal com o modelo de mistura Gaussianas por processos de *Dirichlet* para o tecido substância cinzenta e com o modelo de mistura *Powernorm* por processos de *Dirichlet* para o tecido substância branca afeta negativamente a classificação dos tecidos líquido cerebrospinal e substância cinzenta embora houvesse uma pequena melhoria, mas com uma perda de definição, a classificação do tecido substância branca, devido à alteração da distribuição para *Powernorm* para o caso do tecido substância branca. E por isso, a distribuição *Powernorm* não contribui para a segmentação do tecido substância branca excluindo a distribuição do modelo para a segmentação do tecido substância branca.

## 5.5.7.5 Abordagem Johnson Su

	Sujeito 1	Sujeito 2	Sujeito 3	Sujeito 4	Sujeito 5	Média±Desvio Padrão
<b>Líquido Cerebrospinal</b>						
<b>Jaccard</b>	0,8051	0,7606	0,7714	0,8079	0,7307	0,7752±0,0289
<b>Dice</b>	0,8920	0,8640	0,8710	0,8938	0,8444	0,8730±0,0184
<b>PPV</b>	0,9322	0,8298	0,8832	0,9188	0,7826	0,8693±0,0560
<b>Sensibilidade</b>	0,8552	0,9012	0,8591	0,8700	0,9168	0,8805±0,0243
<b>Especificidade</b>	0,9651	0,9184	0,9442	0,9589	0,9038	0,9381±0,0235
<b>Substância Cinzenta</b>						
<b>Jaccard</b>	0,6800	0,6321	0,6607	0,6935	0,7065	0,6746±0,0261
<b>Dice</b>	0,8095	0,7746	0,7957	0,8190	0,8280	0,8054±0,0187
<b>PPV</b>	0,8236	0,7291	0,8207	0,7966	0,7758	0,7892±0,0347
<b>Sensibilidade</b>	0,7959	0,8261	0,7721	0,8427	0,8878	0,8249±0,0398
<b>Especificidade</b>	0,8236	0,7291	0,8207	0,7966	0,7758	0,7892±0,0347
<b>Substância Branca</b>						
<b>Jaccard</b>	0,7764	0,6688	0,7776	0,7625	0,8238	0,7618±0,0509
<b>Dice</b>	0,8741	0,8015	0,8749	0,8653	0,9034	0,8638±0,0337
<b>PPV</b>	0,9012	0,8397	0,8549	0,9365	0,9186	0,8902±0,0370
<b>Sensibilidade</b>	0,8486	0,7667	0,8957	0,8041	0,8888	0,8408±0,0494
<b>Especificidade</b>	0,9012	0,8397	0,8549	0,9365	0,9186	0,8902±0,0370

Tabela 21 - Resultados das métricas de avaliação para a segmentação pela abordagem Johnson Su

Comparando as Tabelas 4 e 21, observa-se que os resultados obtidos são quase semelhantes. Mas com a diferença entre as médias das métricas de avaliação dos diferentes tecidos cerebrais dos dois métodos é possível analisar se houve melhorias por parte da abordagem Johnson Su em relação à segmentação por modelos de mistura Gaussianas por processos de Dirichlet. A Tabela 22 apresenta as diferenças das médias das métricas de avaliação entre a segmentação por modelos de mistura Gaussianas com a abordagem Johnson Su. Analisando a Tabela 22, as diferenças das médias das métricas para o tecido líquido cerebrospinal revela que a segmentação, em geral, dos dois métodos são idênticas porque o valor de Dice é igual a 0 e Jaccard tem uma diferença de 0,0001 a favor da abordagem Johnson Su. Mas a abordagem Johnson Su tem um PPV e Especificidade melhor que a segmentação por modelos de mistura Gaussianas embora tenha uma Sensibilidade pior, o que significa que a abordagem Johnson Su tem uma melhor precisão a classificar as intensidades que pertencem à classe e consegue identificar melhor as intensidades que não pertencem à classe mas perde na classificação de intensidades que realmente pertencem à classe mas são classificados como outro tecido.

JOHNSON SU	Líquido Cerebrospinal	Substância Cinzenta	Substância Branca
<b>Jaccard</b>	-0,0001	-0,0012	0,0004
<b>Dice</b>	0,0000	-0,0008	0,0003
<b>PPV</b>	-0,0036	0,0013	0,0016
<b>Sensibilidade</b>	0,0037	-0,0026	-0,0013
<b>Especificidade</b>	-0,0023	0,0014	0,0008

Tabela 22 - Diferença das médias das métricas de avaliação entre a segmentação por modelos de mistura Gaussianas e pela abordagem Johnson Su

Para o caso da substância cinzenta, a avaliação da segmentação, no geral, revela que a abordagem *Johnson Su* consegue classificar melhor que a segmentação por modelos de mistura Gaussianas porque o valor da diferença da métrica *Dice* e *Jaccard* é 0,0008 e 0,0012, respetivamente a favor da abordagem *Johnson Su*. Mas a segmentação por modelos de mistura Gaussianas tem PPV e Especificidade melhor que a abordagem *Johnson Su* mas perde para a abordagem *Johnson Su* na métrica Sensibilidade, o que significa que a abordagem *Johnson Su* tem uma melhor classificação de intensidades que realmente pertencem à classe mas são classificados como outro tecido mas classifica mais pontos que não pertencem à classe e tem uma pior precisão das intensidades. Em relação à substância branca, os resultados das diferenças das médias das métricas revelam que a segmentação por modelos de mistura Gaussianas, no geral, é quase idêntica à segmentação do modelo de mistura *Johnson Su* porque o valor de *Dice* tem uma diferença 0,0003 e *Jaccard* tem uma diferença de 0,0004. Mas a abordagem *Johnson Su* tem um PPV e Especificidade pior que a segmentação por modelos de mistura Gaussianas embora tenha uma Sensibilidade melhor, o que indica a abordagem *Johnson Su* identifica melhor as intensidades que realmente são substância branca mas tem uma pior precisão das intensidades e classifica pior as intensidades que não pertencem a substância branca.

Uma vez feita a análise da tabela, fomos comprovar a análise dos resultados das métricas dos dois métodos com um teste estatístico. O teste escolhido foi um teste *Wilcoxon* que comparou as médias de cada métrica da segmentação por modelos de mistura Gaussianas por processos de *Dirichlet* com as correspondentes da abordagem *Johnson Su* para cada tipo de tecido cerebral. A hipótese nula do teste é se a diferença das médias das amostras é igual a zero, ou seja, se a diferença é nula. O resultado obtido do teste *Wilcoxon* foi de  $p - valor = 0,715$ , para o caso das médias das métricas do tecido líquido cerebrospinal, para o caso do tecido substância cinzenta, foi de  $p - valor = 0,893$  e para o caso do tecido substância branca, foi de  $p - valor = 0,345$ , o que indica que a diferença para esta amostra não é estatisticamente significativa entre os métodos de segmentação para cada tipo de tecido, porque o p-valor de cada tecido é maior que 0.05. Com o resultado deste teste não podemos concluir que a segmentação pela abordagem *Johnson Su* seja melhor que a segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*.

Com isso, fomos comparar as matrizes de classificação de ambas as segmentações dos dois métodos para termos uma análise mais detalhada do desempenho de cada método. A Tabela 23 representa a matriz de classificação da segmentação pela abordagem *Johnson Su*, onde o valor 0 corresponde aos pontos que não interessam ser classificados, o valor 1 corresponde ao tecido líquido cerebrospinal, o valor 2 corresponde ao tecido substância cinzenta e o valor 3 corresponde ao tecido substância branca.

<i>Johnson Su</i>	<i>Previsão 0</i>	<i>Previsão 1</i>	<i>Previsão 2</i>	<i>Previsão 3</i>
<i>Verdadeiro 0</i>	11519575	0	0	0
<i>Verdadeiro 1</i>	0	653440	87668	2170
<i>Verdadeiro 2</i>	0	96329	732034	103900
<i>Verdadeiro 3</i>	0	3361	67167	558356

*Tabela 23 - Matriz de classificação da segmentação pela abordagem Johnson Su*

Comparando as tabelas 7 e 23, verifica-se que para o caso do tecido líquido cerebrospinal (valor 1), a segmentação pela abordagem *Johnson Su* classifica menos pontos verdadeiros como líquido cerebrospinal (653440 para a abordagem *Johnson Su* e 656285 para a segmentação por modelos de mistura Gaussianas) e deteta mais falsos positivos que deviam ser tecido substância cinzenta e branca que a segmentação por modelos de mistura Gaussianas (87668 que deviam ser substância cinzenta e 2170 substância branca, ambos da abordagem *Johnson Su* em relação aos 85161 que deviam ser substância cinzenta e 1832 substância branca, ambos da segmentação por modelos de mistura Gaussianas) mas classifica menos pontos que não pertencem ao tecido líquido cerebrospinal mas que deviam pertencer, do que a segmentação por modelos de mistura de Gaussianas (96329 como substância cinzenta e 3361 como substância branca da abordagem *Johnson Su* para 99974 como substância cinzenta e 3500 como substância branca da segmentação por modelos de mistura Gaussianas). O que significa que a abordagem *Johnsons Su* piora a segmentação do tecido líquido cerebrospinal, com o compromisso de perder alguns pontos como líquido cerebrospinal para outros tecidos mas prevê melhor que a segmentação por modelos de mistura Gaussianas. Para o caso do tecido substância cinzenta (valor 2), a segmentação pela abordagem *Johnson Su* classifica mais pontos corretamente como esse tecido (732034 para a abordagem *Johnson Su* e 729411 para a segmentação por modelos de mistura Gaussianas) e deteta menos falsos positivos que deviam ser tecido líquido cerebrospinal mas para a substância branca classifica mais que a segmentação por modelos de mistura Gaussianas (96329 que deviam ser líquido cerebrospinal e 103900 substância branca, ambos da abordagem *Johnson Su* em relação aos 99974 que deviam ser líquido cerebrospinal e 102878 substância branca, ambos da segmentação por modelos de mistura Gaussianas) e classifica mais pontos que não pertencem à classe mas que deviam e que foram classificados como tecido líquido cerebrospinal mas menos pontos que não pertencem à classe mas que deviam e são marcados como substância branca do que a segmentação por modelos de mistura de Gaussianas (87668 como líquido cerebrospinal e 67167 como substância branca da abordagem *Johnson Su* para 85161 como líquido cerebrospinal e 67807 como substância branca da segmentação por modelos de mistura Gaussianas). E com isso, a abordagem *Johnson Su* tem uma pequena melhoria da segmentação do tecido substância cinzenta com o aspecto negativo de agravar ligeiramente a definição da segmentação. E para o caso do tecido substância branca (valor 3), a segmentação pela abordagem *Johnson Su* classifica mais pontos verdadeiros como esse tecido (558356 para a abordagem *Johnson Su* e 557577 para a segmentação por modelos de mistura Gaussianas)

e deteta menos falsos positivos que deviam ser tecido líquido cerebrospinal e substância cinzenta que a segmentação por modelos de mistura Gaussianas (3361 para o líquido cerebrospinal e 67167 para a substância cinzenta, ambos da abordagem *Johnson Su* em relação aos 3500 para o líquido cerebrospinal e 67807 para a substância cinzenta, ambos da segmentação modelo de mistura Gaussianas) mas classifica mais pontos que não pertencem à substância branca mas que deviam pertencer do que a segmentação por modelos de mistura de Gaussianas (2170 como líquido cerebrospinal e 103900 como substância cinzenta da abordagem *Johnson Su* para 1832 como líquido cerebrospinal e 102878 como substância cinzenta da segmentação por modelos de mistura Gaussianas). E com isso, a abordagem *Johnson Su* tem uma ligeira melhoria na segmentação do tecido substância branca do que a segmentação por modelos de mistura Gaussianas mas com a cedência de pormenores na segmentação.

Concluindo, a utilização da segmentação de tecidos cerebrais por combinação do modelo de mistura Ricianas por processos de *Dirichlet* para o tecido líquido cerebrospinal com o modelo de mistura Gaussianas por processos de *Dirichlet* para o tecido substância cinzenta e com o modelo de mistura *Johnson Su* por processos de *Dirichlet* para o tecido substância branca interfere ligeiramente na classificação dos tecidos líquido cerebrospinal mas para o tecido substância cinzenta contribui para a sua segmentação. E para o tecido substância branca embora seja pior em relação à segmentação por modelos de mistura Gaussianas, a sua diferença é muito pequena, mas consegue detetar mais intensidades corretas com a introdução de poucas intensidades erradas, devido à alteração da forma da distribuição que o modelo usa para *Johnson Su*. E na globalidade, a abordagem *Johnson Su* contribui para a segmentação do tecido substância cinzenta e substância branca.

Resumindo, das abordagens apresentadas para adaptarmos o modelo de mistura por processos de *Dirichlet* às intensidades do tecido substância branca e que melhora a segmentação dos tecidos cerebrais, o que apresentou melhores resultados para a segmentação dos tecidos cerebrais foi a abordagem *Johnson Su*, devido à melhoria na segmentação do tecido substância cinzenta e branca embora a melhoria seja mínima. Porém, mesmo com a pequena melhoria que a distribuição *Johnson Su* no modelo de mistura por processos de *Dirichlet* introduz, a sua segmentação apresenta os mesmos defeitos que a segmentação por modelos de mistura Gaussianas por processos de *Dirichlet*, já apresentadas na seção anterior. A figura 39 apresenta uma comparação entre a segmentação do modelo de mistura Gaussianas com a abordagem *Johnson Su*, com ambos os métodos assinalados com marcações das zonas mal classificadas, onde a cor azul corresponde ao tecido líquido cerebrospinal, a cor verde corresponde ao tecido substância cinzenta e a cor vermelha corresponde ao tecido substância branca, e com as marcações do médico, onde apresentam os mesmos defeitos.

---

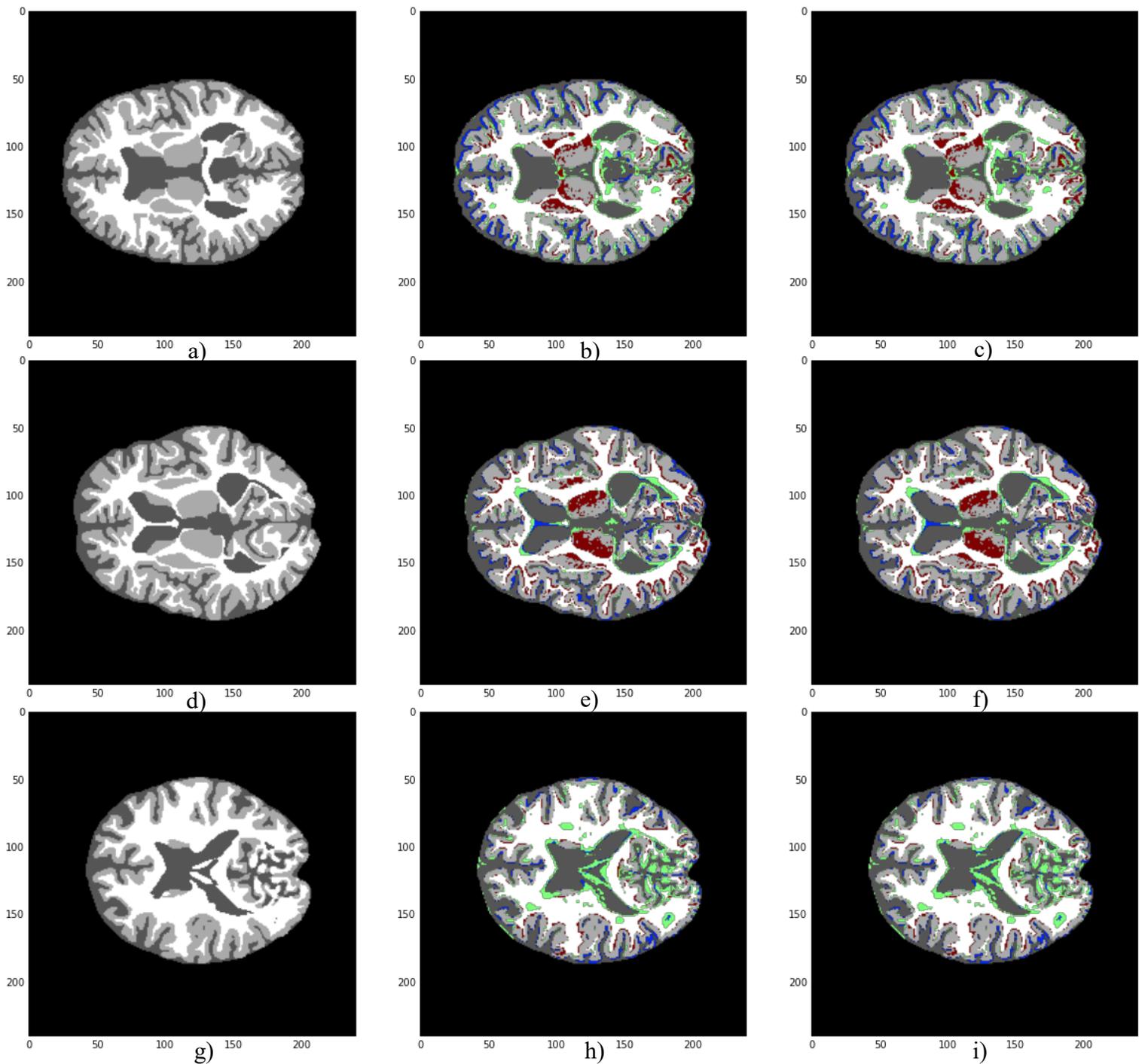


Figura 39 - Comparação da segmentação por modelos de mistura Gaussianas por processos de Dirichlet com a segmentação pela abordagem Johnson Su, onde ambos têm marcações das zonas mal classificadas, e com a marcação do médico. a), d) e g) Marcação do médico; b), e) e h) Segmentação por modelos de mistura Gaussianas; c), f) e i) Segmentação pela abordagem Johnson Su

## 5.6 Sumário

Neste capítulo foi apresentado o trabalho experimental realizado. Foram apresentados exemplos para que o modelo de mistura de Gaussianas por processos de *Dirichlet* calculasse a distribuição preditiva dos dados, onde foi possível demonstrar que o modelo é capaz de estimar, a partir de amostras, a distribuição preditiva. Também foi apresentado o procedimento que é necessário para calcularmos o modelo para obtermos a probabilidade preditiva dos dados. De seguida foi demonstrado a aplicação do modelo de mistura Gaussianas por processos de *Dirichlet* na segmentação de tecidos cerebrais de imagens de ressonância magnética, onde se discutiu as imagens utilizadas, do seu pré-processamento, dos histogramas das intensidades dos tecidos cerebrais e dos resultados obtidos da modelação das intensidades dos tecidos cerebrais segundo o nosso modelo e da segmentação dos tecidos. A partir da análise da segmentação obtida verificaram-se certas limitações e para superarmos essas limitações, sugerimos alterações às distribuições utilizadas pelo modelo para se adaptarem melhor segundo a distribuição dos dados das intensidades dos tecidos líquido cefalorraquidiano e substância branca. E também apresentamos os resultados obtidos do trabalho experimental. Foi avaliado o uso de modelos de mistura de Gaussianas por processos de *Dirichlet* na segmentação de tecidos cerebrais de imagens de ressonância magnética e revelou-se que apresenta resultados melhores que métodos que usam abordagens paramétricas, comparando com o método *Atropos*. Também se mostrou os resultados obtidos da combinação de modelos de mistura para adaptarmos às intensidades do líquido cefalorraquidiano e comparou-se os resultados aos dos modelos de mistura Gaussianas por processos de *Dirichlet* e revelou que a combinação de modelos, ou seja, com a introdução da distribuição Rician no modelo de mistura para as intensidades do tecido líquido cefalorraquidiano consegue uma segmentação com uma pequena diferença para melhor que os modelos de mistura Gaussianas. E, do mesmo modo, foram apresentados os resultados obtidos da combinação de modelos de mistura para adaptarmos às intensidades do substância branca e comparou-se os resultados aos dos modelos de mistura Gaussianas por processos de *Dirichlet*. Das 5 distribuições apresentadas para se adaptar ao modelo de mistura por processos de *Dirichlet* para as intensidades do tecido substância branca, só a distribuição *Johnson Su* é que contribui para a melhoria da segmentação dos tecidos cerebrais embora tenha contribuído mais para a segmentação do tecido substância cinzenta do que a substância branca com uma diferença pequena e interfere ligeiramente na do tecido líquido cefalorraquidiano. Mas mesmo com estas melhorias a segmentação apresenta problemas que modelos estatísticos não conseguem superar, como por exemplo, na imagem de ressonância magnética apresentar zonas onde as intensidades são muito semelhantes o que leva a uma sobreposição de intensidades.

## 6. CONCLUSÃO E TRABALHO FUTURO

Neste capítulo é resumido todo o trabalho efetuado e desenvolvido, apresentando as conclusões retiradas dos resultados obtidos, eventuais melhorias aos modelos e outras abordagens, em trabalho futuro.

### 6.1 Conclusões

Apresentamos, nesta dissertação, a segmentação de tecidos cerebrais por modelos de mistura Gaussianos por processos de *Dirichlet* para imagens de ressonância magnética. O principal objetivo era obter um modelo de mistura Gaussianas para cada tecido cerebral. Para isso, foi estudado o modelo de mistura de Gaussianas por processos de *Dirichlet*, onde explicou-se a estrutura do modelo para calcular-se a distribuição preditiva dos dados. Uma vez que o modelo é capaz de calcular a distribuição preditiva ou *a posteriori* de dados foi demonstrado que a aplicação do modelo para segmentação de tecidos cerebrais em imagens cerebrais de ressonância magnética é possível e com bons resultados embora como algumas limitações. Atingido objetivo principal foi-se à procura de melhorias e otimizações ao modelo de mistura proposto. Foi apresentado uma modificação do modelo para os dados das intensidades do líquido cerebrospinal e da substância branca devido à propriedade que o modelo tem de trabalhar com distribuições de famílias exponenciais e convexas é possível alterar a distribuição para outras distribuições da família exponencial. Para o caso do tecido líquido cerebrospinal modificou-se o modelo para trabalhar com a distribuição Rician. Essa modificação foi avaliado na segmentação e mostrou-se que consegue segmentar melhor os tecidos que o modelo de mistura Gaussianas mas a sua contribuição é pequena. E para o caso do tecido substância branca foram propostas cinco distribuições alternativas ao modelo de mistura Gaussianas para as intensidades da substância branca. Esses modelos modificados foram avaliados e mostrou-se que só uma das modificações, a do modelo de mistura *Johnson Su*, tem uma contribuição reduzida, similarmente ao modelo que usa a distribuição Rician.

### 6.2 Trabalho Futuro

Na eventualidade de haver uma continuação do desenvolvimento deste trabalho, existem possíveis abordagens a considerar para melhorar o modelo ou de implementação. Para melhorar o modelo deve-se ir estudar uma forma de diminuir a sobreposição das intensidades da substância cinzenta com os restantes tecidos, como por exemplo, adaptar uma nova distribuição ao modelo para trabalhar melhor as intensidades da substância cinzenta e fazer mais testes com mais imagens de ressonância magnética. Para a implementação do modelo deveria ser introduzido num sistema como característica e verificar a sua contribuição para a classificação dos tecidos cerebrais ou ser implementado como um classificador e utilizado com outros classificadores e testar a segmentação dos tecidos cerebrais.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Suzuki, Kenji (2014). *Computational Intelligence in Biomedical Imaging*, Springer, página 295
- [2] Silva, Adelino R. Ferreira da (2006). *Dirichlet process mixture model for brain MRI tissue classification*, Medical Image Analysis 11 2007, pág. 169-182
- [3] Balafar, M.A (2012). *Gaussian mixture model based segmentation methods for brain MRI images*, Artificial Intelligence Review (2014), Springer Science + Business Media B.V. 2012
- [4] BS (2009). [http://www.scholarpedia.org/article/Bayesian\\_statistics](http://www.scholarpedia.org/article/Bayesian_statistics), Bayesian Statistics, Scholarpedia
- [5] ISBA (2009). <https://bayesian.org/Bayes-Explained>, International Society for Bayesian Analysis
- [6] Pereira Polidoro, Maria João Fernandes (2013). *Metodologia Bayesiana e Adequação de Modelos*. Tese de Doutoramento, Departamento de Estatística e Investigação Operacional da F.C da Universidade de Lisboa, Lisboa
- [7] Orbanz, Peter (2014) *Lecture Notes on Bayesian Nonparametrics*. [http://stat.columbia.edu/~porbanz/papers/porbanz\\_BNP\\_draft.pdf](http://stat.columbia.edu/~porbanz/papers/porbanz_BNP_draft.pdf)
- [8] Fernando da Paz, Rosineide (2013) *Aspectos práticos da estimação do modelo de mistura via processo de Dirichlet*. Departamento de Estatística da Universidade Federal de São Carlos
- [9] Distribuição Dirichlet (Outubro 2015). <https://docs.scipy.org/doc/numpy-1.6.0/reference/generated/numpy.random.mtrand.dirichlet.html>, Scipy Org
- [10] Distribuição Dirichlet (Outubro 2015). [https://en.wikipedia.org/wiki/Dirichlet\\_distribution](https://en.wikipedia.org/wiki/Dirichlet_distribution), Wikipedia
- [11] Distribuição Dirichlet (Outubro 2015). <http://docs.scipy.org/doc/scipy-0.16.0/reference/generated/scipy.stats.dirichlet.html>, Scipy Org
- [12] Distribuição Rician (Outubro 2015). <http://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.stats.rice.html>, Scipy Org
- [13] Distribuição Rician (Outubro 2015). [https://en.wikipedia.org/wiki/Rice\\_distribution](https://en.wikipedia.org/wiki/Rice_distribution), Wikipedia
- [14] Distribuição Log Gama (Outubro 2015). <http://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.stats.loggamma.html#scipy.stats.loggamma>, Scipy Org

- [15] Distribuição Johnson Su (Outubro 2015). <http://docs.scipy.org/doc/scipy-0.16.0/reference/generated/scipy.stats.johnsonsu.html>, Scipy Org
- [16] Distribuição Normal (Outubro 2015). [https://en.wikipedia.org/wiki/Normal\\_distribution](https://en.wikipedia.org/wiki/Normal_distribution), Wikipedia
- [17] Distribuição T-Student multivariável/multivariada (Outubro 2015)-  
[https://en.wikipedia.org/wiki/Multivariate\\_t-distribution](https://en.wikipedia.org/wiki/Multivariate_t-distribution), Wikipedia
- [18] Distribuição Gumbel enviesada à esquerda- [http://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.stats.gumbel\\_1.html#scipy.stats.gumbel\\_1](http://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.stats.gumbel_1.html#scipy.stats.gumbel_1), Scipy Org
- [19] Distribuição GenLogistic (Outubro 2015). <http://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.stats.genlogistic.html#scipy.stats.genlogistic>, Scipy Org
- [20] Distribuição Powernorm (Outubro 2015). <http://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.stats.powernorm.html#scipy.stats.powernorm>, Scipy Org
- [21] Blei, David. e M. Jordan, Michael I. (2006). *Variational Inference for Dirichlet Process Mixtures*. International Society for Bayesian Analysis 2006, pp 121-144
- [22] Das, Rajarshi (2014). *Fast Collapsed Gibbs Sampler for Dirichlet Process Gaussian Mixture Models using Rank 1 Cholesky updates*. C-lab lecture.
- [23] Ferguson, T. (1973). *A Bayesian analysis of some nonparametric problems*. The Annals of Statistics, 1:209-230.
- [24] El-Arini, Khalid (2008). *Dirichlet Processes, A gentle tutorial*, SELECT Lab Meeting, [https://www.cs.cmu.edu/~kbe/dp\\_tutorial.pdf](https://www.cs.cmu.edu/~kbe/dp_tutorial.pdf)
- [25] Frigyik, Bela A., Kapila, e Gupta, Maya R. (2010). *Introduction to the Dirichlet Distribution and Related Processes*, UWEE Technical Report, Department Electrical Engineering, University of Washington
- [26] Chen, Edwin (2012). *Infinite Mixture Models with Nonparametric Bayes and the Dirichlet Process*. <http://blog.echen.me/2012/03/20/infinite-mixture-models-with-nonparametric-bayes-and-the-dirichlet-process/>
- [27] The Dirichlet Process and the Chinese Restaurant Process and other representations (2014). <http://blog.datumbox.com/the-dirichlet-process-the-chinese-restaurant-process-and-other-representations/>
- [28] Teh, Yee Whye (2007). *Dirichlet Processes: Tutorial and Practical Course (updated)*. Gatsby Computational Neuroscience Unit, University College London.
- [29] Jordan, Michael I. (2005). *Dirichlet Processes, Chinese Restaurant Processes and All That*. Department of Statistics and Computer Science Division, University of California, Berkeley.

- [30] Alemu, Tilaye Yismaw (2007). *From Mixtures to Mixing*. Master of Science Artificial Intelligence, School of Informatics, University of Edinburgh
- [31] Popuri et al (2011). *3D variational brain tumor segmentation using Dirichlet priors on a clustered feature set*.
- [32] Schmidt et al (2013). *Fully Bayesian Inference for Structural MRI: Application to Segmentation and Statistical Analysis of T2-Hypointensities*. PLOS ONE
- [33] Yisu Lu et al (2014). *Multimodal Brain-Tumor Segmentation Based on Dirichlet Process Mixture Model with Anisotropic Diffusion and Markov Random Field Prior*. Computational and Mathematical Methods in Medicine, Volume 2014 (2014), Article ID 717206.
- [34] Zhang et al (2001). *Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation – Maximization Algorithm*. IEEE Transactions on Medical Imaging, Vol 20, N°1, January 2001
- [35] Bricq et al (2008). *MS Legion Segmentation based on Hidden Markov Chains*.
- [36] Liang et al (2009). *An EM Approach to MAP Solution of Segmenting Tissue Mixture: A Numerical Analysis*. IEEE Trans Med Imaging. 2009 February
- [37] Balafar (2013). *Fast and Robust Gaussian Mixture Model for MRI Brain Image Segmentation*. International Journal on “Technical and Physical Problems of Engineering”, Issue 15, Volume 5, N° 2, June 2013.
- [38] Kurihara, Kenichi *et al* (2006). *Accelerated Variational Dirichlet Process Mixtures*. Neural Information Processing Systems, Neural Information Processing Systems 2006.
- [39] Sengupta (2004). *Kolmogorov Theorem*. Lecture Notes
- [40] Blackwell e MacQueen (1973). *Ferguson distributions via Pólya urn schemes*. *The Annals of Statistics*, 1(2):353-355
- [41] Sethuraman (1994). *A constructive definition of Dirichlet priors*. *Statistica Sinica*, 4:639-650
- [42] Rolin (1992). *Some useful properties of the Dirichlet Process* (Tech. Rep.). Université Catholique de Louvain, Belgium.
- [43] Johnson e Sinanovic. *Symmetrizing the Kullback-Leibler distance*. Computer and Information Technology Institute, Department of Electrical and Computer Engineering, Rice University, Houston, Texas

- [44] Distribuição Multinomial (Outubro 2015). <http://www.portalaction.com.br/probabilidades/55-distribuicao-multinomial>, Portal Action
- [45] Murayama, Hitoshi. *Dirac Delta Function*. <http://hitoshi.berkeley.edu/221a/delta.pdf>
- [46] Função Delta Dirac (Outubro 2015). <https://jkogler.wordpress.com/2008/06/14/delta-de-dirac-ou-impulso-unitario/>, Wordpress
- [47] Função Delta Dirac (Outubro 2015). [https://pt.wikipedia.org/wiki/Ficheiro:Dirac\\_distribution\\_PDF.png](https://pt.wikipedia.org/wiki/Ficheiro:Dirac_distribution_PDF.png), Wikipedia
- [48] Base de Dados MRBrainS (Outubro 2014). <http://mrbrains13.isi.uu.nl/details.php>
- [49] Pereira, Sérgio (2013). *Detecção Automática de Espaços de Virchow-Robin em imagens de ressonância magnética*. Dissertação de Mestrado, Mestrado Integrado em Engenharia Biomédica, Ramo Electrónica Médica, Escola de Engenharia, Universidade do Minho.
- [50] Kabir et al (2009). *Multimodal MRI segmentation of ischemic stroke lesions*. Proceedings of The 29th Annual International Conference of the IEEE Engineering in Medicine in Biology Society, 2007.
- [51] Birenbaum, Dale, Bancroft, Laura W. e Felsberg, Gary J. (2011). *Imaging in Acute Stroke*. West J Emerg Med, February 2011.
- [52] Chupin et al. (2009). *Fully Automatic Hippocampus Segmentation and Classification in Alzheimer's 2 Disease and Mild Cognitive Impairment Applied on Data from ADNI*. Hippocampus, 19: 579-587.
- [53] Carmichael et al (2011). *Atlas-Based Hippocampus Segmentation In Alzheimer's Disease and Mild Cognitive Impairment*. Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania, November 2004.
- [54] Kinnunen et al (2014). *Automated Segmentation of Thalamus from MRI: Method Validation and Comparison for Volumetric Measurement in Familial Alzheimer's Disease*. Alzheimer's & Dementia: The Journal of the Alzheimer's Association, Vol. 10, Issue 4, P107-108, July 2014.
- [55] Matriz de classificação (Março 2016) [http://www.saedsayad.com/model\\_evaluation\\_c.htm](http://www.saedsayad.com/model_evaluation_c.htm)
- [56] Matriz de classificação (Março 2016) [http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion\\_matrix/confusion\\_matrix.html](http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html)
- [57] Adrienne M.Mendrik, Koel L. Vincken, Hugo J. Kuijf, et al (2015). *MRBrainS Challenge: Online Evaluation Framework for Brain Image Segmentation in 3T MRI Scans*. Computational Intelligence and Neuroscience, Vol. 2015, Article ID 813696, 16 pages. doi:10.1155/2015/813696

