**Aproximação numérica à
convolução de Mellin
via mistura de exponenciais**

Jorge Luis Torrejón Matos

TESE APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
DOUTOR EM CIÊNCIAS

Programa: Matemática Aplicada
Orientador: Prof. Dr. Julio Michael Stern

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CAPES

São Paulo, Outubro de 2015

**Aproximação numérica à
convolução de Mellin
via mistura de exponenciais**

Este exemplar corresponde à redação
final da tese devidamente corrigida
e defendida por Jorge Luis Torrejón Matos
e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Prof. Dr. Julio Michael Stern (orientador) - IME-USP.
- Prof. Dr. Adilson Simonis - IME-USP.
- Prof. Dr. Marcelo de Souza Lauretto - EACH-USP.
- Prof. Dr. Laura Leticia Ramos Rifo - IMECC-UNICAMP.
- Prof. Dr. Adriano Polpo de Campos - UFSCAR.

Agradecimentos

Agradeço a Deus, por estar sempre ao meu lado nas minhas quedas, lutas, vitórias e derrotas.

Aos meus pais, Jorge Luis Torrejón Meza e Sonia Ruth Matos de Torrejón, aos meus irmãos e meus familiares mais próximos, por estarem sempre comigo apoiando e incentivando para seguir sempre em frente. Esta vitória também é de vocês.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro.

Ao programa de Pós-graduação em Matemática Aplicada do Instituto de Matemática e Estatística (IME-USP), por me oferecer a oportunidade de realizar este doutorado.

Ao meu orientador, o professor Dr. Julio Stern e ao professor Dr. Adilson Simonis, por direcionar minha pesquisa e auxiliar na execução e nas soluções de dúvidas que surgiram ao longo do trabalho.

Aos professores do IME, que de alguma maneira contribuíram nesta minha caminhada, em especial aos professores que ministraram as disciplinas que cursei durante estes anos.

Aos meus antigos professores, que me ensinaram com prazer e dedicação parte do que sei e abriram a porta do meu futuro com o estudo.

A minha companheira, Diane Rizzotto Rossetto por ter me apoiado incondicionalmente durante este longo caminho, já que sem o apoio dela não teria alcançado este grande objetivo.

A todos os meus amigos, pelo incentivo e torcida. Registro aqui as saudades dos momentos inesquecíveis e deixo um abraço imenso a todos eles.

Vou fazer menção também a dois dos meus amigos, Pablo Laura e Nelson Vidaurre, que me deram apoio e dicas durante as minhas aplicações numéricas deste trabalho.

Resumo

A finalidade deste trabalho é calcular a composição de modelos no FBST (the Full Bayesian Significance Test) descrito por Borges e Stern [6]. Nosso objetivo foi encontrar um método de aproximação numérica mais eficiente que consiga substituir o método de condensação descrita por Kaplan. Três técnicas foram comparadas: a primeira é a aproximação da convolução de Mellin usando discretização e condensação descrita por Kaplan [11], a segunda é a aproximação da convolução de Mellin usando mistura de exponenciais, descrita por Dufresne [8], para calcular a convolução de Fourier mediante a aproximação de mistura de convoluções exponenciais, usando a estrutura algébrica descrita por Hogg [10], mais a aplicação do operador descrito por Collins [7], para transformar a convolução de Fourier para a convolução de Mellin, a terceira é a aproximação da convolução de Mellin usando mistura de exponenciais, descrita por Dufresne [8], para aproximar diretamente via mistura de exponenciais a convolução de Fourier, mais a aplicação do operador descrito por Collins [7], para transformar a convolução de Fourier para a convolução de Mellin.

Palavras-chave: Aproximação Numérica, Convolução, Mistura de exponenciais.

Abstract

The purpose of this work is to calculate the compositional models of FBST (the Full Bayesian Significance Test) studied by Borges and Stern [6]. The objective of this work was to find an approximation method numerically efficient that can replace the condensation methods described by Kaplan. Three techniques were compared: First, the approximation of Mellin convolution using discretization and condensation described by Kaplan [11], second, the approximation of Mellin convolution using mixtures of exponentials, described by Dufresne [8], to calculate the Fourier convolution by approximation of mixtures of exponential convolutions, using the algebraic structure described by Hogg [10], and then to apply the operator described by Collins [7], to transform the usual convolution to Mellin convolution, third, the approximation of Mellin convolution using mixtures of exponentials, described by Dufresne [8], to calculate the Fourier convolution by direct approximation of mixtures of exponentials, and then to apply the operator described by Collins [7], to transform the usual convolution to Mellin convolution.

Keywords: Numerical Approximation, Convolution, Mixtures of exponentials.

Sumário

Lista de Símbolos	ix
Lista de Figuras	xi
Lista de Tabelas	xiii
Introdução	1
1 Teste de Significância Totalmente Bayesiano	5
1.1 O e-valor	5
1.2 O valor-verdade para conjunções	7
2 Discretização e Condensação de funções de distribuição contínua	11
2.1 Caso horizontal	11
2.1.1 Discretização horizontal	11
2.1.2 Convolução horizontal	12
2.1.3 Condensação horizontal	13
2.2 Caso vertical	14
2.2.1 Discretização vertical	14
2.2.2 Convolução vertical	15
2.2.3 Condensação vertical	16
2.3 Exemplo de convolução de Mellin	19
3 Relação entre as transformadas de Fourier e Mellin	21
3.1 A Transformada de Laplace	21
3.1.1 Propriedades	22
3.2 A Transformada de Fourier	23
3.2.1 Série de Fourier	24
3.2.2 Transformada de Fourier	31
3.2.3 Propriedades	33
3.2.4 Convolução	34
3.3 A Transformada de Mellin	36
3.3.1 Convolução de Mellin	36

3.3.2	Transformada de Mellin	37
3.4	Relação entre as convoluções e as funções de distribuição	41
3.4.1	Soma de variáveis aleatórias	41
3.4.2	Produto de variáveis aleatórias	45
4	Aproximação a distribuições via mistura de exponenciais	49
4.1	Polinômio de Jacobi	50
4.2	Misturas de exponenciais usando o polinômio de Jacobi	55
4.2.1	Usando a estrutura algébrica de Hogg para as convoluções	62
4.3	Métodos numéricos	71
4.3.1	Aproximações via mistura de exponenciais	71
4.3.2	Aproximação da convolução	77
4.4	Método direto para o cálculo das convoluções	78
5	O Teste de independência condicionada	83
5.1	O Modelo Multinomial	83
5.2	O exemplo	86
	Conclusões	93
	Referências Bibliográficas	95

Lista de Símbolos

F_T	Função de distribuição da v.a. T .
$\bar{F}_T$	Função de cauda da v.a. T .
f_T	Função de densidade da v.a. T .
F_X	Função de distribuição da v.a. X .
$\bar{F}_X$	Função de cauda da v.a. X .
f_X	Função de densidade da v.a. X .
f, g, h	Funções padrões definidas dependendo do contexto.
$f^{[c]}(x)$	Função de variável x definida como $x^{-c}f(x)$, com $c \geq 0$.
$\mathcal{L}[f]$	Transformada de Laplace da função f .
$\mathcal{F}\{f\}$	Transformada de Fourier da função f .
$\mathcal{M}[f]$	Transformada de Mellin da função f .
$\mathcal{T}$	Operador de Collins.
$f * g$	Convolução de Fourier de f e g .
$f \otimes g$	Convolução de Mellin de f e g .

Lista de Figuras

2.1	Tabela I	16
2.2	Função de distribuição discreta	17
2.3	Tabela II	18
4.1	Cauda da Uniforme	72
4.2	Distribuição da Uniforme	73
4.3	Densidade da Uniforme	73
4.4	Cauda da Pareto	74
4.5	Distribuição da Pareto	75
4.6	Densidade da Pareto	75
4.7	Cauda da Lognormal	76
4.8	Distribuição da Lognormal	76
4.9	Densidade da Lognormal	77
4.10	Normal Truncada	77
4.11	Normal Truncada - Comparações das convoluções	78
4.12	Convolução de Mellin Alg Mix of Exp vs Direct Mix of Exp	80
5.1	Distribuição da função verdade	88
5.2	Distribuição da função verdade Híbrida Horizontal	89
5.3	Distribuição da função verdade Híbrida Vertical	90
5.4	Convolução de Mellin das Distribuições da função verdade	91

Lista de Tabelas

4.1	Distribuição Uniforme	72
4.2	Distribuição de Pareto - 500 repetições	74
4.3	Distribuição Lognormal - 500 repetições	76
5.1	X=k	84
5.2	X=k	85
5.3	X=1	86
5.4	X=2	87
5.5	X=3	87

Introdução

A função verdade W , para um tipo de hipótese simples H dada, é definida por Borges e Stern [6] para o cálculo da evidência $ev(H)$ que é a região de aceitação da hipótese simples H .

Para uma hipótese composta H , Borges e Stern [6] analisaram a relação entre a hipótese composta H e suas respectivas hipóteses simples H_j , $j = 1, \dots, N$. Se a hipótese composta H for constituída por conjunções de suas respectivas hipóteses simples H_j , $j = 1, \dots, N$, então, a função verdade W , para o cálculo da evidência $ev(H)$, pode ser calculada pela suas componentes W_j , as funções verdade de cada H_j , $j = 1, \dots, N$, como a convolução de Mellin definida como

$$W = W_1 \otimes W_2 \otimes \dots \otimes W_N,$$

onde, para duas funções de densidade (pdf), a convolução de Mellin é dada por

$$(f_1 \otimes f_2)(x) = \int_0^\infty f_2\left(\frac{x}{w}\right) f_1(w) \frac{1}{w} dw \quad (1)$$

onde

$$f_1, f_2 \in \mathcal{M}_s(\mathbf{R}_+) = \left\{ f \in L^1(\mathbf{R}) / \int_0^\infty |f(x)| x^{s-1} dx < \infty, \text{ para algum } s \right\}. \quad (2)$$

O e-valor da hipótese H , $ev(H)$, é uma epistêmica Bayesiana, valor de credibilidade ou verdade, definido como o teste de significância totalmente Bayesiano, “full Bayesian significance test” (FBST).

Existem distribuições cuja convolução de Mellin não é conhecida de forma simples. Para contornar esse problema, em Andrade e Stern [3] utilizam uma técnica de aproximação onde usam discretização e condensação, descrita em Kaplan [11].

Hogg [10] investiga um método para descrever analiticamente a distribuição do tempo de vida de um sistema mediante a construção de uma estrutura algébrica. Para isto, Hogg [10] usou propriedades inerentes de convolução e mistura de distribuições do tempo de vida.

Uma vez que a estrutura algébrica é construída, Hogg [10] propõe usar esta estrutura para desenvolver algoritmos para ir desde o diagrama esquemático de um sistema até sua respectiva função de sobrevivência.

Mistura de componentes da taxa de falha de um sistema pode ser descrita por mistura de convoluções. Essas expressões podem ser simplificadas usando as propriedades da estrutura algébrica encontrada.

O modelo de mistura adotada por Hogg [10] é a de mistura de exponenciais. Isto se deve ao fato que a distribuição exponencial possui o modelo mais simples possível do tempo de vida, no sentido de ser caracterizada por uma taxa de falha homogênea no tempo.

Hogg [10], como exemplo, usou o modelo exponencial, com taxa de falha λ , mostrando uma adaptação flexível na estrutura algébrica por mistura de convoluções.

Se a função de sobrevivência for representada por mistura de exponenciais, aproveitando a estrutura algébrica junto com as boas propriedades do modelo exponencial, poderíamos calcular a convolução aproximando-a por mistura de convoluções exponenciais.

Para o cálculo da evidência, usando a convolução de Mellin, investigamos o uso de um método numérico descrito por Dufresne [8] para aproximar funções de distribuição contínua via mistura de exponenciais com o objetivo de aproximar a convolução de Fourier usando a estrutura algébrica encontrada por Hogg [10]. Collins [7] representa a convolução de Mellin mediante um operador de transformação, este operador transforma a convolução de Fourier para a convolução de Mellin e foi aplicado no método numérico que aproxima a convolução de Fourier para aproximar a convolução de Mellin.

No capítulo 1 apresentaremos conceitos básicos sobre o FBST. Definiremos a função verdade W para um tipo de hipótese simples H dada, consequentemente, veremos como calcular a região de aceitação da hipótese simples H chamada de evidência, $ev(H)$.

Para a função verdade W da hipótese composta H dada, veremos que, se a hipótese H é constituída por conjunção de suas respectivas hipóteses simples H_j , $j = 1, \dots, N$, então, a relação entre a função verdade W e suas respectivas componentes W_j , $j = 1, \dots, N$ será mediante a convolução de Mellin, ou seja, encontraremos a função verdade inteira, $W = \bigotimes_{\{1 \leq j \leq N\}} W_j$. Encontrando a função verdade inteira W , a região de aceitação da hipótese H , $ev(H)$, será calculada.

Em muitas situações, calcular convoluções pode ser uma tarefa difícil. Por esse motivo, no capítulo 2 apresentaremos um método numérico chamado de discretização e condensação, aritmética probabilística. Esse procedimento substitui operações aritméticas com números por operações aritméticas sobre variáveis aleatórias com a finalidade de calcular a convolução de Mellin de distribuições de probabilidade. Serão apresentados dois casos, Horizontal e Vertical.

Esse método consegue obter uma aproximação eficiente. Porém, a região avaliada para aproximar a distribuição é limitada pelo seu mínimo e máximo valor das simulações. Portanto, não poderemos calcular valores da distribuição fora do mínimo e máximo valor das simulações.

Conceitos de Análise de Fourier e de Mellin serão apresentados no capítulo 3. O operador de transformação, que permite representar a convolução de Mellin a partir da convolução de Fourier será tratado aqui. A convolução de Fourier representa a distribuição da soma de duas variáveis aleatórias independentes, a convolução de Mellin representa o produto de duas variáveis aleatórias independentes.

O espaço de distribuições contínuas sobre o operador convolução de Fourier define uma estrutura algébrica. Consequentemente, o operador de transformação permite adotar essa estrutura algébrica no mesmo espaço sobre o operador convolução de Mellin.

Conhecendo o operador de transformação, nosso foco foi tratar numericamente uma aproximação eficiente para a convolução de Fourier. A técnica que usamos para aproximar distribuições contínuas foi via mistura de exponenciais e foi tratada no capítulo 4. A técnica de mistura de exponenciais foi calculada a partir da expansão do polinômio de Jacobi “deslocado”, $R_k^{(\alpha, \beta)}(x)$, sob $[0, 1]$ e parâmetros $\alpha, \beta > -1$.

Mistura de exponenciais é uma técnica de aproximação para a função de sobrevivência, $\bar{F}_T(t) = 1 - F_T(t) = \mathbb{P}(\{T > t\})$, de uma variável aleatória não negativa T , com função de distribuição acumulada $F_T(t)$ e função de densidade de probabilidade $f_T(t)$.

Representando via mistura de exponenciais os elementos do espaço de distribuições contínuas conseguimos obter, usando a estrutura algébrica munido da operação convolução, uma representação da convolução das distribuições via mistura de convoluções exponenciais. Assim, conseguimos dar nossa primeira contribuição nesse trabalho. Teoricamente obtemos um resultado interessante, porém, numericamente tivemos complicações.

Pelas complicações numéricas obtidas na forma algébrica para o cálculo da convolução de Fourier, como nova contribuição, propusemos um método direto para calcular a convolução de Fourier via mistura de exponenciais. Veremos comparações numéricas entre o uso do método algébrico e o método direto.

Por último, no capítulo 5, mostramos uma aplicação fazendo um teste de independência condicionada usando o modelo multinomial. Calculamos a função verdade inteira para obter a evidência de nossa hipótese. Nessa aplicação vemos que a melhor aproximação da função verdade inteira e suas respectivas componentes foi mediante a hibridização de ambos métodos, discretização e condensação na região central, e o método direto de mistura de exponenciais na região das caudas. Os motivos serão explicados no decorrer do capítulo.

Capítulo 1

Teste de Significância Totalmente Bayesiano

O teste de significância totalmente Bayesiano, “full Bayesian significance test” (FBST), é apresentado em Borges e Stern [6] como um coerente teste de significância Bayesiano para hipóteses precisas.

O FBST é um teste intuitivo, com caracterização geométrica, e que pode ser implementado através de técnicas de otimização e integração numérica. O método é considerado **totalmente Bayesiano** por requerer apenas o conhecimento do espaço paramétrico representado pela sua função de densidade a posteriori, e consiste na análise de conjuntos de credibilidade. Uma vantagem do FBST é a de não necessitar de premissas adicionais, como probabilidade positiva para hipóteses precisas.

1.1 O e-valor

No FBST, a evidência (e-valor) para hipóteses precisas é calculada.

O e-valor é dado pelo complementar da probabilidade do conjunto de credibilidade chamado de conjunto tangente, sendo um subconjunto do espaço paramétrico, onde a função surpresa a posteriori, definida como a razão entre a densidade a posteriori e uma função de densidade de referência, de cada um de seus elementos é maior que o máximo dessa função surpresa a posteriori sobre todo o espaço da hipótese nula.

Formalmente:

Seja o espaço triplo (Ω, Σ, Θ) , onde Ω é o espaço amostral, Σ uma σ -álgebra de Ω e $\Theta \subset \mathbb{R}^N$ o espaço paramétrico. Esses conceitos podem ser vistos detalhadamente em Billingsley [5] e em Shiryaev [15].

Para $\theta \in \Theta$ e a verossimilhança $L(\theta|\eta)$ associados aos dados observados η de tamanho n , seja $p_n(\theta)$ a densidade a posteriori que é proporcional ao produto da verossimilhança e a densidade a priori $p_0(\theta)$, ou seja:

$$p_n(\theta) \propto L(\theta|\eta)p_0(\theta).$$

A hipótese nula Θ_0 , que por simplicidade será denotado por H , é definida como:

$$H = \Theta_0 = \{\theta \in \Theta / g(\theta) \leq 0 \wedge h(\theta) = 0\},$$

onde g e h são funções vetoriais que representam as restrições de desigualdade e igualdade dentro do espaço paramétrico.

Os tipos de hipóteses que estamos interessados, são chamadas de hipóteses precisas ($\dim(H) < \dim(\Theta)$), ou seja, existe pelo menos uma restrição de igualdade.

A hipótese H é precisa quando o subconjunto Θ_0 tem medida de Lebesgue nula, ver Royden [13]. Na construção do FBST todos os conjuntos de mesma natureza são tratados da mesma maneira e, como consequência disso, os conjuntos que definem hipóteses precisas têm sempre medida nula. Nesta construção, ao invés de mudar a natureza de H supondo probabilidade positiva nesse conjunto, trabalhamos com o conjunto tangente que definiremos em seguida.

O FBST define $ev(H)$, o e-valor, a evidência da região de aceitação da hipótese H , e $\bar{ev}(H)$, o e-valor de rejeição da hipótese H .

Definimos:

$$T(s) = \{\theta \in \Theta / S(\theta) > s\} \quad \text{e} \quad \bar{T}(s) = \Theta - T(s), \quad (1.1)$$

onde

$$S(\theta) = \frac{p_n(\theta)}{r(\theta)}. \quad (1.2)$$

A função $S(\theta)$ é conhecida como função surpresa a posteriori relativa a densidade de referência $r(\theta)$.

A credibilidade de $T(s)$ é definida pela sua probabilidade a posteriori como:

$$\kappa(s) = \int_{T(s)} p_n(\theta) d\theta = \int_{\Theta} p_n(\theta) \mathbb{I}_{T(s)}(\theta) d\theta$$

onde $\mathbb{I}_{T(s)}(\theta)$ é a função indicadora de $T(s)$, ou seja:

$$\mathbb{I}_{T(s)}(\theta) = \begin{cases} 1 & \text{se } \theta \in T(s) \\ 0 & \text{c.c.} \end{cases}$$

Sejam:

$$\theta^* \in \operatorname{argmax}_{\theta \in H} S(\theta) \quad \text{e} \quad s^* = S(\theta^*),$$

onde s^* o máximo da função surpresa a posteriori sob a hipótese H , θ^* o ponto de máximo, e $T^* = T(s^*)$ o espaço tangente da hipótese H e a credibilidade de T^* como sendo κ^* .

Portanto, o e-valor, sendo o complementar da probabilidade de T^* , é definido como:

$$ev(H) = 1 - \kappa^* = 1 - \int_{\Theta} p_n(\theta) \mathbb{I}_{T^*}(\theta) d\theta = \int_{\Theta} p_n(\theta) \mathbb{I}_{\bar{T}^*}(\theta) d\theta.$$

Observe que

$$\text{se } S(\theta) \leq s^*, \quad \forall \theta \in H \quad \text{então } H \subset \bar{T}^*.$$

Ou seja, se a probabilidade de T^* for grande, o espaço nulo estaria dentro da região de baixa probabilidade e a evidência rejeita a hipótese H . Mas, se a probabilidade de T^* for pequena, então o espaço nulo estaria dentro da região de alta probabilidade e a evidência aceita a hipótese H .

Definimos a função verdade ou função de acumulação surpresa, $W : \mathbb{R}_+ \rightarrow [0, 1]$ como:

$$W(s) = \int_{\bar{T}(s)} p_n(\theta) d\theta \quad \text{e} \quad \bar{W}(s) = 1 - W(s).$$

Pela equação (1.1) obtemos uma nova expressão da função verdade, desde que a função surpresa definida em (1.2) possa ser vista como uma variável aleatória que depende do parâmetro $\theta \in \Theta$. Assim

$$W(s) = \int_{\bar{T}(s)} p_n(\theta) d\theta = \int_{\Theta} \mathbb{I}_{\bar{T}(s)}(\theta) p_n(\theta) d\theta = \mathbb{P}(\{S \leq s\}),$$

e

$$ev(H) = W(s^*), \quad \bar{ev}(H) = \bar{W}(s^*) = 1 - ev(H). \quad (1.3)$$

A consistência e invariância da evidência $ev(H)$ foi analisada em Borges e Stern [6].

1.2 O valor-verdade para conjunções

Para um conjunto independente de parâmetros $(\theta_1, \dots, \theta_N) \in \Theta_1 \times \dots \times \Theta_N$, a hipótese composta H é equivalente a:

$$H : \theta_1 \in H_1 \wedge \theta_2 \in H_2 \wedge \cdots \wedge \theta_N \in H_N$$

com $H_j \subset \Theta_j$, $\forall j = 1, \dots, N$, e pode ser decomposta por hipóteses simples:

$$\begin{aligned} H_1 & : \theta_1 \in H_1 \\ H_2 & : \theta_2 \in H_2 \\ & \vdots \\ H_N & : \theta_N \in H_N. \end{aligned}$$

Para esse caso, vamos supor que H é equivalente a $H_1 \wedge H_2 \wedge \cdots \wedge H_N$.

Sendo p_n e r tais que:

$$p_n(\theta) = \prod_{j=1}^N p_n^j(\theta_j) \quad \text{e} \quad r(\theta) = \prod_{j=1}^N r_j(\theta_j), \quad \forall \theta = (\theta_1, \dots, \theta_N) \in \Theta,$$

temos que

$$S(\theta) = \prod_{j=1}^N S_j(\theta_j), \quad \forall \theta = (\theta_1, \dots, \theta_N) \in \Theta.$$

Sendo assim, temos o seguinte resultado

Resultado 1 *Se H é equivalente a $H_1 \wedge H_2 \wedge \cdots \wedge H_N$, então*

$$s^* = \sup_{\theta \in H} S(\theta) = \prod_{j=1}^N \sup_{\theta_j \in H_j} S_j(\theta_j) = \prod_{j=1}^N s_j^*.$$

Assim, chamando F_S como a distribuição de S , temos

$$\begin{aligned} W(s) &= \mathbb{P}(\{S \leq s\}) = F_S(s) = F_{\prod_{j=1}^N S_j}(s) \\ &= F_{S_1} \otimes F_{S_2} \otimes \cdots \otimes F_{S_N}(s) \\ &= W_1 \otimes W_2 \otimes \cdots \otimes W_N(s) \end{aligned}$$

onde

$$W_1 \otimes W_2(s) = \int_0^s W_1\left(\frac{s}{w}\right) W_2(dw)$$

que é chamada de “convolução de Mellin” que representa o produto de variáveis aleatórias independentes. Portanto

$$ev(H) = W_1 \otimes W_2 \otimes \cdots \otimes W_N \left(\prod_{j=1}^N s_j^* \right). \quad (1.4)$$

Com isto, no próximo capítulo, veremos um método numérico, usado no artigo de Andrade e Stern [3], para calcular a convolução de Mellin definida acima. Esse método numérico foi comparado com um novo método numérico desenvolvido nesse trabalho e que iremos mencionar nos próximos capítulos. O objetivo foi calcular a convolução de Mellin de maneira mais eficiente.

Capítulo 2

Discretização e Condensação de funções de distribuição contínua

Andrade e Stern [3] trabalharam com um método numérico chamado de discretização e condensação, baseado nos resultados em aritmética probabilística de Williamson [16, 17]. Esse procedimento substitui operações aritméticas com números por operações aritméticas sobre variáveis aleatórias. O objetivo é calcular a convolução de Mellin de distribuições de probabilidade. Apresentamos dois casos.

2.1 Caso horizontal

2.1.1 Discretização horizontal

O artifício de aproximar uma variável aleatória contínua X por uma variável aleatória discreta Y , é feito da seguinte maneira:

1. Dividimos os valores de X em classes (intervalos 2 a 2 disjuntos) da seguinte forma: para $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$ com $-\infty \leq a < b \leq \infty$. As classes são definidas como $I_1 = [x_0, x_1)$, $I_2 = [x_1, x_2)$, $\dots$, $I_{n-1} = [x_{n-2}, x_{n-1})$, $I_n = [x_{n-1}, x_n]$.
2. O representante de cada classe $\{I_i\}$, é dado por $y_i = \frac{x_{i-1} + x_i}{2}$, $\forall i = 1, \dots, n$.
3. Sendo $\Delta x_i = x_i - x_{i-1}$, então $\mathbb{P}(\{Y = y_i\}) = f_X(y_i)\Delta x_i$, $\forall i = 1, \dots, n$, onde f_X representa a função de densidade de X .

A ideia segue diretamente da construção do histograma em estatística descritiva, onde $\{I_i\}$ representa o intervalo de classe, Δx_i representa a amplitude, $f_i = \mathbb{P}(\{Y = y_i\})$ representa a frequência relativa e y_i seria o representante da classe I_i , $\forall i = 1, \dots, n$. As classes também são conhecidas como bins.

O seguinte pseudocódigo pode ser usado para implementar esse procedimento.

Algoritmo 1 Subrotina discretização horizontal

```

1: Dado:  $na$       % tamanho da amostra da v.a. contínua  $X$ .
2: Dado:  $nb$       % número de bins ou classes.
3: Dados:  $x(1) < x(2) < \dots < x(na - 1) < x(na)$       % amostra ordenada.
4:  $\Delta x = \frac{x(na) - x(1)}{nb}$ ,  $breaksh(0) = x(1)$ ,
5: Para  $i = 1 : nb$  fazer:
6:      $breaksh(i) = breaksh(i - 1) + \Delta x$ 
7:      $xh(i) = \frac{breaksh(i-1) + breaksh(i)}{2}$ 
8:      $fh(i) = 0$ 
9:     Para  $j = 1 : na$  fazer:
10:        Se  $\{breaksh(i - 1) \leq x(j) < breaksh(i)\}$  fazer:
11:             $fh(i) = fh(i) + \frac{1}{na}$ 
12:        Fim
13:    Fim
14: Fim
15:  $Fh(1) = fh(1)$ 
16: Para  $i = 2 : nb$  fazer:
17:      $Fh(i) = fh(i) + Fh(i - 1)$ 
18: Fim
19: Saidas:  $breaksh$ ,  $xh$ ,  $fh$ ,  $Fh$ .

```

2.1.2 Convolução horizontal

A convolução de Mellin que representa o produto de duas variáveis aleatórias independentes $X_1 X_2$ é dada por:

$$F_{X_1 X_2}(x) = \int_0^x F_{X_1}\left(\frac{x}{w}\right) dF_{X_2}(w). \quad (2.1)$$

Usaremos f_{X_1}, f_{X_2} , discretizados pelo procedimento anterior, como dados de entrada para desenvolver um método para o cálculo da convolução horizontal:

Algoritmo 2 Subrotina convolução de Mellin horizontal

```

1: Dados:  $xh1, xh2, fh1, fh2$       % discretizados.
2: Dado:  $n$       % número de bins.
3:  $m = n * n$ 
4:  $F = f = x = vetor(0, tamanho = m)$ 
5: Para  $i = 1 : n$  fazer:
6:     Para  $j = 1 : n$  fazer:
7:          $x((i - 1) * n + j) = xh1(i) * xh2(j)$ 
8:          $f((i - 1) * n + j) = fh1(i) * fh2(j)$ 
9:     Fim
10: Fim
11: Seja  $xh$  a ordenação de menor a maior valor de  $x$ 
12: Seja  $fh$  a ordenação de  $f$  com respeito de  $xh$ 
13:  $Fh(1) = fh(1)$ 
14: Para  $i = 2 : m$  fazer:
15:      $Fh(i) = fh(i) + Fh(i - 1)$ 
16: Fim
17: Saidas:  $xh, fh, Fh$ .

```

2.1.3 Condensação horizontal

A convolução numérica de duas distribuições com n bins retorna uma distribuição com n^2 bins, o que poderia ser um problema para uma sequência de operações. Para contornar isto, os autores apresentaram um método simples que reduz o retorno de n^2 bins para n bins, sem introduzir erro no resultado.

Esta operação é chamada de “condensação horizontal” e retorna os limites superior e inferior de cada um dos n bins da distribuição obtida da convolução.

Segue o pseudocódigo para este procedimento:

Algoritmo 3 Subrotina condensação horizontal

```

1: Dados:  $xh$   $Fh$       % discretizados.
2: Dado:  $m$       % tamanho dos dados.
3:  $n = \sqrt{m}$ 
4:  $Fhl = Fhu = xhl = xhu = \text{vetor}(0, \text{tamanho} = n)$ 
5: Para  $i = 1 : n$  fazer:
6:    $Fhl(i) = Fh((i - 1) * n + 1)$ 
7:    $Fhu(i) = Fh(i * n)$ 
8:    $xhl(i) = xh((i - 1) * n + 1)$ 
9:    $xhu(i) = xh(i * n)$ 
10:   $xhm(i) = \frac{xhu(i) + xhl(i)}{2}$ 
11:   $Fhm(i) = \left( \frac{xhm(i) - xhl(i)}{xhu(i) - xhl(i)} \right) * (Fhu(i) - Fhl(i)) + Fhl(i)$ 
12: Fim
13:  $fhm(1) = Fhm(1)$ 
14: Para  $i = 2 : n$  fazer:
15:    $fhm(i) = Fhm(i) - Fhm(i - 1)$ 
16: Fim
17: Saidas:  $xhl, xhu, xhm, Fhl, Fhu, Fhm, fhm$ .

```

2.2 Caso vertical**2.2.1 Discretização vertical**

A aproximação da variável aleatória contínua X por uma variável aleatória discreta Y pode ser construída da seguinte maneira:

1. Dividimos os valores de $U = F_X(X)$ em classes. Para $0 = u_0 < u_1 < \dots < u_{n-1} < u_n = 1$, as classes de X são definidas como $I_1 = F_X^{-1}([u_0, u_1))$, $I_2 = F_X^{-1}([u_1, u_2))$, $\dots$, $I_{n-1} = F_X^{-1}([u_{n-2}, u_{n-1}))$, $I_n = F_X^{-1}([u_{n-1}, u_n])$.
2. O representante de cada classe $\{I_i\}$, é dado por $y_i = \frac{F_X^{-1}(u_{i-1}) + F_X^{-1}(u_i)}{2}$, $\forall i = 1, \dots, n$.
3. Sendo $\Delta u_i = u_i - u_{i-1}$, então $\mathbb{P}(\{Y = y_i\}) = \Delta u_i$, $\forall i = 1, \dots, n$.

Para n conhecido como o número de bins.

A nossa seleção do intervalo, como sendo o bins, é dada no eixo vertical da função de distribuição acumulada distribuída no intervalo $[0, 1]$. Com isso, podemos nos focar em algum tipo de região de interesse, por exemplo, se nosso interesse está no comportamento da cauda da distribuição, o tamanho dos bins pode ser reduzido nessa região. Consequentemente, aumentamos o número de bins necessário para representar a cauda da distribuição.

O próximo pseudocódigo resume este procedimento:

Algoritmo 4 Subrotina discretização vertical

```

1: Dado:  $na$            % tamanho da amostra da v.a. contínua  $X$ .
2: Dado:  $nb$            % número de bins ou classes.
3: Dados:  $x(1) < x(2) < \dots < x(na - 1) < x(na)$       % amostra ordenada.
4:  $c = \frac{na}{nb}$ ,  $\Delta u = \frac{1}{nb}$ ,  $breaksv(0) = x(1)$ ,
5: Para  $i = 1 : (nb - 1)$  fazer:
6:      $breaksv(i) = x(i * c + 1)$ 
7: Fim
8:  $breaksv(nb) = x(na)$ 
9: Para  $i = 1 : nb$  fazer:
10:     $xv(i) = \frac{breaksv(i-1) + breaksv(i)}{2}$ 
11:     $fv(i) = \Delta u$ 
12:     $Fv(i) = i * \Delta u$ 
13: Fim
14: Saidas:  $breaksv$ ,  $xv$ ,  $fv$ ,  $Fv$ .

```

2.2.2 Convolução vertical

Dadas as v.a's contínuas X_1, X_2 , usaremos f_{X_1}, f_{X_2} discretizados pelo procedimento anterior como dados de entrada para desenvolver um método para o cálculo da convolução vertical.

O pseudocódigo abaixo sintetiza o método para determinar a convolução para o caso vertical.

Algoritmo 5 Subrotina convolução de Mellin vertical

```

1: Dados:  $xv1, xv2, fv1, fv2, breaksv1, breaksv2$       % discretizados.
2: Dado:  $n$       % número de bins.
3:  $m = n * n$ 
4:  $F = f = x = \text{vetor}(0, \text{tamanho} = m)$ 
5:  $breaks = \text{vetor}(0, \text{tamanho} = m + 1)$ 
6: Para  $i = 1 : n$  fazer:
7:     Para  $j = 1 : n$  fazer:
8:          $x((i - 1) * n + j) = xv1(i) * xv2(j)$ 
9:          $f((i - 1) * n + j) = fv1(i) * fv2(j)$ 
10:         $breaks((i - 1) * n + j - 1) = breaksv1(i - 1) * breaksv2(j - 1)$ 
11:    Fim
12: Fim
13:  $breaks(m) = breaksv1(n) * breaksv2(n)$ 
14: Seja  $breaksv$  a ordenação de menor a maior valor de  $breaks$ 
15: Sejam  $xv, fv$  a ordenação de  $x$  e  $f$  com respeito de  $breaksv$ 
16:  $Fv(1) = fv(1)$ 
17: Para  $i = 2 : m$  fazer:
18:      $Fv(i) = fv(i) + Fv(i - 1)$ 
19: Fim
20: Saídas:  $xv, fv, Fv, breaksv$ .
```

2.2.3 Condensação vertical

Ilustramos este procedimento através de um exemplo apresentado por Kaplan [11]. As figuras exibidas foram extraídas neste artigo.

Seja $\{(x_i, p_i)\}_{i=1}^{15}$, onde $x_1, \dots, x_{15}$ são os dados da amostra e $p_1, \dots, p_{15}$ as suas respectivas massas de probabilidade. Queremos condensar os dados de tamanho 15 para tamanho 6. Mostramos os 15 dados abaixo

x_i	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0	
p_i	.02	.03	.05	.06	.10	.09	.10	.05	.20	.05	.10	.05	.06	.02	.02	
P_i	0	.02	.05	.10	.16	.26	.35	.45	.50	.70	.75	.85	.90	.96	.98	1.0

Figura 2.1: Tabela I

Agora, vamos definir o eixo vertical (P) em 6 classes de intervalos, ou bins vertical,

$$\{[0, 0.1), [0.1, 0.2), [0.2, 0.4), [0.4, 0.6), [0.6, 0.8), [0.8, 1]\}.$$

Os 2 primeiros bins possuem 10% de probabilidade, o restante possui 20% de probabilidade.

Figura 2.2: Função de distribuição discreta

Para obter os novos dados condensados $\{(\tilde{x}_i, \tilde{p}_i)\}_{i=1}^6$ procedemos da seguinte maneira:

Podemos observar que

$$\begin{aligned} \tilde{p}_1 &= p_1 + p_2 + p_3 = 0.1 \quad \text{e} \\ \tilde{x}_1 &= \frac{1}{\tilde{p}_1} (p_1 x_1 + p_2 x_2 + p_3 x_3) = \frac{1}{0.1} (0.23) = 2.3, \end{aligned} \tag{2.2}$$

o par (x_4, p_4) consegue encaixar dentro do segundo bins vertical, mas (x_5, p_5) deve ser dividido para colocá-lo no segundo e terceiro bins vertical, segundo o raio:

$$R = \frac{0.2 - P_4}{P_5 - P_4} = \frac{0.2 - 0.16}{0.26 - 0.16} = 0.4.$$

Assim

$$\begin{aligned}
\tilde{p}_2 &= p_4 + 0.4p_5 = 0.06 + 0.4 \cdot 0.1 = 0.1 \\
\tilde{x}_2 &= \frac{1}{\tilde{p}_2} (p_4x_4 + 0.4p_5x_5) \\
&= \frac{1}{0.1} (0.06 \cdot 4 + 0.4 \cdot 0.1 \cdot 5) = 4.4.
\end{aligned}$$

Novamente

$$R = \frac{0.4 - P_6}{P_7 - P_6} = \frac{0.4 - 0.35}{0.45 - 0.35} = 0.5.$$

Logo

$$\begin{aligned}
\tilde{p}_3 &= 0.6p_5 + p_6 + 0.5p_7 = 0.6 \cdot 0.1 + 0.09 + 0.5 \cdot 0.1 = 0.2 \\
\tilde{x}_3 &= \frac{1}{\tilde{p}_3} (0.6p_5x_5 + p_6x_6 + 0.5p_7x_7) \\
&= \frac{1}{0.2} (0.6 \cdot 0.1 \cdot 5 + 0.09 \cdot 6 + 0.5 \cdot 0.1 \cdot 7) = 5.95.
\end{aligned}$$

De forma similar obtemos

$$\begin{aligned}
\tilde{x}_4 &= \frac{1}{0.2} (0.5 \cdot 0.1 \cdot 7 + 0.05 \cdot 8 + 0.5 \cdot 0.2 \cdot 9) = 8.25 \\
\tilde{x}_5 &= \frac{1}{0.2} (0.5 \cdot 0.2 \cdot 9 + 0.05 \cdot 10 + 0.5 \cdot 0.1 \cdot 11) = 9.75 \\
\tilde{x}_6 &= \frac{1}{0.2} (0.5 \cdot 0.1 \cdot 11 + 0.05 \cdot 12 + 0.06 \cdot 13 + 0.02 \cdot 14 + 0.02 \cdot 15) = 12.55.
\end{aligned}$$

Então

$\tilde{x}_j$	2.3	4.4	5.95	8.25	9.75	12.55	
$\tilde{p}_j$	0.1	0.1	.20	.20	.20	.20	

Figura 2.3: Tabela II

Observe que a média de ambos pares é igual a $\bar{x} = 7.97$. Portanto, esse tipo de processo tem a virtude de “preservar a média”. Escolhendo os bins verticais de forma apropriada, sempre é possível manter as propriedades de interesse.

Mostramos o seguinte pseudocódigo para um tamanho uniforme dos bins verticais.

Algoritmo 6 Subrotina condensação vertical

```

1: Dados:  $breaksv, Fv, fv$       % discretizados.
2: Dado:  $m$       % tamanho dos dados.
3:  $n = \sqrt{m}$ 
4:  $w = [ \ ]$   $breakstil = vetor(0, tamanho = n + 1)$ 
5: Para  $i = 1 : n$  fazer:
6:      $v = find(Fv \geq \frac{i}{n})$ 
7:      $w = [w \ v(1)]$ 
8:      $\tilde{F}(i) = \frac{i}{n}$ 
9: Fim
10:  $breakstil(0) = breaksv(0)$ 
11:  $last = 1$ 
12: Para  $i = 1 : n$  fazer:
13:      $R = \frac{\frac{i}{n} - Fv(w(i)-1)}{Fv(w(i)) - Fv(w(i)-1)}$ 
14:      $breakstil(i) = n * (sum(fv(last : (w(i) - 1)) * breaksv((last - 1) : (w(i) - 2)))$ 
15:      $+ R * fv(w(i)) * breaksv(w(i) - 1))$ 
16:      $fv(w(i) - 1) = fv(w(i) - 1) + R * fv(w(i))$ 
17:      $fv(w(i)) = (1 - R) * fv(w(i))$ 
18:      $\tilde{x}(i) = \frac{breakstil(i) + breakstil(i-1)}{2}$ 
19:      $last = w(i)$ 
20: Fim
21: Saidas:  $breakstil, \tilde{F}, \tilde{x}$ .
```

2.3 Exemplo de convolução de Mellin

Um exemplo de convolução de Mellin para encontrar o produto de duas variáveis aleatórias X_1 e X_2 , (seção 3.2 de Andrade e Stern [3]), ambas com distribuição Log-normal, é apresentado a seguir.

Suponha X_1 e X_2 duas variáveis aleatórias contínuas tais que

$$X_1 \sim \ln\mathcal{N}(\mu_1, \sigma_1^2) \quad \text{e} \quad X_2 \sim \ln\mathcal{N}(\mu_2, \sigma_2^2)$$

e

$$F_{X_1}(x) = \int_{-\infty}^x f_{X_1}(w)dw \quad \text{e} \quad F_{X_2}(x) = \int_{-\infty}^x f_{X_2}(w)dw.$$

Essas distribuições podem ser escritas como

$$\log(X_1) = T_1 \sim \mathcal{N}(\mu_1, \sigma_1^2) \quad \text{e} \quad \log(X_2) = T_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

então

$$\log(X_1 X_2) = \log(X_1) + \log(X_2) = T_1 + T_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

que equivale a

$$X_1 X_2 \sim \ln \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

com função de distribuição definida por

$$F_{X_1 X_2}(x) = \int_{-\infty}^x f_{X_1 X_2}(w) dw.$$

Como acabamos de ver, encontramos uma relação entre a distribuição do produto de variáveis aleatórias independentes da Log-normal, associada a convolução de Mellin, e a distribuição da soma de variáveis aleatórias independentes da normal, associada a convolução de Fourier. No próximo capítulo, veremos que essa relação entre ambas convoluções pode ser generalizada através de um operador, descrito por Collins [7], que transforma a convolução de Fourier para a convolução de Mellin.

Portanto, podemos utilizar métodos numéricos que aproximem a convolução de Fourier e depois usamos o operador de Collins [7] para aproximar a convolução de Mellin através da aproximação da convolução de Fourier.

Capítulo 3

Relação entre as transformadas de Fourier e Mellin

Com o propósito de inserir um novo método numérico para o cálculo da convolução de Mellin das funções verdade definidas no capítulo 1, procurarmos alguma relação entre a convolução de Fourier e a convolução de Mellin. Collins [7] representa a convolução de Mellin mediante um operador de transformação, este operador transforma a convolução de Fourier para a convolução de Mellin. Assim, nosso objetivo de aproximar a convolução de Mellin foi via aproximação a convolução de Fourier.

Introduzimos alguns conceitos prévios de transformada de Laplace, análise de Fourier, tais como séries de Fourier, transformada de Fourier e convolução. Também introduziremos os conceitos de convolução de Mellin e transformada de Mellin. Maiores detalhes podem ser encontrados em Almira [1], Andrade [2], Gonzalez [9], Royden [13], Rodney [12] e Saenz [14].

3.1 A Transformada de Laplace

As definições e resultados podem ser vistas em Andrade [2].

Seja $f : [0, \infty) \rightarrow \mathbb{R}$ e considere

$$\int_0^{\infty} e^{-st} f(t) dt, \quad (3.1)$$

onde s é uma variável real. Quando f é suficientemente bem comportada, que será feito preciso mais adiante, esta integral convergirá para certos valores de s , definindo uma função de s , chamada de Transformada de Laplace de f , e será representada por $\mathcal{L}[f]$ ou $\mathcal{L}[f](s)$.

Observe que para (3.1) existir, uma possível condição é exigir que f seja dominada por alguma exponencial, tal como $e^{-st} f(t)$ que tende a zero quando t cresce. Mais precisamente, vamos introduzir o seguinte conceito.

Definição 3.1.1 Dizemos que f é de ordem exponencial em $[0, \infty)$, se existem constantes $C > 0$ e α tais que

$$|f(t)| \leq Ce^{\alpha t}, \quad \forall t > 0. \quad (3.2)$$

São exemplos de funções de ordem exponencial: $f(t) = C$ (constante) e

$$t^n, e^{at}, \sin(bt), \cos(at), e^{at}t^n \sin(bt), e^{at}t^n \cos(bt).$$

Como consequência também são funções de ordem exponencial, os polinômios e os polinômios trigonométricos.

Para simplificar a notação, uma função de ordem exponencial satisfazendo (3.2) será chamada de função de ordem exponencial α .

Vamos apresentar uma condição suficiente para a existência de $\mathcal{L}$.

Teorema 3.1.1 *Se f é contínua por partes e de ordem exponencial em $[0, \infty)$, então existe um número real α tal que*

$$\int_0^\infty e^{-st} f(t) dt,$$

converge para todos os valores de $s > \alpha$.

Demonstração: Como f é de ordem exponencial, existem $C > 0$ e α reais tais que

$$|f(t)| \leq Ce^{\alpha t}.$$

Logo, temos que

$$\begin{aligned} \left| \int_0^\infty e^{-st} f(t) dt \right| &\leq C \int_0^\infty e^{-(s-\alpha)t} dt \\ &= \lim_{n \rightarrow \infty} \frac{C}{s-\alpha} \left[1 - e^{-(s-\alpha)n} \right] \\ &= \frac{C}{s-\alpha}, \quad s > \alpha. \end{aligned}$$

□

A transformada de Laplace de toda função de ordem exponencial em $[0, \infty)$ existe, mas a recíproca não é verdadeira. Existem funções que não são de ordem exponencial mas suas transformadas de Laplace existem.

Exemplo de funções mencionadas acima: $f(t) = \frac{1}{\sqrt{t}}$ cuja transformada de Laplace é $\frac{1}{2} \frac{\sqrt{\pi}}{s^{3/2}}$, e também e^{t^2} que não é de ordem exponencial, mas existe sua transformada de Laplace.

3.1.1 Propriedades

Vamos definir χ , o conjunto de todas as funções contínuas por partes e de ordem exponencial em $[0, \infty)$. Note que χ é um espaço vetorial real.

Também vamos definir ζ sendo o conjunto de todas as funções reais definidas em intervalos da forma (a, ∞) ou $[a, \infty)$, $a \geq -\infty$.

Para qualquer par de funções de ζ a soma será definida como a função cujo domínio é a intersecção dos domínios de cada função e essa função será considerada como um elemento de ζ .

Podemos considerar que duas funções de ζ são idênticas quando elas coincidem em algum intervalo da forma (a, ∞) . Desta maneira segue que $\mathcal{L}$ é um operador linear entre χ e ζ .

Vamos resumir este comentário com o seguinte resultado.

Teorema 3.1.2 (*Linearidade $\mathcal{L}$*). *Sejam $f, g \in \chi$ e $K \in \mathbb{R}$. Então, $\mathcal{L}[Kf + g](s) = K\mathcal{L}[f](s) + \mathcal{L}[g](s)$.*

Agora vamos apresentar um resultado importante na teoria de transformada de Laplace, pois garante que $\mathcal{L}$ é injetora.

Teorema 3.1.3 (*Teorema de Lerch*). *Sejam $f, g \in \chi$. Suponha que existe $s_0 \in \mathbb{R}$ tal que $\mathcal{L}[f](s) = \mathcal{L}[g](s)$, $\forall s > s_0$. Então $f(t) = g(t)$, $\forall t > 0$, exceto possivelmente nos pontos de descontinuidade de f e g .*

As próximas propriedades serão enunciadas brevemente.

1. **Comportamento assintótico de $\mathcal{L}[f]$:** Se $f \in \chi$. Então, $\lim_{s \rightarrow \infty} \mathcal{L}[f](s) = 0$.
2. **Teorema do deslocamento na variável s :** Seja $f \in \chi$. Então, $\mathcal{L}[e^{at}f](s) = \mathcal{L}[f](s - a)$.
3. **Mudança de Escala:** Seja $f \in \chi$ e $a \neq 0$. Então, $\mathcal{L}[f(at)](s) = \frac{1}{a}\mathcal{L}[f]\left(\frac{s}{a}\right)$.
4. **Transformada de Laplace da derivada:** Se $f' \in \chi$. Então, $\mathcal{L}[f'](s) = s\mathcal{L}[f](s) - f(0^+)$.

O operador que retorna $f(t)$ a partir de $\mathcal{L}[f(t)](s)$ é a “transformada inversa de Laplace”. Denotamos

$$f(t) = \mathcal{L}^{-1}[F(s)](t) \quad \text{ou} \quad f(t) = \mathcal{L}^{-1}[\mathcal{L}[f(t)](s)](t).$$

Observe que o Teorema de Lerch, afirma que a transformação linear $\mathcal{L}[f(t)](s)$ é injetora. Assim, se denotamos por $\mathcal{L}(\chi)$ a sua imagem temos que

$$\begin{aligned} \mathcal{L} &: \chi \rightarrow \mathcal{L}(\chi) \\ f &\rightarrow \mathcal{L}(f) \end{aligned}$$

é inversível.

Teorema 3.1.4 *Se $F = \mathcal{L}[f]$ e $G = \mathcal{L}[g]$ então, $\mathcal{L}^{-1}[KF + G](t) = Kf + g$.*

3.2 A Transformada de Fourier

Para estudar a transformada de Fourier vamos definir série de Fourier. Os detalhes podem ser vistas em Royden [13] e em Saenz [14].

3.2.1 Série de Fourier

Seja

$$L^2 = \left\{ f \text{ mensurável} / \int |f(t)|^2 dt < \infty \right\}.$$

Sabemos que L^2 é um espaço de Hilbert separável (possui base ortonormal enumerável completa) com produto interno definido por

$$\langle f, g \rangle = \int f(t) \overline{g(t)} dt.$$

Seja $\{f_k\}_{k \in \mathbb{N}}$ um conjunto ortonormal de L^2 , se $f \in SPAN\{f_1, \dots, f_n\}$ quando n é fixo, então

$$f = \sum_{k=1}^n \alpha_k f_k.$$

Assim, fazendo o produto interno com um dos f_j , $j = 1, \dots, n$ temos

$$\begin{aligned} \langle f, f_j \rangle &= \left\langle \sum_{k=1}^n \alpha_k f_k, f_j \right\rangle \\ &= \sum_{k=1}^n \alpha_k \langle f_k, f_j \rangle \\ &= \alpha_j \end{aligned}$$

obtemos

$$f = \sum_{k=1}^n \langle f, f_k \rangle f_k. \quad (3.3)$$

De forma mais geral, consideremos qualquer $f \in L^2$ e $G_n = SPAN\{f_1, \dots, f_n\}$, $n \in \mathbb{N}$, definimos $g_n \in G_n$ como sendo

$$g_n = \sum_{k=1}^n \langle f, f_k \rangle f_k. \quad (3.4)$$

Por (3.3) temos

$$g_n = \sum_{k=1}^n \langle g_n, f_k \rangle f_k, \quad \forall n \in \mathbb{N}. \quad (3.5)$$

Definimos $u_n = f - g_n$, assim por (3.4) e (3.5) temos

$$\begin{aligned} \sum_{k=1}^n \langle u_n, f_k \rangle f_k &= \sum_{k=1}^n \langle f - g_n, f_k \rangle f_k \\ &= \sum_{k=1}^n \langle f, f_k \rangle f_k - \sum_{k=1}^n \langle g_n, f_k \rangle f_k \\ &= 0 \end{aligned}$$

por serem $f_1, \dots, f_n$ L.I., então $\langle u_n, f_k \rangle = 0, \forall k = 1, \dots, n$.

Também

$$\begin{aligned} \|g_n\|^2 &= \left\langle \sum_{k=1}^n \langle f, f_k \rangle f_k, \sum_{j=1}^n \langle f, f_j \rangle f_j \right\rangle \\ &= \sum_{k=1}^n |\langle f, f_k \rangle|^2 \end{aligned} \quad (3.6)$$

e

$$\begin{aligned} \langle u_n, g_n \rangle &= \langle f - g_n, g_n \rangle \\ &= \langle f, g_n \rangle - \langle g_n, g_n \rangle \\ &= \left\langle f, \sum_{k=1}^n \langle f, f_k \rangle f_k \right\rangle - \|g_n\|^2 \\ &= \sum_{k=1}^n |\langle f, f_k \rangle|^2 - \|g_n\|^2 \\ &= 0 \end{aligned}$$

portanto $u_n \perp f_k, \forall k = 1, \dots, n$ e $u_n \perp g_n, \forall n \in \mathbb{N}$.

Pela relação pitagórica $\|f\|^2 = \|g_n\|^2 + \|u_n\|^2$, então

$$\begin{aligned} \|u_n\|^2 &= \|f\|^2 - \|g_n\|^2 \\ &= \|f\|^2 - \sum_{k=1}^n |\langle f, f_k \rangle|^2 \geq 0 \end{aligned}$$

ou seja

$$\sum_{k=1}^n |\langle f, f_k \rangle|^2 \leq \|f\|^2, \quad \forall n \in \mathbb{N}.$$

Assim quando $n \rightarrow \infty$, obtemos a desigualdade de Bessel

$$\sum_{k=1}^{\infty} |\langle f, f_k \rangle|^2 \leq \|f\|^2, \quad \forall n \in \mathbb{N}$$

$\{\langle f, f_k \rangle\}_{k \in \mathbb{N}}$ são conhecidos como **coeficientes de Fourier**.

Se $\{f_k\}_{k \in \mathbb{N}}$ for um conjunto ortonormal completo, então $\{f_k\}_{k \in \mathbb{N}}$ é uma base ortonormal de L^2 .

Por ser L^2 espaço de Hilbert e $G_n \subset L^2$, $\forall n \in \mathbb{N}$, existe $g \in L^2$ tal que $g_n \rightarrow g$, quando $n \rightarrow \infty$, (convergência em L^2), e pela continuidade do produto interno

$$\begin{aligned} \langle g, f_k \rangle &= \left\langle \lim_{n \rightarrow \infty} g_n, f_k \right\rangle \\ &= \lim_{n \rightarrow \infty} \langle g_n, f_k \rangle \\ &= \lim_{n \rightarrow \infty} \langle f, f_k \rangle \\ &= \langle f, f_k \rangle. \end{aligned}$$

Assim $\langle f - g, f_k \rangle = 0$, $\forall k \in \mathbb{N}$. Por ser $\{f_k\}_{k \in \mathbb{N}}$ completa então $f - g = 0$ ou $f = g$, portanto

$$f = \sum_{k=1}^{\infty} \langle f, f_k \rangle f_k \quad \text{e} \quad \|f\|^2 = \sum_{k=1}^{\infty} |\langle f, f_k \rangle|^2. \quad (3.7)$$

Sejam $f, g \in L^2$, denotemos por $\hat{f}_k = \langle f, f_k \rangle$ e $\hat{g}_k = \langle g, f_k \rangle$, $\forall k \in \mathbb{N}$, então por (3.7) temos

$$f = \sum_{k=1}^{\infty} \hat{f}_k f_k \quad \text{e} \quad g = \sum_{k=1}^{\infty} \hat{g}_k f_k$$

também

$$\langle f, g \rangle = \left\langle \sum_{k=1}^{\infty} \hat{f}_k f_k, \sum_{j=1}^{\infty} \hat{g}_j f_j \right\rangle = \sum_{k=1}^{\infty} \hat{f}_k \overline{\hat{g}_k} = \left\langle (\hat{f}_k)_{k \in \mathbb{N}}, (\hat{g}_k)_{k \in \mathbb{N}} \right\rangle_{l_2(\mathbb{N})} \quad (3.8)$$

e

$$\|f\|^2 = \sum_{k=1}^{\infty} |\hat{f}_k|^2 = \left\| \left(\hat{f}_k \right)_{k \in \mathbb{N}} \right\|_{l_2(\mathbb{N})}^2 \quad (3.9)$$

onde

$$l_2(\mathbb{N}) = \left\{ (x_k)_{k \in \mathbb{N}} / \sum_{k=1}^{\infty} |x_k|^2 < \infty \right\}.$$

Sabemos que $l_2(\mathbb{N})$ é um espaço de Hilbert separável para o produto interno dado em (3.8) e $(\hat{f}_k)_{k \in \mathbb{N}} \in l_2(\mathbb{N})$.

Portanto, depois de obter (3.7), concluímos que existe uma relação um a um entre f e $(\hat{f}_k)_{k \in \mathbb{N}}$, e de (3.8) e (3.9) obtemos o seguinte teorema.

Teorema 3.2.1 *Seja $\mathbb{K}$ um conjunto enumerável de índices, se $\{u_k\}_{k \in \mathbb{K}}$ é uma base ortonormal do espaço de Hilbert $\mathcal{H}$, então a aplicação $x \longrightarrow (\hat{x}_k)_{k \in \mathbb{K}}$ com $\hat{x}_k = \langle x, u_k \rangle$, $\forall k \in \mathbb{K}$, é um isomorfismo entre $\mathcal{H}$ e $l_2(\mathbb{K})$.*

Isto estabelece que todos os espaços de Hilbert separáveis são isomorfos.

Assim, para $a, b \in \mathbb{R}$ com $a < b$ temos

$$L^2([a, b]) \cong l_2(\mathbb{Z}) \cong l_2(\mathbb{N}).$$

O seguinte conjunto constitui uma base ortonormal para $L^2([a, b])$

$$\left\{ \frac{1}{\sqrt{b-a}}, \frac{\sqrt{2}}{\sqrt{b-a}} \cos \left(k\pi \left[\frac{2t-a-b}{b-a} \right] \right), \frac{\sqrt{2}}{\sqrt{b-a}} \sin \left(k\pi \left[\frac{2t-a-b}{b-a} \right] \right) \right\}_{k \in \mathbb{N}}$$

com produto interno

$$\langle f, g \rangle = \int_a^b f(t) \overline{g(t)} dt \quad \forall f, g \in L^2([a, b]).$$

Para $a = -L$ e $b = L$, temos a base ortonormal de $L^2([-L, L])$ dada por

$$\left\{ \frac{1}{\sqrt{2L}}, \frac{1}{\sqrt{L}} \cos \left(k \frac{\pi}{L} t \right), \frac{1}{\sqrt{L}} \sin \left(k \frac{\pi}{L} t \right) \right\}_{k \in \mathbb{N}}$$

com produto interno

$$\langle f, g \rangle = \int_{-L}^L f(t) \overline{g(t)} dt \quad \forall f, g \in L^2([-L, L]).$$

De (3.7) temos que para cada $f \in L^2([-L, L])$

$$f(t) = A_0 \left(\frac{1}{\sqrt{2L}} \right) + \sum_{k=1}^{\infty} \left\{ A_k \left[\frac{1}{\sqrt{L}} \cos \left(k \frac{\pi}{L} t \right) \right] + B_k \left[\frac{1}{\sqrt{L}} \sin \left(k \frac{\pi}{L} t \right) \right] \right\} \quad (3.10)$$

com

$$\begin{aligned} A_0 &= \left\langle f, \frac{1}{\sqrt{2L}} \right\rangle = \frac{1}{\sqrt{2L}} \int_{-L}^L f(t) dt \\ A_k &= \left\langle f, \frac{1}{\sqrt{L}} \cos \left(k \frac{\pi}{L} (\cdot) \right) \right\rangle = \frac{1}{\sqrt{L}} \int_{-L}^L f(t) \cos \left(k \frac{\pi}{L} t \right) dt \\ B_k &= \left\langle f, \frac{1}{\sqrt{L}} \sin \left(k \frac{\pi}{L} (\cdot) \right) \right\rangle = \frac{1}{\sqrt{L}} \int_{-L}^L f(t) \sin \left(k \frac{\pi}{L} t \right) dt, \quad \forall k \in \mathbb{N}. \end{aligned}$$

Sejam $a_0 = \sqrt{\frac{2}{L}} A_0$, $a_k = \frac{A_k}{\sqrt{L}}$ e $b_k = \frac{B_k}{\sqrt{L}}$, $\forall k \in \mathbb{N}$, então

$$\begin{aligned} a_0 &= \frac{1}{L} \int_{-L}^L f(t) dt \\ a_k &= \frac{1}{L} \int_{-L}^L f(t) \cos \left(k \frac{\pi}{L} t \right) dt \\ b_k &= \frac{1}{L} \int_{-L}^L f(t) \sin \left(k \frac{\pi}{L} t \right) dt, \quad \forall k \in \mathbb{N}. \end{aligned} \quad (3.11)$$

Assim, substituindo (3.11) em (3.10) obtemos a série de Fourier

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left\{ a_k \cos \left(k \frac{\pi}{L} t \right) + b_k \sin \left(k \frac{\pi}{L} t \right) \right\}. \quad (3.12)$$

Sabendo que para $i = \sqrt{-1}$

$$\begin{aligned} \cos \left(k \frac{\pi}{L} t \right) &= \frac{e^{ik \frac{\pi}{L} t} + e^{-ik \frac{\pi}{L} t}}{2} \\ \sin \left(k \frac{\pi}{L} t \right) &= \frac{e^{ik \frac{\pi}{L} t} - e^{-ik \frac{\pi}{L} t}}{2i} \end{aligned}$$

temos

$$\begin{aligned}
f(t) &= \frac{a_0}{2} + \sum_{k=1}^{\infty} \left\{ a_k \left(\frac{e^{ik\frac{\pi}{L}t} + e^{-ik\frac{\pi}{L}t}}{2} \right) - ib_k \left(\frac{e^{ik\frac{\pi}{L}t} - e^{-ik\frac{\pi}{L}t}}{2} \right) \right\} \\
&= \frac{a_0}{2} + \sum_{k=1}^{\infty} \left\{ \left(\frac{a_k - ib_k}{2} \right) e^{ik\frac{\pi}{L}t} + \left(\frac{a_k + ib_k}{2} \right) e^{-ik\frac{\pi}{L}t} \right\} \\
&= c_0 + \sum_{k=1}^{\infty} \left\{ c_k e^{ik\frac{\pi}{L}t} + c_{(-k)} e^{-ik\frac{\pi}{L}t} \right\} \\
&= \sum_{k=-\infty}^{\infty} c_k e^{ik\frac{\pi}{L}t}
\end{aligned} \tag{3.13}$$

usando (3.11) em c_k , $k \in \mathbb{Z}$, temos

$$\begin{aligned}
c_k &= \frac{a_k - ib_k}{2}, \quad \forall k \in \mathbb{N} \\
&= \frac{1}{L} \int_{-L}^L f(t) \left[\frac{\cos(k\frac{\pi}{L}t) - i \sin(k\frac{\pi}{L}t)}{2} \right] dt \\
&= \frac{1}{2L} \int_{-L}^L f(t) e^{-ik\frac{\pi}{L}t} dt
\end{aligned}$$

também

$$\begin{aligned}
c_{(-k)} &= \frac{a_k + ib_k}{2}, \quad \forall k \in \mathbb{N} \\
&= \frac{1}{L} \int_{-L}^L f(t) \left[\frac{\cos(k\frac{\pi}{L}t) + i \sin(k\frac{\pi}{L}t)}{2} \right] dt \\
&= \frac{1}{2L} \int_{-L}^L f(t) e^{ik\frac{\pi}{L}t} dt \\
&= \frac{1}{2L} \int_{-L}^L f(t) e^{-i(-k)\frac{\pi}{L}t} dt
\end{aligned}$$

e

$$\begin{aligned}
c_0 &= \frac{a_0}{2} \\
&= \frac{1}{2L} \int_{-L}^L f(t) dt \\
&= \frac{1}{2L} \int_{-L}^L f(t) e^{-i(0)\frac{\pi}{L}t} dt.
\end{aligned}$$

Portanto

$$c_k = \frac{1}{2L} \int_{-L}^L f(t) e^{-ik \frac{\pi}{L} t} dt, \quad \forall k \in \mathbb{Z}.$$

Definindo $e_k^L(t) = \frac{e^{ik \frac{\pi}{L} t}}{\sqrt{2L}}$, $\forall k \in \mathbb{Z}$, então, $\{e_k^L\}_{k \in \mathbb{Z}}$ constitui uma base ortonormal para $L^2([-L, L])$ e

$$c_k = \frac{1}{\sqrt{2L}} \langle f, e_k^L \rangle = \frac{\hat{f}_k}{\sqrt{2L}}, \quad \forall k \in \mathbb{Z}.$$

Assim de (3.7) e (3.13) temos que para cada $f \in L^2([-L, L])$

$$f = \sum_{k=-\infty}^{\infty} \hat{f}_k e_k^L = \lim_{n \rightarrow \infty} \sum_{k=-n}^n \hat{f}_k e_k^L \quad (3.14)$$

com convergência em $L^2([-L, L])$.

O conjunto das funções contínuas em $[-L, L]$, em relação a norma $\|\cdot\|_{L^2([-L, L])}$, é um conjunto denso em $L^2([-L, L])$. Portanto, para cada função f contínua em $[-L, L]$, f pode ser aproximada pela série (3.14). O seguinte teorema nos dá uma condição necessária para que a série (3.14) aproxime uniformemente f quando ela é contínua em $[-L, L]$.

Teorema 3.2.2 *Se $f \in \mathcal{C}([-L, L])$ onde*

$$\mathcal{C}([-L, L]) = \{f : \mathbb{R} \rightarrow \mathbb{R} / f \text{ é contínua em } [-L, L]\}$$

e suponha que $\sum_{k=-\infty}^{\infty} |\hat{f}_k| < \infty$; então a série de Fourier

$$\sum_{k=-\infty}^{\infty} \hat{f}_k e_k^L$$

converge uniformemente a f .

Segue que (ver Almira [1])

$$f(t) = \sum_{k=-\infty}^{\infty} \left[\int_{-L}^L f(t) \frac{e^{-ik \frac{\pi}{L} t}}{\sqrt{2L}} dt \right] \frac{e^{ik \frac{\pi}{L} t}}{\sqrt{2L}} = \sum_{k=-\infty}^{\infty} \left[\int_{-L}^L f(t) e^{-ik \frac{\pi}{L} t} dt \right] \frac{e^{ik \frac{\pi}{L} t}}{2L}$$

para $\xi_k = \frac{k}{2L}$ e $\Delta \xi = \xi_{k+1} - \xi_k = \frac{1}{2L}$, $\forall k \in \mathbb{Z}$. Então para L suficientemente grande temos

$$\begin{aligned}
f(t) &= \sum_{k=-\infty}^{\infty} \left[\int_{-L}^L f(s) e^{-2\pi i \xi_k s} ds \right] e^{2\pi i \xi_k t} \Delta \xi \\
&\approx \sum_{k=-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(s) e^{-2\pi i \xi_k s} ds \right] e^{2\pi i \xi_k t} \Delta \xi.
\end{aligned}$$

Isso é possível se f for absolutamente integrável, ou seja

$$\int_{-\infty}^{\infty} |f(t)| dt < \infty$$

isto é, $\forall k \in \mathbb{Z}$ temos

$$\left| \int_{-\infty}^{\infty} f(t) e^{-2\pi i \xi_k t} dt \right| \leq \int_{-\infty}^{\infty} |f(t) e^{-2\pi i \xi_k t}| dt \leq \int_{-\infty}^{\infty} |f(t)| dt < \infty$$

Fazendo $L \rightarrow \infty$, então, $\Delta \xi \rightarrow 0$ e $\{\xi_k\}_{k \in \mathbb{Z}}$ pode ser interpretado como nós equidistantes de uma partição de Riemann para a integral limite

$$\int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(s) e^{-2\pi i \xi s} ds \right] e^{2\pi i \xi t} d\xi$$

ou seja, podemos concluir (para certas condições de suavidade da função f) que

$$f(t) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(s) e^{-2\pi i \xi s} ds \right] e^{2\pi i \xi t} d\xi. \quad (3.15)$$

Esse resultado, conhecido como teorema integral de Fourier, será enunciado abaixo.

Teorema 3.2.3 *Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ uma função suave por partes, absolutamente integrável. Então f tem uma representação por integral de Fourier dada em (3.15) que converge para $f(t)$ nos pontos de continuidade de f e para a média dos limites laterais nos pontos de descontinuidade de f .*

3.2.2 Transformada de Fourier

Os detalhes podem ser vistas em Rodney [12], em Almira [1] e em González [9].

Seja $f : \mathbb{R} \rightarrow \mathbb{R}$, definimos a transformada de Fourier como

$$\mathcal{F}\{f\}(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}} f(t) e^{-2\pi i \xi t} dt, \quad i = \sqrt{-1} \quad e \quad f \in L^1(\mathbb{R}) \quad (3.16)$$

onde

$$L^1(\mathbb{R}) = \left\{ f \text{ mensurável} / \int_{\mathbb{R}} |f(t)| dt < \infty \right\}.$$

Consequentemente, se f for suave por partes, então pelo teorema 3.2.3, a expressão que permite obter $f(t)$ a partir de $\mathcal{F}\{f\}(\xi)$ é chamada de “transformada de Fourier inversa” e é definida por

$$f(t) = \mathcal{F}^{-1}\{\hat{f}\}(t) = \int_{\mathbb{R}} \hat{f}(\xi) e^{2\pi i t \xi} d\xi. \quad (3.17)$$

Observe que apesar da função f ser uma função definida na reta (isto é, uma função de variável real) tomando valores reais, em geral a função $\hat{f}$ é uma função definida na reta tomando valores complexos. Portanto, no estudo de transformadas de Fourier é inevitável o aparecimento de funções de $\mathbb{R}$ em $\mathbb{C}$.

Isso nos leva a seguinte definição: definimos a transformada de Fourier de f , como a função $\mathcal{F}$ que associa a cada função absolutamente integrável $f : \mathbb{R} \rightarrow \mathbb{R}$ a função $\hat{f} : \mathbb{R} \rightarrow \mathbb{C}$ definida pela expressão (3.16); a sua inversa, chamada a transformada de Fourier inversa, é a função $\mathcal{F}^{-1}$ que associa a cada função $\hat{f} : \mathbb{R} \rightarrow \mathbb{C}$ que pertence ao conjunto imagem de $\mathcal{F}$ a função absolutamente integrável $f : \mathbb{R} \rightarrow \mathbb{R}$ definida pela expressão (3.17). Assim, se f é contínua,

$$\mathcal{F}^{-1}\{\mathcal{F}\{f\}\} = f. \quad (3.18)$$

Isso é consequência imediata de (3.15).

Para funções que assintoticamente não tendem para zero quando $t \rightarrow \infty$ ou $-\infty$ em geral a transformada de Fourier não está definida. Para que isto não ocorra apresentamos o seguinte resultado.

Teorema 3.2.4 *Seja $f \in L^1(\mathbb{R})$ e suave por partes, então a sua transformada de Fourier $\hat{f}$ é uma função contínua e seu operador $\mathcal{F} : L^1(\mathbb{R}) \rightarrow \mathcal{C}(\mathbb{R})$ é um operador limitado. Se, além disso, $\hat{f} \in L^1(\mathbb{R})$, então f é contínua e $\mathcal{F}^{-1} : L^1(\mathbb{R}) \rightarrow \mathcal{C}(\mathbb{R})$ é um operador limitado.*

Algumas funções do espaço $L^2(\mathbb{R})$ poderiam não possuir transformada de Fourier, onde

$$L^2(\mathbb{R}) = \left\{ f \text{ mensurável} / \int_{\mathbb{R}} |f(t)|^2 dt < \infty \right\},$$

pela simples razão que estas funções poderiam não estar dentro do espaço $L^1(\mathbb{R})$. Mesmo assim, seria útil dispor de uma extensão da transformada de Fourier a todo $L^2(\mathbb{R})$ porque este espaço é o substituto natural do espaço $l_2(\mathbb{Z})$ que proporciona boas propriedades para a série de Fourier. Portanto, poderia oferecer boas propriedades para a transformada de Fourier.

Esta extensão pode ser sintetizada pelo seguinte resultado.

Lema 3.2.1 *Se $f, g \in L^1(\mathbb{R})$ são tais que $\hat{f}, \hat{g} \in L^1(\mathbb{R})$, então tanto f, g como $\hat{f}, \hat{g}$ pertencem a $L^2(\mathbb{R})$ e também*

$$\langle f, g \rangle_{L^2(\mathbb{R})} = \langle \hat{f}, \hat{g} \rangle_{L^2(\mathbb{R})} \text{ e } \|f\|_{L^2(\mathbb{R})} = \|\hat{f}\|_{L^2(\mathbb{R})}.$$

Agora, se $f \in L^2(\mathbb{R})$, então é possível encontrar uma sequência $\{f_n\}_{n \in \mathbb{N}}$ tal que $f_n, \hat{f}_n \in L^1(\mathbb{R})$, $\forall n \in \mathbb{N}$, e $f_n \rightarrow f$ em $L^2(\mathbb{R})$ quando $n \rightarrow \infty$. Pelo lema 3.2.1 temos que $\{\hat{f}_n\}_{n \in \mathbb{N}}$ é de Cauchy em $L^2(\mathbb{R})$, assim, existe $F \in L^2(\mathbb{R})$ tal que $\hat{f}_n \rightarrow F$ em $L^2(\mathbb{R})$ quando $n \rightarrow \infty$.

F não depende da sequência $\{f_n\}_{n \in \mathbb{N}}$. Com efeito, se escolhermos outra sequência $\{g_n\}_{n \in \mathbb{N}}$ tal que $g_n, \hat{g}_n \in L^1(\mathbb{R})$, $\forall n \in \mathbb{N}$, e $g_n \rightarrow f$ em $L^2(\mathbb{R})$ quando $n \rightarrow \infty$, então temos que

$$\lim_{n \rightarrow \infty} \|\hat{f}_n - \hat{g}_n\|_{L^2(\mathbb{R})} = \lim_{n \rightarrow \infty} \|f_n - g_n\|_{L^2(\mathbb{R})} = 0$$

portanto

$$\lim_{n \rightarrow \infty} \|\hat{g}_n - F\|_{L^2(\mathbb{R})} = 0.$$

Isso mostra que F depende somente da função $f \in L^2(\mathbb{R})$.

Definimos a transformada de Fourier de $f \in L^2(\mathbb{R})$ como $\hat{f} = \lim_{n \rightarrow \infty} \hat{f}_n$ em $L^2(\mathbb{R})$, onde $\{f_n\}_{n \in \mathbb{N}} \subset L^2(\mathbb{R})$ tal que $f_n, \hat{f}_n \in L^1(\mathbb{R})$, $\forall n \in \mathbb{N}$, e $f_n \rightarrow f$ em $L^2(\mathbb{R})$ quando $n \rightarrow \infty$.

Dessa maneira obtemos o seguinte resultado conhecido como o teorema de Plancherel.

Teorema 3.2.5 *A transformada de Fourier definida originalmente em $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, pode ser estendida de forma única a um operador $\mathcal{F} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ ($\mathcal{F}\{f\}(\xi) = \hat{f}(\xi)$). Também*

$$\langle f, g \rangle_{L^2(\mathbb{R})} = \langle \hat{f}, \hat{g} \rangle_{L^2(\mathbb{R})} \text{ e } \|f\|_{L^2(\mathbb{R})} = \|\hat{f}\|_{L^2(\mathbb{R})}$$

cumpra $\forall f, g \in L^2(\mathbb{R})$.

3.2.3 Propriedades

Mencionaremos brevemente algumas propriedades da transformada de Fourier

1. **Linearidade $\mathcal{F}$:** Sejam $f, g \in L^1(\mathbb{R})$ e $K \in \mathbb{R}$. Então, $\mathcal{F}\{Kf + g\}(\xi) = K\mathcal{F}\{f\}(\xi) + \mathcal{F}\{g\}(\xi)$.

2. **Teorema do deslocamento na variável ξ :** Seja $f \in L^1(\mathbb{R})$. Então, $\mathcal{F}\{f\}(\xi - a) = \mathcal{F}\{e^{2\pi i a t} f\}(\xi)$ e $\mathcal{F}\{f(t - a)\}(\xi) = e^{-2\pi i a \xi} \mathcal{F}\{f\}(\xi)$.
3. **Mudança de Escala:** Seja $f \in L^1(\mathbb{R})$ e $a \neq 0$. Então, $\mathcal{F}\{f(at)\}(\xi) = \frac{1}{a} \mathcal{F}\{f\}\left(\frac{\xi}{a}\right)$ e $\mathcal{F}\{f(t)\}(a\xi) = \mathcal{F}\left\{\frac{1}{a} f\left(\frac{t}{a}\right)\right\}(\xi)$.
4. **Transformada de Fourier da derivada:** Se $f' \in L^1(\mathbb{R})$. Então, $\mathcal{F}\{f'\}(\xi) = 2\pi i \xi \mathcal{F}\{f\}(\xi)$.

3.2.4 Convolução

Sejam $f, g \in L^1(\mathbb{R})$, definimos a convolução como

$$(f * g)(t) = \int_{-\infty}^{\infty} f(s)g(t-s)ds. \quad (3.19)$$

A convolução está bem definida desde que uma das funções seja limitada (em $\mathbb{R}$) e a outra seja absolutamente integrável.

Proposição 3.2.1 *A convolução é uma operação comutativa.*

Demonstração:

Fazendo a mudança de variável $t - s = u$, tal que

$$\begin{aligned} (f * g)(t) &= \int_{-\infty}^{\infty} f(s)g(t-s)ds = - \int_{\infty}^{-\infty} f(t-u)g(u)du \\ &= \int_{-\infty}^{\infty} f(t-u)g(u)du = (g * f)(t). \end{aligned}$$

□

Proposição 3.2.2 (Desigualdade de Young). *Sejam $f, g \in L^1(\mathbb{R})$, suave por partes, então a sua convolução $(f * g) \in L^1(\mathbb{R})$ e verifica-se que*

$$\|(f * g)\|_{L^1(\mathbb{R})} \leq \|f\|_{L^1(\mathbb{R})} \|g\|_{L^1(\mathbb{R})}.$$

Demonstração:

Desde que $f, g \in L^1(\mathbb{R})$, usamos o teorema de Fubini e assim

$$\begin{aligned} \|(f * g)\|_{L^1(\mathbb{R})} &= \int_{-\infty}^{\infty} |(f * g)(t)| dt \\ &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} f(s)g(t-s)ds \right| dt \end{aligned}$$

$$\begin{aligned}
&\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(s)| |g(t-s)| ds dt \\
&= \int_{-\infty}^{\infty} |f(s)| \left\{ \int_{-\infty}^{\infty} |g(t-s)| dt \right\} ds \quad \text{por Fubini} \\
&= \left[\int_{-\infty}^{\infty} |f(s)| ds \right] \|g\|_{L^1(\mathbb{R})} \\
&= \|f\|_{L^1(\mathbb{R})} \|g\|_{L^1(\mathbb{R})}
\end{aligned}$$

□

Proposição 3.2.3 *Sejam $f, g, h \in L^1(\mathbb{R})$ e λ um escalar, então verifica-se que*

1. $(f * (g * h)) = ((f * g) * h)$
2. $(f * [g + h]) = (f * g) + (f * h)$
3. $([f + g] * h) = (f * h) + (g * h)$
4. $\lambda (f * g) = ([\lambda f] * g) = (f * [\lambda g])$

Demonstração:

Segue diretamente da definição (3.19)

□

Observe que, pela proposição 3.2.3 e a proposição 3.2.2, o conjunto $\{L^1(\mathbb{R}), +, *\}$, dotado da operação de produto binária (*), chamada de convolução em (3.19), define uma álgebra de Banach associativa e pela proposição 3.2.1 também é comutativa.

Teorema 3.2.6 *Sejam $f, g \in L^1(\mathbb{R})$, então*

$$\widehat{(f * g)}(\xi) = \hat{f}(\xi) \hat{g}(\xi)$$

Demonstração:

Já sabemos que $\widehat{(f * g)} \in L^1(\mathbb{R})$, então usando Fubini, temos

$$\begin{aligned}
\widehat{(f * g)}(\xi) &= \int_{-\infty}^{\infty} (f * g)(t) e^{-2\pi i \xi t} dt \\
&= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} f(s) g(t-s) ds \right\} e^{-2\pi i \xi t} dt \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s) e^{-2\pi i \xi s} g(t-s) e^{-2\pi i \xi (t-s)} ds dt \\
&= \int_{-\infty}^{\infty} f(s) e^{-2\pi i \xi s} \left\{ \int_{-\infty}^{\infty} g(t-s) e^{-2\pi i \xi (t-s)} dt \right\} ds
\end{aligned}$$

$$\begin{aligned}
&= \left[\int_{-\infty}^{\infty} f(s) e^{-2\pi i \xi s} ds \right] \hat{g}(\xi) \\
&= \hat{f}(\xi) \hat{g}(\xi)
\end{aligned}$$

□

3.3 A Transformada de Mellin

Os resultados e as definições podem ser vistas em Collins [7].

3.3.1 Convolução de Mellin

Para $f \in L^1(\mathbb{R})$, consideremos o operador $\mathcal{T}$ como $\mathcal{T}(f) = \mathcal{T}f = f \circ \exp$. Definamos a seguinte norma

$$\|f\|_{\mathcal{T}} = \int_{-\infty}^{\infty} |\mathcal{T}f(t)| dt = \int_{-\infty}^{\infty} |f(e^t)| dt = \int_0^{\infty} |f(u)| \frac{1}{u} du$$

Veja que, para $f \in L^1(\mathbb{R})$ não necessariamente a $\mathcal{T}$ -norma é finita. Por exemplo, $f(t) = e^{-t} \in L^1(\mathbb{R})$, mas $\int_0^{\infty} |e^{-u}| \frac{1}{u} du$ não converge.

Para evitar esse tipo de problema, vamos definir um novo tipo de norma chamada a $\mathcal{M}_c$ -norma como

$$\|f\|_{\mathcal{M}_c} = \int_0^{\infty} |f(t)| t^{c-1} dt$$

onde c é escolhido de maneira que garanta a convergência. Por exemplo as funções de densidade de uma variável aleatória T , $f_T(t)$, satisfazem $\int_0^{\infty} |f_T(t)| dt < \infty$. As funções que possuem melhor comportamento caem rapidamente para zero quando $t \rightarrow \infty$ e $\int_0^{\infty} |f_T(t)| t^p dt < \infty$ para $p \geq 1$. Assim $\|f_T\|_{\mathcal{M}_c} < \infty$ se $c \geq 1$ para o tipo de funções de densidade f_T com melhor comportamento.

Definimos a convolução de Mellin a partir da convolução usual usando o operador $\mathcal{T}$ como

$$\begin{aligned}
(f \otimes g)(x) &= (\mathcal{T}f * \mathcal{T}g)(t) \quad \text{onde } x = e^t \\
&= \int_{-\infty}^{\infty} \mathcal{T}f(t-s) \mathcal{T}g(s) ds \\
&= \int_{-\infty}^{\infty} f(e^{t-s}) g(e^s) ds \\
&= \int_0^{\infty} f(e^{\log x - \log w}) g(e^{\log w}) \frac{1}{w} dw \\
&= \int_0^{\infty} f\left(\frac{x}{w}\right) g(w) \frac{1}{w} dw.
\end{aligned} \tag{3.20}$$

Observe que da fórmula (3.20), para $x = e^t$ e da definição de $\mathcal{T}$ temos que

$$(f \otimes g)(x) = (f \otimes g)(e^t) = \mathcal{T}(f \otimes g)(t) = (\mathcal{T}f * \mathcal{T}g)(t). \quad \forall t \in \mathbb{R}$$

Assim

$$\mathcal{T}(f \otimes g) = (\mathcal{T}f * \mathcal{T}g). \quad (3.21)$$

Portanto o operador $\mathcal{T}$ se comporta como um homomorfismo para as operações $(\otimes)$ e $(*)$.

Definimos

$$\mathcal{M}(\mathbb{R}_+) = \left\{ f \in L^1(\mathbb{R}) / \exists S \text{ com } \int_0^\infty |f(t)| t^{s-1} dt < \infty, \quad \forall s \in S \right\}$$

e

$$\|f\|_{\mathcal{M}_s} = \int_0^\infty |f(t)| t^{s-1} dt.$$

O conjunto $\mathcal{M}_c(\mathbb{R}_+)$, tal que a $\mathcal{M}_c$ -norma é finita para algum $c \in \mathbb{R}$ fixo, é um espaço de Banach.

Por ser $\mathcal{T}$ um homomorfismo para as operações $(\otimes)$ e $(*)$, equação (3.21), temos que, pela proposição 3.2.2 a convolução de Mellin como sendo $(\otimes)$, é uma operação de produto binária fechada no conjunto $\mathcal{M}_c(\mathbb{R}_+)$, e a desigualdade de normas com a $\mathcal{M}_c$ -norma é adotada. A convolução de Mellin, dentro do conjunto $\mathcal{M}_c(\mathbb{R}_+)$ também adota as propriedades da proposição 3.2.1 e da proposição 3.2.3.

Assim, o conjunto $\{\mathcal{M}_c(\mathbb{R}_+), +, \otimes\}$ é uma álgebra de Banach associativa e comutativa. Ver os detalhes dos conceitos algébricos em Barata [4].

3.3.2 Transformada de Mellin

Se $f \in \mathcal{M}_c(\mathbb{R}_+)$, $\forall c \in [a, b]$, diremos que $f \in \mathcal{M}_{[a,b]}(\mathbb{R}_+)$ (nossas funções de densidade bem comportadas pertencem a $\mathcal{M}_{[1,\infty)}(\mathbb{R}_+)$). Definimos a transformada de Mellin da variável $s \in \mathbb{C}$ como

$$\mathcal{M}[f](s) = \int_0^\infty f(t) t^{s-1} dt \quad (3.22)$$

onde $a \leq \operatorname{Re}(s) \leq b$.

Para $F(s) = \mathcal{M}[f](s)$, a “transformada de Mellin inversa” é definida como

$$f(x) = \mathcal{M}^{-1} \{ \mathcal{M}[f] \} (x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} F(s) x^{-s} ds. \quad (3.23)$$

A condição para que a inversa exista é que a função $F(s)x^{-s}$ seja analítica na faixa $(a, b) \times (-i\infty, i\infty)$ tal que $c \in (a, b)$.

Podemos encontrar uma relação entre a transformada de Fourier definida em (3.16) e a transformada de Mellin definida em (3.22) usando o operador $\mathcal{T}$ e a substituição de $\xi = -\frac{\eta-c}{2\pi i}$ para $c \geq 0$:

$$\begin{aligned} \widehat{\mathcal{T}f}(\xi) &= \int_{-\infty}^{\infty} \mathcal{T}f(t) e^{-2\pi i \xi t} dt \\ \widehat{\mathcal{T}f}\left(-\frac{\eta-c}{2\pi i}\right) &= \int_{-\infty}^{\infty} f(e^t) e^{(\eta-c)t} dt \\ &= \int_0^{\infty} f(x) e^{(\eta-c) \log x} \frac{1}{x} dx \\ &= \int_0^{\infty} f(x) x^{(\eta-c)} \frac{1}{x} dx \\ &= \int_0^{\infty} [f(x) x^{-c}] x^{(\eta-1)} dx \\ &= \int_0^{\infty} [f^{[c]}(x)] x^{(\eta-1)} dx \quad \text{sendo } f^{[c]}(x) = f(x) x^{-c} \\ &= \mathcal{M} [f^{[c]}] (\eta). \end{aligned}$$

Assim

$$\mathcal{F} \{ \mathcal{T}f \} (\xi) = \mathcal{M} [f^{[c]}] (\eta) \quad (3.24)$$

para $\xi = -\frac{\eta-c}{2\pi i}$ e $f^{[c]}(x) = f(x)x^{-c}$ com $c \geq 0$ e $x > 0$.

Definindo o operador inverso $\mathcal{T}^{-1}$ como $\mathcal{T}^{-1}(f) = \mathcal{T}^{-1}f = f \circ \log$ temos que para $x = e^t$

$$(\mathcal{T}^{-1} \circ \mathcal{T}) f(x) = \mathcal{T}^{-1} (\mathcal{T}f) (x) = (\mathcal{T}f) \circ \log (x) = \mathcal{T}f(\log x) = \mathcal{T}f(t) = f(e^t) = f(x)$$

e

$$(\mathcal{T} \circ \mathcal{T}^{-1}) f(t) = \mathcal{T} (\mathcal{T}^{-1}f) (t) = (\mathcal{T}^{-1}f) \circ \exp (t) = \mathcal{T}^{-1}f(e^t) = \mathcal{T}^{-1}f(x) = f(\log x) = f(t).$$

Ou seja

$$(\mathcal{T}^{-1} \circ \mathcal{T}) f|_{\mathbb{R}_+} = f|_{\mathbb{R}_+} \quad \wedge \quad (\mathcal{T} \circ \mathcal{T}^{-1}) f = f.$$

Agora para $\xi = -\frac{\eta-c}{2\pi i}$ e $f^{[c]}(x) = f(x)x^{-c}$ com $c \geq 0$ e $x > 0$, podemos usar a transformada de Fourier inversa definida em (3.17) e a transformada de Mellin inversa definida em (3.23) para encontrar uma relação entre as transformadas de Fourier e Mellin usando o operador inverso $\mathcal{T}^{-1}$ como

$$\begin{aligned} \mathcal{T}^{-1} f(x) &= \mathcal{T}^{-1} \left[\mathcal{F}^{-1} \left\{ \hat{f} \right\} \right] (x) \\ &= \mathcal{F}^{-1} \left\{ \hat{f} \right\} (\log x) \\ &= \int_{-\infty}^{\infty} \hat{f}(\xi) e^{2\pi i \log(x) \xi} d\xi \\ &= \int_{-\infty}^{\infty} \hat{f}(\xi) x^{2\pi i \xi} d\xi \\ &= -\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \hat{f} \left(-\frac{\eta-c}{2\pi i} \right) x^{-(\eta-c)} d\eta \\ &= \frac{x^c}{2\pi i} \int_{c-i\infty}^{c+i\infty} \hat{f} \left(-\frac{\eta-c}{2\pi i} \right) x^{-\eta} d\eta. \end{aligned}$$

Então

$$\begin{aligned} (\mathcal{T}^{-1} f)^{[c]}(x) &= x^{-c} (\mathcal{T}^{-1} f)(x) \\ &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \hat{f} \left(-\frac{\eta-c}{2\pi i} \right) x^{-\eta} d\eta \\ &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \mathcal{M} \left[(\mathcal{T}^{-1} f)^{[c]} \right] (\eta) x^{-\eta} d\eta \quad \text{por (3.24)} \\ &= \mathcal{M}^{-1} \left\{ \mathcal{M} \left[(\mathcal{T}^{-1} f)^{[c]} \right] \right\} (x) \quad \text{por (3.23)}. \end{aligned}$$

Ou seja

$$\mathcal{F} \{ f \} (\xi) = \mathcal{M} \left[(\mathcal{T}^{-1} f)^{[c]} \right] (\eta). \quad (3.25)$$

Com isso vemos que, de (3.24) e (3.25) o operador $\mathcal{T}$, descrito por Collins [7], está bem definido e proporciona uma relação entre as transformadas de Fourier e Mellin.

Em alguns casos o operador $\mathcal{T}$ proporciona uma maneira mais simples de inverter a transformada de Mellin através do uso da transformada de Fourier inversa.

Lema 3.3.1 Se $f^{[c]}(x) = f(x)x^{-c}$, para $x > 0$ e $c \geq 0$, então $(f \otimes g)^{[c]}(x) = (f^{[c]} \otimes g^{[c]})(x)$.

Demonstração:

Com efeito, usando (3.20)

$$\begin{aligned}
 (f^{[c]} \otimes g^{[c]})(x) &= \int_0^\infty f^{[c]}\left(\frac{x}{w}\right) g^{[c]}(w) \frac{1}{w} dw \\
 &= \int_0^\infty \left[\left(\frac{x}{w}\right)^{-c} f\left(\frac{x}{w}\right) \right] \left[w^{-c} g(w) \right] \frac{1}{w} dw \\
 &= x^{-c} \int_0^\infty f\left(\frac{x}{w}\right) g(w) \frac{1}{w} dw \\
 &= x^{-c} (f \otimes g)(x) \\
 &= (f \otimes g)^{[c]}(x).
 \end{aligned}$$

□

Teorema 3.3.1 Se $f, g \in \mathcal{M}(\mathbb{R}_+)$, então

$$\mathcal{M}[f \otimes g](s) = \mathcal{M}[f](s) \mathcal{M}[g](s).$$

Demonstração:

Com efeito, usando o teorema de Fubini e substituindo $y = \frac{x}{w}$, temos

$$\begin{aligned}
 \mathcal{M}[f \otimes g](s) &= \int_0^\infty \left[\int_0^\infty f\left(\frac{x}{w}\right) g(w) \frac{1}{w} dw \right] x^{s-1} dx \\
 &= \int_0^\infty \left[\int_0^\infty f\left(\frac{x}{w}\right) x^{s-1} dx \right] g(w) \frac{1}{w} dw \\
 &= \int_0^\infty \left[\int_0^\infty f(y) y^{s-1} w^{s-1} w dy \right] g(w) \frac{1}{w} dw \\
 &= \int_0^\infty \left[\int_0^\infty f(y) y^{s-1} dy \right] g(w) w^{s-1} dw \\
 &= \left[\int_0^\infty f(y) y^{s-1} dy \right] \left[\int_0^\infty g(w) w^{s-1} dw \right] \\
 &= \mathcal{M}[f](s) \mathcal{M}[g](s)
 \end{aligned}$$

□

Pelos teoremas 3.2.6 e 3.3.1, pelo lema 3.3.1, e pelas equações (3.21) e (3.24) podemos verificar que

$$\mathcal{F}\{\mathcal{T}(f \otimes g)\}(\xi) = \mathcal{F}\{\mathcal{T}f * \mathcal{T}g\}(\xi)$$

$$\begin{aligned}
&= \mathcal{F}\{\mathcal{T}f\}(\xi)\mathcal{F}\{\mathcal{T}g\}(\xi) \\
&= \mathcal{M}\left[f^{[c]}\right](\eta)\mathcal{M}\left[g^{[c]}\right](\eta) \\
&= \mathcal{M}\left[f^{[c]} \otimes g^{[c]}\right](\eta) \\
&= \mathcal{M}\left[(f \otimes g)^{[c]}\right](\eta).
\end{aligned}$$

Assim encontramos uma relação entre as transformadas e convoluções de Fourier e Mellin junto com o operador $\mathcal{T}$ de Collins.

O teorema 3.3.1 pode ser estendido para n funções. Ou seja, dado $\{f_k\}_{k=1}^n \subset \mathcal{M}(\mathbb{R}_+)$ temos

$$\mathcal{M}[f_1 \otimes \cdots \otimes f_n](s) = \mathcal{M}[f_1](s) \cdots \mathcal{M}[f_n](s). \quad (3.26)$$

3.4 Relação entre as convoluções e as funções de distribuição

A análise estatística de sobrevivência estuda o risco de falha ou morte ao longo do tempo. Esse tempo, chamado de tempo de falha, pode ser representado por uma variável aleatória não negativa, sendo usualmente especificada pela sua função de sobrevivência.

A função de sobrevivência é definida como a probabilidade de uma observação não falhar até um certo tempo t . Ou seja, para uma variável aleatória $T \geq 0$, definimos a função de sobrevivência, $\bar{F}_T(t)$, como: $\bar{F}_T(t) = \mathbb{P}(\{T > t\})$, $\forall t \geq 0$.

Hogg [10] investiga um método para descrever analiticamente a distribuição do tempo de vida de um sistema mediante a construção de uma estrutura algébrica. Para isto, Hogg [10] usou propriedades inerentes de convolução e mistura de distribuições do tempo de vida.

Uma vez que a estrutura algébrica é construída, Hogg [10] propõe usar esta estrutura para desenvolver algoritmos para ir desde o diagrama esquemático de um sistema até sua respectiva função de sobrevivência.

Mistura de componentes da taxa de falha de um sistema pode ser descrita por mistura de convoluções. Essas expressões podem ser simplificadas usando as propriedades da estrutura algébrica encontrada.

3.4.1 Soma de variáveis aleatórias

Seja $T > 0$ uma variável aleatória, denotemos a sua respectiva função de distribuição como F_T e a de densidade como f_T .

Hogg [10] expressa a convolução de Fourier em termos da distribuição da soma de duas variáveis aleatórias independentes $T_1, T_2 > 0$ para a função de sobrevivência e a função de densidade.

$$\bar{F}_{T_1+T_2}(t) = \bar{F}_{T_1}(t) + \int_0^t \bar{F}_{T_2}(t-s)f_{T_1}(s)ds \quad (3.27)$$

e

$$f_{T_1+T_2}(t) = \int_0^t f_{T_2}(t-s)f_{T_1}(s)ds. \quad (3.28)$$

Com efeito

$$\begin{aligned} F_{T_1+T_2}(t) &= \mathbb{P}(\{T_1 + T_2 \leq t\}) \\ &= \int_0^t \mathbb{P}(\{T_2 \leq t-s\}|\{T_1 = s\})f_{T_1}(s)ds \\ &= \int_0^t \mathbb{P}(\{T_2 \leq t-s\})f_{T_1}(s)ds \quad \text{independencia} \\ &= \int_0^t F_{T_2}(t-s)f_{T_1}(s)ds. \end{aligned}$$

Então

$$\begin{aligned} \bar{F}_{T_1+T_2}(t) &= 1 - \int_0^t F_{T_2}(t-s)f_{T_1}(s)ds \\ &= \bar{F}_{T_1}(t) + F_{T_1}(t) - \int_0^t F_{T_2}(t-s)f_{T_1}(s)ds \\ &= \bar{F}_{T_1}(t) + \int_0^t f_{T_1}(s)ds - \int_0^t F_{T_2}(t-s)f_{T_1}(s)ds \\ &= \bar{F}_{T_1}(t) + \int_0^t [1 - F_{T_2}(t-s)]f_{T_1}(s)ds \\ &= \bar{F}_{T_1}(t) + \int_0^t \bar{F}_{T_2}(t-s)f_{T_1}(s)ds. \end{aligned}$$

Segue que (3.27) é fechada no espaço das distribuições contínuas com a operação convolução.

Analogamente obtemos

$$F_{T_1+T_2}(t) = \int_0^t F_{T_1}(t-s)f_{T_2}(s)ds = F_{T_2+T_1}(t)$$

e

$$\bar{F}_{T_1+T_2}(t) = \bar{F}_{T_2}(t) + \int_0^t \bar{F}_{T_1}(t-s)f_{T_2}(s)ds = \bar{F}_{T_2+T_1}(t).$$

Assim (3.27) segue a propriedade comutativa.

Usando o teorema de Fubini vemos também que

$$\begin{aligned}
\bar{F}_{(T_1+T_2)+T_3}(t) &= \bar{F}_{T_1+T_2}(t) + \int_0^t \bar{F}_{T_3}(t-s)f_{T_1+T_2}(s)ds \\
&= \bar{F}_{T_1+T_2}(t) + \int_0^t \bar{F}_{T_3}(s)f_{T_1+T_2}(t-s)ds \\
&= \bar{F}_{T_1}(t) + \int_0^t \bar{F}_{T_2}(t-s)f_{T_1}(s)ds + \int_0^t \bar{F}_{T_3}(s) \left[\int_0^{t-s} f_{T_2}(t-s-u)f_{T_1}(u)du \right] ds \\
&= \bar{F}_{T_1}(t) + \int_0^t \bar{F}_{T_2}(t-s)f_{T_1}(s)ds + \int_0^t \int_0^{t-s} [\bar{F}_{T_3}(s)f_{T_2}(t-s-u)f_{T_1}(u)] dud s \\
&= \bar{F}_{T_1}(t) + \int_0^t \bar{F}_{T_2}(t-s)f_{T_1}(s)ds + \int_0^t \left[\int_0^{t-u} \bar{F}_{T_3}(s)f_{T_2}(t-s-u)ds \right] f_{T_1}(u)du \\
&= \bar{F}_{T_1}(t) + \int_0^t \bar{F}_{T_2}(t-s)f_{T_1}(s)ds + \int_0^t \left[\int_0^{t-u} \bar{F}_{T_3}(t-s-u)f_{T_2}(s)ds \right] f_{T_1}(u)du \\
&= \bar{F}_{T_1}(t) + \int_0^t \bar{F}_{T_2}(t-s)f_{T_1}(s)ds + \int_0^t \left[\int_0^{t-s} \bar{F}_{T_3}(t-s-u)f_{T_2}(u)du \right] f_{T_1}(s)ds \\
&= \bar{F}_{T_1}(t) + \int_0^t \left[\bar{F}_{T_2}(t-s) + \int_0^{t-s} \bar{F}_{T_3}(t-s-u)f_{T_2}(u)du \right] f_{T_1}(s)ds \\
&= \bar{F}_{T_1}(t) + \int_0^t \bar{F}_{T_2+T_3}(t-s)f_{T_1}(s)ds \\
&= \bar{F}_{T_1+(T_2+T_3)}(t).
\end{aligned}$$

Com isso segue que (3.27) é associativa.

Também, se $u = t - s$, então definimos $\phi(s, u) = (s, s + u) = (s, t)$, então, $\phi^{-1}(s, t) = (s, u) = (s, t - s)$ e

$$\mathcal{J}_{\phi^{-1}}(s, t) = \begin{vmatrix} \frac{\partial \phi_1^{-1}}{\partial s} & \frac{\partial \phi_1^{-1}}{\partial t} \\ \frac{\partial \phi_2^{-1}}{\partial s} & \frac{\partial \phi_2^{-1}}{\partial t} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1.$$

Assim

$$\begin{aligned}
f_{(T_1, T_1+T_2)}(s, t) &= (f_{(T_1, T_2)} \circ \phi^{-1})(s, t) \mathcal{J}_{\phi^{-1}}(s, t) \\
&= f_{(T_1, T_2)}(s, t - s) \\
&= f_{T_1}(s)f_{T_2}(t - s).
\end{aligned}$$

Portanto

$$f_{T_1+T_2}(t) = \int_0^t f_{(T_1, T_1+T_2)}(s, t)ds$$

$$\begin{aligned}
&= \int_0^t f_{T_1}(s) f_{T_2}(t-s) ds \\
&= \int_0^t f_{T_2}(t-s) f_{T_1}(s) ds
\end{aligned}$$

e

$$f_{T_1+T_2}(t) = \int_0^t f_{T_2}(t-s) f_{T_1}(s) ds = \int_0^t f_{T_1}(t-s) f_{T_2}(s) ds = f_{T_2+T_1}(t)$$

ou

$$f_{T_1+T_2}(t) = (f_{T_1} * f_{T_2})(t). \quad (3.29)$$

Dentro do espaço das distribuições contínuas, as funções de densidade da soma de variáveis aleatórias independentes relacionam-se diretamente com a operação convolução definida em (3.19). Portanto, temos que (3.28) adota as propriedades de uma álgebra de Banach associativa e comutativa.

Assim (3.28) é fechada, é associativa e é comutativa no espaço das distribuições contínuas com a operação convolução.

Hogg [10] também expressa a operação de mistura ou combinação, MIX, das distribuições de duas variáveis aleatórias independentes $T_1, T_2 > 0$ para a função de sobrevivência e a função de densidade.

$$MIX[a_1 \bar{F}_{T_1}(t), a_2 \bar{F}_{T_2}(t)] = a_1 \bar{F}_{T_1}(t) + a_2 \bar{F}_{T_2}(t) \quad (3.30)$$

e

$$MIX[a_1 f_{T_1}(t), a_2 f_{T_2}(t)] = a_1 f_{T_1}(t) + a_2 f_{T_2}(t) \quad (3.31)$$

onde $a_1, a_2 > 0$ e $a_1 + a_2 = 1$, chamadas de probabilidades a priori de mistura.

A operação MIX definida em (3.30) e (3.31) também adota as propriedades de ser fechada, associativa e comutativa no espaço das distribuições contínuas.

Podemos definir a função nula como sendo $zero(t) = 0$. Se anexamos esta função nula com o espaço das distribuições contínuas, torna-se um elemento neutro para a operação convolução e MIX, assim obtemos o seguinte resultado:

Teorema 3.4.1 *O espaço de distribuições contínuas com a função zero(t) adjunta, forma um monoide abeliano com a operação convolução definida em (3.27) e (3.28) e a operação MIX definida em (3.30) e (3.31).*

3.4.2 Produto de variáveis aleatórias

Sejam $X_1, X_2 > 1$ v.a. independentes com função de densidade f_{X_1}, f_{X_2} . Vamos determinar qual é a função de densidade de $X_1 X_2$. Para isso definimos $\varphi(w, y) = (w, wy) = (w, x)$, então $\varphi^{-1}(w, x) = (w, y) = (w, x/w)$ e

$$\mathcal{J}_{\varphi^{-1}}(w, x) = \begin{vmatrix} \frac{\partial \varphi_1^{-1}}{\partial w} & \frac{\partial \varphi_1^{-1}}{\partial x} \\ \frac{\partial \varphi_2^{-1}}{\partial w} & \frac{\partial \varphi_2^{-1}}{\partial x} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -\frac{x}{w^2} & \frac{1}{w} \end{vmatrix} = \frac{1}{w}.$$

Assim

$$\begin{aligned} f_{(X_1, X_1 X_2)}(w, x) &= (f_{(X_1, X_2)} \circ \varphi^{-1})(w, x) \mathcal{J}_{\varphi^{-1}}(w, x) \\ &= f_{(X_1, X_2)}\left(w, \frac{x}{w}\right) \frac{1}{w} \\ &= f_{X_1}(w) f_{X_2}\left(\frac{x}{w}\right) \frac{1}{w}. \end{aligned}$$

Portanto

$$\begin{aligned} f_{X_1 X_2}(x) &= \int_1^x f_{(X_1, X_1 X_2)}(w, x) dw \\ &= \int_1^x f_{X_1}(w) f_{X_2}\left(\frac{x}{w}\right) \frac{1}{w} dw \\ &= \int_1^x f_{X_2}\left(\frac{x}{w}\right) f_{X_1}(w) \frac{1}{w} dw \end{aligned}$$

e

$$f_{X_1 X_2}(x) = \int_1^x f_{X_2}\left(\frac{x}{w}\right) f_{X_1}(w) \frac{1}{w} dw = \int_1^x f_{X_1}\left(\frac{x}{w}\right) f_{X_2}(w) \frac{1}{w} dw = f_{X_2 X_1}(x)$$

obtemos

$$f_{X_1 X_2}(x) = (f_{X_1} \otimes f_{X_2})(x). \quad (3.32)$$

Segue que, dentro do espaço das distribuições absolutamente contínuas, as funções de densidade do produto de variáveis aleatórias independentes relaciona-se diretamente com a operação convolução de Mellin. Portanto, por ser $\mathcal{T}$ um homomorfismo, como visto em (3.21), temos que (3.32) com a $\mathcal{M}_c$ -norma, para algum $c \in \mathbb{R}$ fixo, adota as propriedades de uma álgebra de Banach associativa e comutativa.

Assim (3.32) é fechada, é associativa e é comutativa no espaço das distribuições absolutamente contínuas com a operação convolução de Mellin.

Usando a fórmula (3.21) para comparar (3.29) e (3.32), vemos que para $X = e^T$ temos

$$F_X(x) = \mathbb{P}(\{X \leq x\}) = \mathbb{P}(\{e^T \leq x\}) = \mathbb{P}(\{T \leq \log x\}) = F_T(\log x) = [\mathcal{T}^{-1}F_T](x). \quad (3.33)$$

Analogamente

$$\bar{F}_X(x) = \bar{F}_T(\log x) = [\mathcal{T}^{-1}\bar{F}_T](x). \quad (3.34)$$

Derivando (3.33) com respeito a x temos

$$f_X(x) = x^{-1}f_T(\log x) = x^{-1}[\mathcal{T}^{-1}f_T](x). \quad (3.35)$$

De (3.33) podemos observar que

$$F_T(t) = F_T(\log x) = F_X(x) = F_X(e^t) = [\mathcal{T}F_X](t) \quad (3.36)$$

analogamente

$$\bar{F}_T(t) = \bar{F}_X(e^t) = [\mathcal{T}\bar{F}_X](t). \quad (3.37)$$

Derivando (3.36) com respeito a t temos

$$f_T(t) = e^t f_X(e^t) = \exp(t) [\mathcal{T}f_X](t)$$

ou

$$[\mathcal{T}f_X](t) = \exp(-t)f_T(t). \quad (3.38)$$

Usando (3.38) obtemos a seguinte fórmula

$$\begin{aligned} [\mathcal{T}f_{X_1 X_2}](t) &= [\mathcal{T}(f_{X_1} \otimes f_{X_2})](t) \\ &= (\mathcal{T}f_{X_1} * \mathcal{T}f_{X_2})(t) \end{aligned}$$

$$\begin{aligned}
&= \left(\left[\frac{f_{T_1}}{\exp} \right] * \left[\frac{f_{T_2}}{\exp} \right] \right) (t) \\
&= \exp(-t) (f_{T_1} * f_{T_2}) (t) \\
&= e^{-t} f_{T_1+T_2}(t)
\end{aligned}$$

ou

$$f_{X_1 X_2}(x) = x^{-1} [\mathcal{T}^{-1} f_{T_1+T_2}] (x).$$

Assim

$$\begin{aligned}
F_{X_1 X_2}(x) &= \int_1^x f_{X_1 X_2}(w) dw \\
&= \int_1^x w^{-1} [\mathcal{T}^{-1} f_{T_1+T_2}] (w) dw \\
&= \int_1^x f_{T_1+T_2}(\log w) \frac{1}{w} dw \\
&= \int_0^t f_{T_1+T_2}(s) ds \\
&= F_{T_1+T_2}(t) \\
&= F_{T_1+T_2}(\log x) \\
&= [\mathcal{T}^{-1} F_{T_1+T_2}] (x)
\end{aligned}$$

e

$$\bar{F}_{X_1 X_2}(x) = \bar{F}_{T_1+T_2}(t) = \bar{F}_{T_1+T_2}(\log x) = [\mathcal{T}^{-1} \bar{F}_{T_1+T_2}] (x).$$

Finalmente obtemos, a partir da convolução de Fourier, a convolução de Mellin como

$$\bar{F}_{X_1 X_2}(x) = [\mathcal{T}^{-1}(\bar{F}_{T_1+T_2})](x) = \bar{F}_{T_1+T_2}(\log x) \quad (3.39)$$

e

$$f_{X_1 X_2}(x) = x^{-1} [\mathcal{T}^{-1}(f_{T_1+T_2})](x) = x^{-1} f_{T_1+T_2}(\log x). \quad (3.40)$$

Desde que $\mathcal{T}zero(t) = 0$, então, pelo teorema 3.4.1, o espaço de distribuições absolutamente contínuas forma um monóide abeliano, veja Barata [4], com as operações convolução de Mellin

e MIX.

O operador $\mathcal{T}$ que foi descrito por Collins [7] é uma forte ferramenta que transforma a convolução usual para a convolução de Mellin. Isto é, mesmo que consigamos obter a convolução de Fourier de forma analítica ou de forma numérica, aplicando o operador $\mathcal{T}$ obtemos a convolução de Mellin.

Portanto, o objetivo será preocupar-se em calcular a convolução usual, pois, como consequência, a convolução de Mellin será calculada.

Capítulo 4

Aproximação a distribuições via mistura de exponenciais

O modelo de mistura adotada por Hogg [10] é a de mistura de exponenciais. Isto se deve ao fato que a distribuição exponencial possui o modelo mais simples possível do tempo de vida, no sentido de ser caracterizada por uma taxa de falha homogênea no tempo.

Hogg [10], como exemplo, usou o modelo exponencial, com taxa de falha λ , mostrando uma adaptação flexível na estrutura algébrica por mistura de convoluções.

Se a função de sobrevivência for representada por mistura de exponenciais, aproveitando a estrutura algébrica junto com as boas propriedades do modelo exponencial, poderíamos calcular a convolução aproximando-a por mistura de convoluções exponenciais.

Usamos mistura de exponenciais, descrita por Dufresne [8], como uma técnica de aproximação para a função de sobrevivência, $\bar{F}_T(t) = 1 - F_T(t) = \mathbb{P}(\{T > t\})$, de uma variável aleatória não negativa T , com função de distribuição acumulada $F_T(t)$ e função de densidade de probabilidade $f_T(t)$. A técnica de mistura de exponenciais é calculada a partir da expansão do polinômio de Jacobi “deslocado”, $R_k^{(\alpha, \beta)}(x)$, sob $[0, 1]$ e parâmetros $\alpha, \beta > -1$. Esse polinômio tem diferentes definições equivalentes, usaremos aquela que é expressa em potências de x como

$$R_k^{(\alpha, \beta)}(x) = \sum_{i=0}^k \rho_{ki} x^i, \quad \forall x \in [0, 1],$$

onde

$$\rho_{ki} = \frac{(-1)^k (\beta + 1)_k (-k)_i (k + \alpha + \beta + 1)_i}{(\beta + 1)_i k! i!}$$

e

$$(z)_m = \frac{\Gamma(z + m)}{\Gamma(z)}, \quad z \in \mathbb{R}, \quad m \in \mathbb{N} \cup \{0\},$$

é o símbolo de Pochhammer e $\Gamma(z)$ é a função gama.

A técnica de mistura de exponenciais que aproxima a função de sobrevivência, $\bar{F}_T(t)$, é dada por

$$\bar{F}_T(t) \approx \sum_{j=0}^n a_j e^{-\lambda_j t}, \quad t \geq 0, \quad \lambda_j > 0, \quad j = 0, \dots, n, \quad 1 \leq n < \infty.$$

Para ser mistura de exponenciais, deve satisfazer $a_j > 0$, $j = 0, \dots, n$, porém, numericamente Dufresne [8] permite que alguns dos a_j sejam negativos.

Se queremos aproveitar a operação MIX da estrutura algébrica expressa por Hogg [10], devemos considerar que

$$\sum_{j=0}^n a_j \approx 1.$$

Estes conceitos serão introduzidos nas próximas seções.

4.1 Polinômio de Jacobi

Seja $x \in [-1, 1]$. Para $\alpha, \beta > -1$ definimos o polinômio de Jacobi como

$$P_n^{(\alpha, \beta)}(x) = \frac{(-1)^n}{2^n n!} (1-x)^{-\alpha} (1+x)^{-\beta} \frac{d^n}{dx^n} \left[(1-x)^{\alpha+n} (1+x)^{\beta+n} \right]. \quad (4.1)$$

Para $k \neq m$, $P_k^{(\alpha, \beta)}$ e $P_m^{(\alpha, \beta)}$ são ortogonais sob $[-1, 1]$ para uma função de peso definida como $(1-x)^\alpha (1+x)^\beta$.

Podemos redefinir o polinômio de Jacobi usando a seguinte igualdade:

$$\begin{aligned} & \frac{d^n}{dx^n} \left[(1-x)^{\alpha+n} (1+x)^{\beta+n} \right] = \\ &= \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} \frac{\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{\Gamma(\alpha+i+1)\Gamma(\beta+n-i+1)} (1-x)^{\alpha+i} (1+x)^{\beta+n-i}, \end{aligned} \quad (4.2)$$

que pode ser provada usando indução matemática.

Substituímos (4.2) em (4.1) e obtemos

$$\begin{aligned} P_n^{(\alpha, \beta)}(x) &= \frac{(-1)^n}{2^n n!} \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} \frac{\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{\Gamma(\alpha+i+1)\Gamma(\beta+n-i+1)} (1-x)^i (1+x)^{n-i} \\ &= \frac{1}{n!} \sum_{i=0}^n (-1)^i \binom{n}{i} \frac{\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{\Gamma(\alpha+i+1)\Gamma(\beta+n-i+1)} \left(\frac{1-x}{2} \right)^i \left(\frac{1+x}{2} \right)^{n-i}. \end{aligned}$$

Deslocando para $x \in [0, 1]$ obtemos

$$P_n^{(\alpha, \beta)}(2x - 1) = \frac{1}{n!} \sum_{i=0}^n (-1)^i \binom{n}{i} \frac{\Gamma(\alpha + n + 1)\Gamma(\beta + n + 1)}{\Gamma(\alpha + i + 1)\Gamma(\beta + n - i + 1)} (1 - x)^i x^{n-i}.$$

Assim definimos o polinômio de Jacobi deslocado

$$R_n^{(\alpha, \beta)}(x) = P_n^{(\alpha, \beta)}(2x - 1), \quad \forall x \in [0, 1], \quad \alpha, \beta > -1.$$

Para $k \neq m$, $R_k^{(\alpha, \beta)}$ e $R_m^{(\alpha, \beta)}$ são ortogonais sob $[0, 1]$ para uma função de peso definido como

$$\omega^{(\alpha, \beta)}(x) = (1 - x)^\alpha x^\beta.$$

Isto é

$$\int_0^1 R_k^{(\alpha, \beta)}(x) R_m^{(\alpha, \beta)}(x) \omega^{(\alpha, \beta)}(x) dx = h_{k,m}$$

onde $h_{k,m} = 0$ para $k \neq m$ e $h_{k,m} = h_k$ para $k = m$, com

$$h_k = \frac{\Gamma(k + \alpha + 1)\Gamma(k + \beta + 1)}{(2k + \alpha + \beta + 1)k!\Gamma(k + \alpha + \beta + 1)}.$$

Então, $\forall x \in [0, 1]$ e $\alpha, \beta > -1$, o polinômio de Jacobi deslocado formula-se como

$$R_n^{(\alpha, \beta)}(x) = \frac{1}{n!} \sum_{i=0}^n (-1)^i \binom{n}{i} \frac{\Gamma(\alpha + n + 1)\Gamma(\beta + n + 1)}{\Gamma(\alpha + i + 1)\Gamma(\beta + n - i + 1)} (1 - x)^i x^{n-i}. \quad (4.3)$$

Pelo binômio de Newton, podemos ver que

$$(1 - x)^i = \sum_{j=0}^i (-1)^j \binom{i}{j} x^j = \sum_{j=0}^i (-1)^{i-j} \binom{i}{i-j} x^{i-j}.$$

Então

$$(1 - x)^i x^{n-i} = \sum_{j=0}^i (-1)^{i-j} \binom{i}{i-j} x^{n-j}. \quad (4.4)$$

Substituindo (4.4) em (4.3), obtemos uma expressão em potências de x da seguinte maneira

$$\begin{aligned}
R_n^{(\alpha, \beta)}(x) &= \frac{1}{n!} \sum_{i=0}^n (-1)^i \binom{n}{i} \frac{\Gamma(\alpha + n + 1) \Gamma(\beta + n + 1)}{\Gamma(\alpha + i + 1) \Gamma(\beta + n - i + 1)} \left(\sum_{j=0}^i (-1)^{i-j} \binom{i}{i-j} x^{n-j} \right) \\
&= \sum_{i=0}^n \sum_{j=0}^i \frac{(-1)^j}{n!} \binom{n}{i} \binom{i}{j} \frac{\Gamma(\alpha + n + 1) \Gamma(\beta + n + 1)}{\Gamma(\alpha + i + 1) \Gamma(\beta + n - i + 1)} x^{n-j} \\
&= \sum_{j=0}^n \sum_{i=j}^n \frac{(-1)^j}{n!} \binom{n}{i} \binom{i}{j} \frac{\Gamma(\alpha + n + 1) \Gamma(\beta + n + 1)}{\Gamma(\alpha + i + 1) \Gamma(\beta + n - i + 1)} x^{n-j} \\
&= \sum_{j=0}^n \left[\sum_{i=n-j}^n \frac{(-1)^{n-j}}{n!} \binom{n}{i} \binom{i}{n-j} \frac{\Gamma(\alpha + n + 1) \Gamma(\beta + n + 1)}{\Gamma(\alpha + i + 1) \Gamma(\beta + n - i + 1)} \right] x^j.
\end{aligned}$$

Seja

$$\rho_{nj} = \sum_{i=n-j}^n \frac{(-1)^{n-j}}{n!} \binom{n}{i} \binom{i}{n-j} \frac{\Gamma(\alpha + n + 1) \Gamma(\beta + n + 1)}{\Gamma(\alpha + i + 1) \Gamma(\beta + n - i + 1)}. \quad (4.5)$$

Podemos encontrar uma expressão simplificada da equação (4.5). Para obter esta expressão, definimos o símbolo de Pochhammer como

$$(z)_n = \frac{\Gamma(z + n)}{\Gamma(z)}, \quad z \in \mathbb{R}, \quad n \in \mathbb{N} \cup \{0\}.$$

Assim

$$\begin{aligned}
(z)_0 &= 1 \\
(z)_n &= z(z+1) \dots (z+n-1), \quad z \in \mathbb{R}, \quad n \in \mathbb{N}.
\end{aligned}$$

Então

$$\frac{\Gamma(\alpha + n + 1)}{\Gamma(\alpha + i + 1)} = \frac{\Gamma(\alpha + i + 1 + (n - i))}{\Gamma(\alpha + i + 1)} = (\alpha + i + 1)_{[n-i]} = \frac{(\alpha + 1)_n}{(\alpha + 1)_i}$$

e

$$\frac{\Gamma(\beta + n + 1)}{\Gamma(\beta + n - i + 1)} = \frac{\Gamma(\beta + (n - i) + 1 + i)}{\Gamma(\beta + n - i + 1)} = (\beta + n - i + 1)_i = \frac{(\beta + 1)_n}{(\beta + 1)_{[n-i]}}.$$

Podemos usar as duas expressões acima para obter

$$\begin{aligned}\rho_{nj} &= \frac{(-1)^n}{n!} \sum_{i=n-j}^n (-1)^j \frac{n!}{(n-i)!i!} \frac{i!}{(n-j)!(i-(n-j))!} \frac{(\alpha+1)_n}{(\alpha+1)_i} \frac{(\beta+1)_n}{(\beta+1)_{[n-i]}} \\ &= \frac{(-1)^n}{n!j!} \sum_{i=n-j}^n (-1)^j \frac{n!}{(n-j)!} \frac{j!}{(n-i)!(j-(n-i))!} \frac{(\alpha+1)_n}{(\alpha+1)_i} \frac{(\beta+1)_n}{(\beta+1)_{[n-i]}}.\end{aligned}$$

Observe que

$$(-1)^j \frac{n!}{(n-j)!} = (-1)^j n(n-1) \dots (n-j+1) = (-n)(-n+1) \dots (-n+(j-1)) = (-n)_j.$$

Logo,

$$\begin{aligned}\rho_{nj} &= \frac{(-1)^n(-n)_j}{n!j!} \sum_{i=n-j}^n \binom{j}{n-i} \frac{(\alpha+1)_n}{(\alpha+1)_i} \frac{(\beta+1)_n}{(\beta+1)_{[n-i]}} \\ &= \frac{(-1)^n(\beta+1)_n(-n)_j}{(\beta+1)_j n!j!} \left[\sum_{i=n-j}^n \binom{j}{n-i} \frac{(\alpha+1)_n}{(\alpha+1)_i} \frac{(\beta+1)_j}{(\beta+1)_{[n-i]}} \right] \\ &= \frac{(-1)^n(\beta+1)_n(-n)_j}{(\beta+1)_j n!j!} \left[\sum_{i=0}^j \binom{j}{j-i} \frac{(\alpha+1)_n}{(\alpha+1)_{[n-j+i]}} \frac{(\beta+1)_j}{(\beta+1)_{[j-i]}} \right] \\ &= \frac{(-1)^n(\beta+1)_n(-n)_j}{(\beta+1)_j n!j!} \left[\sum_{i=0}^j \binom{j}{i} \frac{\Gamma(\alpha+n+1)}{\Gamma(\alpha+n-j+i+1)} \frac{\Gamma(\beta+j+1)}{\Gamma(\beta+j-i+1)} \right]. \quad (4.6)\end{aligned}$$

Seja

$$\binom{a}{b} = \frac{\Gamma(a+1)}{\Gamma(b+1)\Gamma(a-b+1)},$$

então

$$\begin{aligned}& \sum_{i=0}^j \binom{j}{i} \frac{\Gamma(\alpha+n+1)}{\Gamma(\alpha+n-j+i+1)} \frac{\Gamma(\beta+j+1)}{\Gamma(\beta+j-i+1)} = \\ &= \sum_{i=0}^j \frac{j!}{i!(j-i)!} \frac{\Gamma(\alpha+n+1)}{\Gamma(\alpha+n-j+i+1)} \frac{\Gamma(\beta+j+1)}{\Gamma(\beta+j-i+1)} \\ &= \sum_{i=0}^j \frac{j!}{\Gamma(i+1)\Gamma(j-i+1)} \frac{\Gamma(\alpha+n+1)}{\Gamma(\alpha+n-j+i+1)} \frac{\Gamma(\beta+j+1)}{\Gamma(\beta+j-i+1)}\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^j j! \frac{\Gamma(\alpha + n + 1)}{\Gamma(j - i + 1)\Gamma(\alpha + n - j + i + 1)} \frac{\Gamma(\beta + j + 1)}{\Gamma(i + 1)\Gamma(\beta + j - i + 1)} \\
&= j! \sum_{i=0}^j \binom{\alpha + n}{j - i} \binom{\beta + j}{i} \\
&= j! \binom{\alpha + \beta + n + j}{j} \left[\sum_{i=0}^j \frac{\binom{\alpha + n}{j - i} \binom{\beta + j}{i}}{\binom{\alpha + \beta + n + j}{j}} \right] \\
&= \frac{\Gamma(\alpha + \beta + n + j + 1)}{\Gamma(\alpha + \beta + n + 1)} \left[\sum_{i=0}^j \mathbb{P}(\{X = i\}) \right], \quad X \sim \text{Hipergeométrica}(\beta + j, \alpha + n, j) \\
&= (\alpha + \beta + n + 1)_j \cdot [1], \quad \text{por ser } \sum_{i=0}^j \mathbb{P}(\{X = i\}) = 1 \\
&= (\alpha + \beta + n + 1)_j
\end{aligned} \tag{4.7}$$

Assim, substituindo (4.7) em (4.6) obtemos

$$\rho_{nj} = \frac{(-1)^n (\beta + 1)_n (-n)_j}{(\beta + 1)_j n! j!} [(\alpha + \beta + n + 1)_j]$$

ou

$$\rho_{nj} = \frac{(-1)^n (\beta + 1)_n (-n)_j (\alpha + \beta + n + 1)_j}{(\beta + 1)_j n! j!}. \tag{4.8}$$

Portanto o polinômio de Jacobi deslocado pode ser re-escrito como

$$R_n^{(\alpha, \beta)}(x) = \sum_{j=0}^n \rho_{nj} x^j, \quad \forall x \in [0, 1], \quad \alpha, \beta > -1. \tag{4.9}$$

Para o espaço

$$\begin{aligned}
L^2 \left((0, 1), \omega^{(\alpha, \beta)} \right) &= \left\{ f \text{ mensurável} / \int_0^1 [f(x)]^2 \omega^{(\alpha, \beta)}(x) dx < \infty \right\} \\
&= \left\{ f \text{ mensurável} / \int_0^1 [f(x)]^2 (1 - x)^\alpha x^\beta dx < \infty \right\},
\end{aligned}$$

com produto interno para a função de peso $\omega^{(\alpha, \beta)}(x)$ sob $[0, 1]$,

$$\langle f, g \rangle = \int_0^1 f(x) g(x) \omega^{(\alpha, \beta)}(x) dx, \quad \forall f, g \in L^2 \left((0, 1), \omega^{(\alpha, \beta)} \right)$$

e

$$\|f\|_{L^2((0,1),\omega^{(\alpha,\beta)})}^2 = \int_0^1 [f(x)]^2 \omega^{(\alpha,\beta)}(x) dx,$$

possui como base ortogonal a sequência de polinômios de Jacobi deslocado, $\{R_n^{(\alpha,\beta)}(x)\}_{n \in \mathbb{N}}$, sob $[0, 1]$. Assim, como visto no capítulo 3, podemos construir uma série polinômica que aproxime as funções que pertencem a $L^2((0, 1), \omega^{(\alpha,\beta)})$. Dufresne [8] apresenta os seguintes teoremas.

Teorema 4.1.1 *Sejam $f : [0, 1] \rightarrow \mathbb{R}$ contínua e f' contínua por partes e*

$$S_n(x) = \sum_{k=0}^n c_k R_k^{(\alpha,\beta)}(x).$$

Então $S_n \rightarrow f$ quando $n \rightarrow \infty$ uniformemente em $(\epsilon, 1 - \epsilon)$, $\epsilon \in (0, 1)$, onde

$$c_k = \frac{1}{h_k} \int_0^1 f(x) R_k^{(\alpha,\beta)}(x) (1-x)^\alpha x^\beta dx$$

e

$$h_k = \int_0^1 \left[R_k^{(\alpha,\beta)}(x) \right]^2 (1-x)^\alpha x^\beta dx = \frac{\Gamma(k+\alpha+1)\Gamma(k+\beta+1)}{(2k+\alpha+\beta+1)k!\Gamma(k+\alpha+\beta+1)}.$$

Nos pontos de descontinuidade

$$S_n(x) \rightarrow \frac{1}{2} [f(x^-) + f(x^+)].$$

Teorema 4.1.2 *Se $f \in L^2((0, 1), \omega^{(\alpha,\beta)})$, então $S_n \rightarrow f$ quando $n \rightarrow \infty$ em $L^2((0, 1), \omega^{(\alpha,\beta)})$. Ou seja*

$$\|f - S_n\|_{L^2((0,1),\omega^{(\alpha,\beta)})} \rightarrow 0, \quad n \rightarrow \infty.$$

4.2 Misturas de exponenciais usando o polinômio de Jacobi

A aproximação de mistura de exponenciais descrita por Dufresne em [8], é dada pela seguinte equação

$$\bar{F}_T(t) = 1 - F_T(t) = \mathbb{P}(\{T > t\}) \approx \sum_{j=0}^n a_j e^{-\lambda_j t}, \quad t \geq 0. \quad (4.10)$$

Para aproximar a função de sobrevivência $\bar{F}_T(t)$, $t \geq 0$, usando a expansão do polinômio de Jacobi “deslocado” sob $[0, 1]$, devemos transformar a variável $t \geq 0$ para uma variável $x \in [0, 1]$. Para isto, se $X = e^{-rT}$ com $r > 0$ vemos que

$$F_X(x) = \mathbb{P}(\{X < x\}) = \mathbb{P}(\{e^{-rT} < x\}) = \mathbb{P}(\{T > -\frac{1}{r} \log(x)\}) = \bar{F}_T(-\frac{1}{r} \log(x)),$$

onde $t = -\frac{1}{r} \log(x)$. Quando $t \rightarrow \infty$ implica $x \rightarrow 0$ e quando $t \rightarrow 0$ implica $x \rightarrow 1$, assim $x \in (0, 1]$ e $F_X(0) = 0$ por ser $\bar{F}_T(\infty) = 0$.

Se encontrarmos constantes α, β, p e $\{b_k\}$ tal que (informalmente)

$$F_X^{[p]}(x) = x^{-p} F_X(x) = \sum_{k=0}^{\infty} b_k R_k^{(\alpha, \beta)}(x), \quad \forall x \in (0, 1], \quad (4.11)$$

ou

$$F_X(x) = x^p \sum_{k=0}^{\infty} b_k R_k^{(\alpha, \beta)}(x), \quad \forall x \in (0, 1],$$

por (4.9) encontramos uma nova expressão em potências de x como

$$F_X(x) = x^p \sum_{k=0}^{\infty} b_k \left(\sum_{j=0}^k \rho_{kj} x^j \right).$$

Para n suficientemente grande obtemos a seguinte aproximação

$$F_X(x) \approx x^p \sum_{k=0}^n b_k \left(\sum_{j=0}^k \rho_{kj} x^j \right).$$

Desde que $x = e^{-rt}$ e $F_X(x) = \bar{F}_T(-\frac{1}{r} \log(x)) = \bar{F}_T(t)$, como vimos previamente, obtemos a aproximação via mistura de exponenciais como segue

$$\begin{aligned}
\bar{F}_T(t) &\approx e^{-prt} \sum_{k=0}^n b_k \left(\sum_{j=0}^k \rho_{kj} e^{-jrt} \right) \\
&= e^{-prt} \sum_{j=0}^n \left(\sum_{k=j}^n b_k \rho_{kj} \right) e^{-jrt} \\
&= \sum_{j=0}^n \left(\sum_{k=j}^n b_k \rho_{kj} \right) e^{-(j+p)rt}.
\end{aligned}$$

Definindo

$$\lambda_j = (j+p)r, \quad j = 0, \dots, n \quad (4.12)$$

para $p > 0$, dado que $r > 0$, mantemos $\lambda_j > 0, \forall j = 0, \dots, n$ e

$$a_j = \sum_{k=j}^n b_k \rho_{kj}, \quad j = 0, \dots, n \quad (4.13)$$

obtemos a equação (4.10) desejada.

Para $F_X^{[p]}(x) = x^{-p} F_X(x)$, as constantes $\{b_k\}$, dada pela fórmula (4.11), pode ser reformulada pela seguinte expressão

$$\begin{aligned}
b_k &= \frac{1}{h_k} \int_0^1 F_X^{[p]}(x) R_k^{(\alpha, \beta)}(x) (1-x)^\alpha x^\beta dx \\
&= \frac{1}{h_k} \int_0^1 x^{-p} F_X(x) \left[\sum_{j=0}^k \rho_{kj} x^j \right] (1-x)^\alpha x^\beta dx \\
&= \frac{1}{h_k} \int_0^\infty e^{prt} \bar{F}_T(t) \left[\sum_{j=0}^k \rho_{kj} e^{-jrt} \right] (1-e^{-rt})^\alpha e^{-\beta rt} d(e^{-rt}) \\
&= \frac{r}{h_k} \int_0^\infty e^{-(\beta-p+1)rt} (1-e^{-rt})^\alpha \left[\sum_{j=0}^k \rho_{kj} e^{-jrt} \right] \bar{F}_T(t) dt \\
&= \frac{r}{h_k} \sum_{j=0}^k \rho_{kj} \left[\int_0^\infty e^{-(\beta-p+1)rt} (1-e^{-rt})^\alpha e^{-jrt} \bar{F}_T(t) dt \right] \\
&= \frac{r}{h_k} \sum_{j=0}^k \rho_{kj} \left[\int_0^\infty e^{-(\beta-p+j+1)rt} (1-e^{-rt})^\alpha \bar{F}_T(t) dt \right] \quad (4.14)
\end{aligned}$$

com

$$h_k = \frac{\Gamma(k + \alpha + 1)\Gamma(k + \beta + 1)}{(2k + \alpha + \beta + 1)k!\Gamma(k + \alpha + \beta + 1)}. \quad (4.15)$$

As equações (4.14) e (4.15) são dadas pelo teorema 4.1.1. Como consequência, obtem-se o seguinte resultado (veja Dufresne [8]).

Resultado 2 *Suponha $\alpha, \beta > -1$ e $\bar{F}_T(\cdot)$ contínua sob $[0, \infty)$ e a que função $e^{prt}\bar{F}_T(t)$ possui limite finito quando $t \rightarrow \infty$ para algum $p \in \mathbb{R}$. Então, $b_k < \infty, \forall k \in \mathbb{N}$ e*

$$\bar{F}_T(t) = e^{-prt} \sum_{k=0}^{\infty} b_k R_k^{(\alpha, \beta)}(e^{-rt}), \quad \forall t \in (0, \infty). \quad (4.16)$$

A convergência é uniforme sob $[a, b]$ para $0 < a < b < \infty$.

Observe que

$$\begin{aligned} \|F_X^{[p]}\|_{L^2((0,1), \omega^{(\alpha, \beta)})}^2 &= \int_0^1 [F_X^{[p]}(x)]^2 \omega^{(\alpha, \beta)}(x) dx \\ &= r \int_0^\infty [F_X^{[p]}(e^{-rt})]^2 \omega^{(\alpha, \beta)}(e^{-rt}) e^{-rt} dt \\ &= r \int_0^\infty [e^{prt} \bar{F}_T(t)]^2 (1 - e^{-rt})^\alpha e^{-\beta rt} e^{-rt} dt \\ &= r \int_0^\infty e^{-(\beta+1-2p)rt} (1 - e^{-rt})^\alpha [\bar{F}_T(t)]^2 dt. \end{aligned} \quad (4.17)$$

Para $p < (\beta + 1)/2$ temos $\|F_X^{[p]}\|_{L^2((0,1), \omega^{(\alpha, \beta)})} < \infty$.

Também

$$\begin{aligned} &\left\| F_X^{[p]} - \sum_{k=0}^n b_k R_k^{(\alpha, \beta)} \right\|_{L^2((0,1), \omega^{(\alpha, \beta)})}^2 = \\ &= r \int_0^\infty \left[e^{prt} \bar{F}_T(t) - \sum_{k=0}^n b_k R_k^{(\alpha, \beta)}(e^{-rt}) \right]^2 (1 - e^{-rt})^\alpha e^{-\beta rt} e^{-rt} dt \\ &= r \int_0^\infty e^{-(\beta+1-2p)rt} (1 - e^{-rt})^\alpha \left[\bar{F}_T(t) - e^{-prt} \sum_{k=0}^n b_k R_k^{(\alpha, \beta)}(e^{-rt}) \right]^2 dt. \end{aligned} \quad (4.18)$$

Como consequência de (4.17), (4.18) e o teorema 4.1.2, obtemos o seguinte resultado.

Resultado 3 *Suponha $\alpha, \beta > -1$ e que para algum $p \in \mathbb{R}$ e $r > 0$ tal que*

$$\int_0^\infty e^{-(\beta+1-2p)rt} (1 - e^{-rt})^\alpha [\bar{F}_T(t)]^2 dt < \infty. \quad (4.19)$$

Então

$$\int_0^\infty e^{-(\beta+1-2p)rt} (1 - e^{-rt})^\alpha \left[\bar{F}_T(t) - e^{-prt} \sum_{k=0}^n b_k R_k^{(\alpha, \beta)}(e^{-rt}) \right]^2 dt \longrightarrow 0. \text{ quando } n \rightarrow \infty.$$

A convergência (4.16) é mantida quase certamente, com b_k dado em (4.14).

Para $\alpha = 0$ podemos redefinir (4.14) e (4.15) como

$$b_k = \frac{r}{h_k} \sum_{j=0}^k \rho_{kj} \left[\int_0^\infty e^{-(\beta-p+j+1)rt} \bar{F}_T(t) dt \right] \quad (4.20)$$

e

$$h_k = \frac{1}{(2k + \beta + 1)}. \quad (4.21)$$

De (4.20), vemos que

$$\begin{aligned} & r \int_0^\infty e^{-(\beta-p+j+1)rt} \bar{F}_T(t) dt = \\ &= r \left[\frac{-1}{(\beta-p+j+1)r} \int_0^\infty \bar{F}_T(t) d(e^{-(\beta-p+j+1)rt}) \right] \\ &= \frac{-1}{(\beta-p+j+1)} \int_0^\infty \bar{F}_T(t) d(e^{-(\beta-p+j+1)rt}) \\ &= \frac{-1}{(\beta-p+j+1)} \left\{ \left[\bar{F}_T(t) e^{-(\beta-p+j+1)rt} \right]_0^\infty - \int_0^\infty e^{-(\beta-p+j+1)rt} d\bar{F}_T(t) \right\} \\ &= \frac{-1}{(\beta-p+j+1)} \left\{ -1 + \int_0^\infty e^{-(\beta-p+j+1)rt} dF_T(t) \right\} \\ &= \frac{1}{(\beta-p+j+1)} \left\{ 1 - \int_0^\infty e^{-(\beta-p+j+1)rt} dF_T(t) \right\} \\ &= \frac{1}{(\beta-p+j+1)} \left[1 - \mathbb{E} \left(e^{-(\beta-p+j+1)rT} \right) \right]. \end{aligned} \quad (4.22)$$

Substituindo (4.21) e (4.22) em (4.20) temos

$$\begin{aligned}
b_k &= (2k + \beta + 1)r \sum_{j=0}^k \rho_{kj} \left[\int_0^\infty e^{-(\beta-p+j+1)rt} \bar{F}_T(t) dt \right] \\
&= (2k + \beta + 1) \sum_{j=0}^k \rho_{kj} \left[r \int_0^\infty e^{-(\beta-p+j+1)rt} \bar{F}_T(t) dt \right] \\
&= (2k + \beta + 1) \sum_{j=0}^k \frac{\rho_{kj}}{(\beta - p + j + 1)} \left[1 - \mathbb{E} \left(e^{-(\beta-p+j+1)rT} \right) \right]. \tag{4.23}
\end{aligned}$$

Observe que a aproximação para a distribuição de certa variável aleatória depende fortemente da sua Transformada de Laplace, expressa em termos de esperança. Para Transformadas de Laplace sem resultado analítico isso pode ser computado numericamente via método de Monte Carlo.

Também, para $\alpha = 0$, (4.8) é reformulada como

$$\rho_{kj} = \frac{(-1)^k (\beta + 1)_k (-k)_j (\beta + k + 1)_j}{(\beta + 1)_j k! j!}. \tag{4.24}$$

A fórmula (4.24) pode ser vista da seguinte maneira:

$$\begin{aligned}
\rho_{kj} &= \frac{(-1)^k (\beta + 1)_k (-k)_j (\beta + k + 1)_j}{(\beta + 1)_j k! j!} \\
&= (-1)^k \frac{(\beta + 1)_k}{k!} \frac{(-k)_j}{j!} \frac{(\beta + k + 1)_j}{(\beta + 1)_j} \\
&= (-1)^k \left[\left(\frac{\beta + 1}{1} \right) \left(\frac{\beta + 2}{2} \right) \cdots \left(\frac{\beta + k}{k} \right) \right] \\
&\quad \cdot \left[(-1)^j \frac{k!}{(k-j)! j!} \right] \left[\left(\frac{\beta + k + 1}{\beta + 1} \right) \left(\frac{\beta + k + 2}{\beta + 2} \right) \cdots \left(\frac{\beta + k + j}{\beta + j} \right) \right] \\
&= (-1)^k \left[\left(\frac{\beta}{1} + 1 \right) \left(\frac{\beta}{2} + 1 \right) \cdots \left(\frac{\beta}{k} + 1 \right) \right] \\
&\quad \cdot \left[(-1)^j \binom{k}{j} \right] \left[\left(\frac{k}{\beta + 1} + 1 \right) \left(\frac{k}{\beta + 2} + 1 \right) \cdots \left(\frac{k}{\beta + j} + 1 \right) \right] \\
&= (-1)^k (-1)^j \left[\left(\frac{\beta}{1} + 1 \right) \left(\frac{\beta}{2} + 1 \right) \cdots \left(\frac{\beta}{k} + 1 \right) \right] \\
&\quad \cdot \left[\left(\frac{k}{\beta + 1} + 1 \right) \left(\frac{k}{\beta + 2} + 1 \right) \cdots \left(\frac{k}{\beta + j} + 1 \right) \right] \binom{k}{j}. \tag{4.25}
\end{aligned}$$

O procedimento para determinar os valores de ρ pode ser sintetizado no seguinte pseudocódigo:

Algoritmo 7 Subrotina cálculo de ρ

```

1: Dados:  $n, \beta$ .
2:  $\rho(0, 0) = 1, sinalk = -1, x = 1$ .
3: Para  $k = 1 : n$  fazer:
4:    $x = x * \left(\frac{\beta}{k} + 1\right)$ 
5:    $\rho(k, 0) = sinalk * x$ 
6:    $y = 1$ 
7:    $sinalj = -1$ 
8:   Para  $j = 1 : k$  fazer:
9:      $y = y * \left(\frac{k}{\beta+j} + 1\right)$ 
10:    faça uma subrotina para obter  $comb(k, j) = \binom{k}{j}$ 
11:     $\rho(k, j) = sinalk * sinalj * x * y * comb(k, j)$ 
12:     $sinalj = -sinalj$ 
13:  Fim
14:   $sinalk = -sinalk$ 
15: Fim.
```

Também usaremos (4.23) para o cálculo de b e (4.12) para o cálculo de λ com a seguinte subrotina:

Algoritmo 8 Subrotina cálculo de b e λ

```

1: Dados:  $n, \beta, p, r, chooseVarT$ .
2: Para  $k = 0 : n$  fazer:
3:    $x = 2 * k + \beta + 1$ 
4:    $y = 0$ 
5:   Para  $j = 0 : k$  fazer:
6:      $\theta = (\beta - p + j + 1) * r$ 
7:     faça uma subrotina para escolher a v.a.  $T$  usando como criterio  $chooseVarT$ 
8:     faça uma subrotina para obter  $esp(\theta) = \mathbb{E}[e^{-\theta T}]$ 
9:      $y = y + \rho(k, j) * [1 - esp(\theta)] * \frac{r}{\theta}$ 
10:  Fim
11:   $b(k) = x * y$ 
12:   $\lambda(k) = (k + p) * r$ 
13: Fim.
```

Portanto, após serem encontrados ρ e b , podemos encontrar a de (4.13) com a seguinte subrotina:

Algoritmo 9 Subrotina cálculo de a

```

1: Dados:  $n, b, \rho$ .
2: Para  $j = 0 : n$  fazer:
3:      $a(j) = 0$ 
4:     Para  $k = j : n$  fazer:
5:          $a(j) = a(j) + b(k) * \rho(k, j)$ 
6:     Fim
7: Fim.
```

Sendo assim podemos aproximar a função de densidade como:

$$f_T(t) = \sum_{j=0}^n a_j \lambda_j e^{-\lambda_j t}, \quad t \geq 0 \quad (4.26)$$

com λ_j e a_j para $j = 1 \dots n$ definidos em (4.12) e (4.13).

Assim, de (4.10) e (4.26) obtemos a seguinte subrotina de aproximação:

Algoritmo 10 Subrotina aproximação de f_T , F_T e $\bar{F}_T$

```

1: Dados:  $n, a, \lambda$  e  $t$ .
2:  $fT = 0, barFT = 0$ .
3: Para  $j = 0 : n$  fazer:
4:      $fT = fT + a(j) * \lambda(j) * \exp(-\lambda(j) * t)$ 
5:      $barFT = barFT + a(j) * \exp(-\lambda(j) * t)$ 
6: Fim
7:  $FT = 1 - barFT$ .
```

4.2.1 Usando a estrutura algébrica de Hogg para as convoluções

Recorde que, considerando a operação MIX da estrutura algébrica de monóide expressa por Hogg [10], uma condição a mais de aproximação para n suficientemente grande seria

$$\sum_{j=0}^n a_j \approx 1. \quad (4.27)$$

Hogg [10], simplifica convenientemente a fórmula (3.27) para descrever a convolução da função de sobrevivência exponencial da seguinte maneira

$$Exp(\lambda_1) \oplus Exp(\lambda_2) \iff \bar{F}_{T_1+T_2}(t) = \bar{F}_{T_1}(t) + \int_0^t \bar{F}_{T_2}(t-s) f_{T_1}(s) ds$$

e descreve o MIX da fórmula (3.30) com a_1, a_2 sendo as probabilidades a priori de mistura como

$$MIX\{a_1 Exp(\lambda_1), a_2 Exp(\lambda_2)\} \iff MIX[a_1 \bar{F}_{T_1}(t), a_2 \bar{F}_{T_2}(t)] = a_1 \bar{F}_{T_1}(t) + a_2 \bar{F}_{T_2}(t).$$

Substituindo pela fórmula exponencial, obtemos para $\lambda_1 \neq \lambda_2$

$$Exp(\lambda_1) \oplus Exp(\lambda_2) \iff \bar{F}_{T_1+T_2}(t) = \frac{\lambda_2 e^{-\lambda_1 t}}{\lambda_2 - \lambda_1} + \frac{\lambda_1 e^{-\lambda_2 t}}{\lambda_1 - \lambda_2} \quad (4.28)$$

e

$$MIX\{a_1 Exp(\lambda_1), a_2 Exp(\lambda_2)\} \iff MIX[a_1 e^{-\lambda_1 t}, a_2 e^{-\lambda_2 t}] = a_1 e^{-\lambda_1 t} + a_2 e^{-\lambda_2 t}. \quad (4.29)$$

Podemos estender para N variáveis aleatórias considerando todos seus λ_i diferentes $\forall i$ obtendo

$$\bigoplus_{i=1}^N Exp(\lambda_i) \iff \sum_{i=1}^N \left[\prod_{j=1, j \neq i}^N \left(\frac{\lambda_j}{\lambda_j - \lambda_i} \right) \right] e^{-\lambda_i t}.$$

Portanto, podemos expressar a aproximação de mistura de exponenciais (4.10) da seguinte maneira

$$T_i \iff MIX_{k_i=0}^{n_i} [a_{i,k_i} Exp(\lambda_{i,k_i})] \quad \forall i = 1, \dots, N$$

onde n_i é o grau do polinômio de aproximação e $k_i = 0, \dots, n_i$ seu correspondente conjunto de índices da i -ésima variável aleatória não negativa T_i .

Para $N = 2$ obtemos

$$\begin{aligned} T_1 + T_2 &\iff \left\{ MIX_{k_1=0}^{n_1} [a_{1,k_1} Exp(\lambda_{1,k_1})] \right\} \oplus \left\{ MIX_{k_2=0}^{n_2} [a_{2,k_2} Exp(\lambda_{2,k_2})] \right\} \\ &\iff MIX_{k_1, k_2=0}^{n_1, n_2} \{ (a_{1,k_1} a_{2,k_2}) [Exp(\lambda_{1,k_1}) \oplus Exp(\lambda_{2,k_2})] \}. \end{aligned}$$

Para o caso geral obtemos

$$T_1 + T_2 + \dots + T_N \iff MIX_{k_1, \dots, k_N=0}^{n_1, \dots, n_N} \left\{ \left(\prod_{j=1}^N a_{j,k_j} \right) \left[\bigoplus_{i=1}^N Exp(\lambda_{i,k_i}) \right] \right\},$$

Segue que, a convolução que representa a soma de variáveis aleatórias independentes é aproximada por mistura de convoluções exponenciais. As fórmulas seguem de (4.28) e (4.29) para obter

$$\bar{F}_{T_1+T_2}(t) = \sum_{k_1, k_2=0}^{n_1, n_2} \left\{ (a_{1, k_1} a_{2, k_2}) \left[\frac{\lambda_{2, k_2} e^{-\lambda_{1, k_1} t}}{\lambda_{2, k_2} - \lambda_{1, k_1}} + \frac{\lambda_{1, k_1} e^{-\lambda_{2, k_2} t}}{\lambda_{1, k_1} - \lambda_{2, k_2}} \right] \right\}$$

e

$$\bar{F}_{T_1+T_2+\dots+T_N}(t) = \sum_{k_1, \dots, k_N=0}^{n_1, \dots, n_N} \left\{ \left(\prod_{j=1}^N a_{j, k_j} \right) \left(\sum_{i=1}^N \left[\prod_{j=1, j \neq i}^N \left(\frac{\lambda_{j, k_j}}{\lambda_{j, k_j} - \lambda_{i, k_i}} \right) \right] e^{-\lambda_{i, k_i} t} \right) \right\}.$$

Derivando com respeito a t temos

$$f_{T_1+T_2}(t) = \sum_{k_1, k_2=0}^{n_1, n_2} \left\{ (a_{1, k_1} \lambda_{1, k_1} a_{2, k_2} \lambda_{2, k_2}) \left[\frac{e^{-\lambda_{1, k_1} t}}{\lambda_{2, k_2} - \lambda_{1, k_1}} + \frac{e^{-\lambda_{2, k_2} t}}{\lambda_{1, k_1} - \lambda_{2, k_2}} \right] \right\}$$

e

$$f_{T_1+T_2+\dots+T_N}(t) = \sum_{k_1, \dots, k_N=0}^{n_1, \dots, n_N} \left\{ \left(\prod_{j=1}^N a_{j, k_j} \lambda_{j, k_j} \right) \left(\sum_{i=1}^N \left[\prod_{j=1, j \neq i}^N \left(\frac{1}{\lambda_{j, k_j} - \lambda_{i, k_i}} \right) \right] e^{-\lambda_{i, k_i} t} \right) \right\}.$$

As fórmulas obtidas acima, usando mistura de exponenciais, foram encontradas considerando todos os valores dos lambdas diferentes entre si. Porém, pode existir o problema de ter algum caso em que existam pelo menos dois valores dos lambdas iguais. Um caso em que pode existir esse problema, e geralmente é um caso muito usado, é quando queremos calcular a convolução de variáveis aleatórias independentes e identicamente distribuídas.

Primeiro, consideremos o caso de calcular a convolução de duas variáveis aleatórias independentes onde as suas respectivas aproximações, via misturas de exponenciais, possuem os valores dos lambdas iguais e os coeficientes diferentes, ou seja:

$$\begin{aligned} \bar{F}_{T_2}(t-s) f_{T_1}(s) &= \left[\sum_{i=0}^m a_{2, i} e^{-\lambda_i(t-s)} \right] \left[\sum_{j=0}^n a_{1, j} \lambda_j e^{-\lambda_j s} \right] \\ &= \sum_{i=0}^{n \wedge m} a_{2, i} a_{1, i} \lambda_i e^{-\lambda_i t} + \sum_{i \neq j} a_{2, i} a_{1, j} \lambda_j e^{-[\lambda_j s + \lambda_i(t-s)]} \\ &= \sum_{i=0}^{n \wedge m} a_{2, i} a_{1, i} \lambda_i e^{-\lambda_i t} + \sum_{i \neq j} a_{2, i} a_{1, j} \lambda_j e^{-\lambda_i t} e^{-[\lambda_j - \lambda_i] s}, \end{aligned}$$

onde a fórmula obtida acima foi dada usando as aproximações (4.10) e (4.26). Depois para o cálculo

da convolução, usaremos as fórmulas (3.27) e (3.28). Então

$$\begin{aligned} \int_0^t \bar{F}_{T_2}(t-s)f_{T_1}(s)ds &= \sum_{i=0}^{n \wedge m} a_{2,i}a_{1,i}\lambda_i t e^{-\lambda_i t} + \sum_{i \neq j} a_{2,i}a_{1,j}\lambda_j e^{-\lambda_i t} \int_0^t e^{-[\lambda_j - \lambda_i]s} ds \\ &= \sum_{i=0}^{n \wedge m} a_{2,i}a_{1,i}\lambda_i t e^{-\lambda_i t} + \sum_{i \neq j} a_{2,i}a_{1,j}\lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right]. \end{aligned}$$

Logo,

$$\bar{F}_{T_1+T_2}(t) = \sum_{i=0}^n a_{1,i}e^{-\lambda_i t} + \sum_{i=0}^{n \wedge m} a_{2,i}a_{1,i}\lambda_i t e^{-\lambda_i t} + \sum_{i \neq j} a_{2,i}a_{1,j}\lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \quad (4.30)$$

onde $n \wedge m = \min(n, m)$.

Analogamente

$$\begin{aligned} f_{T_2}(t-s)f_{T_1}(s) &= \left[\sum_{i=0}^m a_{2,i}\lambda_i e^{-\lambda_i(t-s)} \right] \left[\sum_{j=0}^n a_{1,j}\lambda_j e^{-\lambda_j s} \right] \\ &= \sum_{i=0}^{n \wedge m} a_{2,i}a_{1,i}\lambda_i^2 e^{-\lambda_i t} + \sum_{i \neq j} a_{2,i}\lambda_i a_{1,j}\lambda_j e^{-[\lambda_j s + \lambda_i(t-s)]} \\ &= \sum_{i=0}^{n \wedge m} a_{2,i}a_{1,i}\lambda_i^2 e^{-\lambda_i t} + \sum_{i \neq j} a_{2,i}\lambda_i a_{1,j}\lambda_j e^{-\lambda_i t} e^{-[\lambda_j - \lambda_i]s}. \end{aligned}$$

Então

$$\begin{aligned} \int_0^t f_{T_2}(t-s)f_{T_1}(s)ds &= \sum_{i=0}^{n \wedge m} a_{2,i}a_{1,i}\lambda_i^2 t e^{-\lambda_i t} + \sum_{i \neq j} a_{2,i}\lambda_i a_{1,j}\lambda_j e^{-\lambda_i t} \int_0^t e^{-[\lambda_j - \lambda_i]s} ds \\ &= \sum_{i=0}^{n \wedge m} a_{2,i}a_{1,i}\lambda_i^2 t e^{-\lambda_i t} + \sum_{i \neq j} a_{2,i}\lambda_i a_{1,j}\lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right]. \end{aligned}$$

Assim

$$f_{T_1+T_2}(t) = \sum_{i=0}^{n \wedge m} a_{2,i}a_{1,i}\lambda_i^2 t e^{-\lambda_i t} + \sum_{i \neq j} a_{2,i}\lambda_i a_{1,j}\lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right]. \quad (4.31)$$

Agora, considerando o caso de duas variáveis aleatórias i.i.d., calcularemos, via mistura de exponenciais, a convolução e aproximaremos numericamente.

O próximo passo será aplicar o operador $\mathcal{T}$, descrito por Collins [7], para finalmente obter a convolução de Mellin.

Usaremos as aproximações (4.10) e (4.26) para aproximar as convoluções das fórmulas (3.27) e (3.28). Então

$$\begin{aligned}\bar{F}_{T_2}(t-s)f_{T_1}(s) &= \left[\sum_{i=0}^n a_i e^{-\lambda_i(t-s)} \right] \left[\sum_{j=0}^n a_j \lambda_j e^{-\lambda_j s} \right] \\ &= \sum_{i=0}^n a_i^2 \lambda_i e^{-\lambda_i t} + \sum_{i \neq j} a_i a_j \lambda_j e^{-[\lambda_j s + \lambda_i(t-s)]} \\ &= \sum_{i=0}^n a_i^2 \lambda_i e^{-\lambda_i t} + \sum_{i \neq j} a_i a_j \lambda_j e^{-\lambda_i t} e^{-[\lambda_j - \lambda_i]s}.\end{aligned}$$

Logo,

$$\begin{aligned}\int_0^t \bar{F}_{T_2}(t-s)f_{T_1}(s)ds &= \sum_{i=0}^n a_i^2 \lambda_i t e^{-\lambda_i t} + \sum_{i \neq j} a_i a_j \lambda_j e^{-\lambda_i t} \int_0^t e^{-[\lambda_j - \lambda_i]s} ds \\ &= \sum_{i=0}^n a_i^2 \lambda_i t e^{-\lambda_i t} + \sum_{i \neq j} a_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right].\end{aligned}$$

Desta forma

$$\bar{F}_{T_1+T_2}(t) = \sum_{i=0}^n a_i e^{-\lambda_i t} + \sum_{i=0}^n a_i^2 \lambda_i t e^{-\lambda_i t} + \sum_{i \neq j} a_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \quad (4.32)$$

ou

$$\begin{aligned}\bar{F}_{T_1+T_2}(t) &= \sum_{i=0}^n a_i e^{-\lambda_i t} + \sum_{i=0}^n a_i^2 \lambda_i t e^{-\lambda_i t} + \sum_{i < j} a_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \\ &\quad + \sum_{i > j} a_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \\ &= \sum_{i=0}^n a_i e^{-\lambda_i t} + \sum_{i=0}^n a_i^2 \lambda_i t e^{-\lambda_i t} + \sum_{i < j} a_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \\ &\quad + \sum_{i < j} a_j a_i \lambda_i \left[\frac{e^{-\lambda_j t} - e^{-\lambda_i t}}{\lambda_i - \lambda_j} \right] \\ &= \sum_{i=0}^n a_i e^{-\lambda_i t} + \sum_{i=0}^n a_i^2 \lambda_i t e^{-\lambda_i t} + \sum_{i < j} a_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right]\end{aligned}$$

$$+ \sum_{i < j} a_i a_j \lambda_i \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right].$$

Portanto

$$\bar{F}_{T_1+T_2}(t) = \sum_{i=0}^n (1 + a_i \lambda_i t) a_i e^{-\lambda_i t} + \sum_{i < j} a_i a_j \left[\frac{\lambda_j + \lambda_i}{\lambda_j - \lambda_i} \right] (e^{-\lambda_i t} - e^{-\lambda_j t}) \quad (4.33)$$

e

$$\begin{aligned} f_{T_2}(t-s)f_{T_1}(s) &= \left[\sum_{i=0}^n a_i \lambda_i e^{-\lambda_i(t-s)} \right] \left[\sum_{j=0}^n a_j \lambda_j e^{-\lambda_j s} \right] \\ &= \sum_{i=0}^n (a_i \lambda_i)^2 e^{-\lambda_i t} + \sum_{i \neq j} a_i \lambda_i a_j \lambda_j e^{-[\lambda_j s + \lambda_i(t-s)]} \\ &= \sum_{i=0}^n (a_i \lambda_i)^2 e^{-\lambda_i t} + \sum_{i \neq j} a_i \lambda_i a_j \lambda_j e^{-\lambda_i t} e^{-[\lambda_j - \lambda_i]s}. \end{aligned}$$

Então

$$\begin{aligned} \int_0^t f_{T_2}(t-s)f_{T_1}(s)ds &= \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} + \sum_{i \neq j} a_i \lambda_i a_j \lambda_j e^{-\lambda_i t} \int_0^t e^{-[\lambda_j - \lambda_i]s} ds \\ &= \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} + \sum_{i \neq j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right]. \end{aligned}$$

Assim

$$f_{T_1+T_2}(t) = \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} + \sum_{i \neq j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \quad (4.34)$$

ou

$$\begin{aligned} f_{T_1+T_2}(t) &= \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} + \sum_{i < j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \\ &+ \sum_{i > j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} + \sum_{i < j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \\
&+ \sum_{i < j} a_j \lambda_j a_i \lambda_i \left[\frac{e^{-\lambda_j t} - e^{-\lambda_i t}}{\lambda_i - \lambda_j} \right] \\
&= \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} + \sum_{i < j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \\
&+ \sum_{i < j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right].
\end{aligned}$$

Portanto

$$f_{T_1+T_2}(t) = \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} + 2 \sum_{i < j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right]. \quad (4.35)$$

Sabemos que $F'_{T_1+T_2}(t) = f_{T_1+T_2}(t)$.

Para verificar isso, vamos comparar (4.32) com (4.34). Assim de (4.32) temos

$$F_{T_1+T_2}(t) = 1 - \sum_{i=0}^n a_i e^{-\lambda_i t} - \sum_{i=0}^n a_i^2 \lambda_i t e^{-\lambda_i t} - \sum_{i \neq j} a_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right].$$

Então

$$\begin{aligned}
f_{T_1+T_2}(t) &= \sum_{i=0}^n a_i \lambda_i e^{-\lambda_i t} + \sum_{i=0}^n \left[(a_i \lambda_i)^2 t e^{-\lambda_i t} - a_i^2 \lambda_i e^{-\lambda_i t} \right] \\
&+ \sum_{i \neq j} a_i a_j \lambda_j \left[\frac{\lambda_i e^{-\lambda_i t} - \lambda_j e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \\
&= \sum_{i=0}^n a_i \lambda_i e^{-\lambda_i t} + \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} - \sum_{i=0}^n a_i^2 \lambda_i e^{-\lambda_i t} \\
&+ \sum_{i \neq j} a_i a_j \lambda_j \left[\frac{\lambda_i e^{-\lambda_i t} - \lambda_i e^{-\lambda_j t} + \lambda_i e^{-\lambda_j t} - \lambda_j e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \\
&= \sum_{i=0}^n a_i \lambda_i e^{-\lambda_i t} + \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} - \sum_{i=0}^n a_i^2 \lambda_i e^{-\lambda_i t} \\
&+ \sum_{i \neq j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] - \sum_{i \neq j} a_i a_j \lambda_j e^{-\lambda_j t} \\
&= \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} + \sum_{i \neq j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right]
\end{aligned}$$

$$\begin{aligned}
& + \sum_{i=0}^n a_i \lambda_i e^{-\lambda_i t} - \sum_{i=0}^n a_i^2 \lambda_i e^{-\lambda_i t} - \sum_{i \neq j} a_i a_j \lambda_j e^{-\lambda_j t} \\
& = \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} + \sum_{i \neq j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \\
& + \sum_{i=0}^n a_i \lambda_i e^{-\lambda_i t} - \sum_{i=j} a_i a_j \lambda_j e^{-\lambda_j t} - \sum_{i \neq j} a_i a_j \lambda_j e^{-\lambda_j t} \\
& = \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} + \sum_{i \neq j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \\
& + \sum_{i=0}^n a_i \lambda_i e^{-\lambda_i t} - \sum_{i,j} a_i a_j \lambda_j e^{-\lambda_j t} \\
& = \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} + \sum_{i \neq j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \\
& + \sum_{i=0}^n a_i \lambda_i e^{-\lambda_i t} - \left(\sum_{j=0}^n a_j \right) \sum_{i=0}^n a_i \lambda_i e^{-\lambda_i t} \\
& = \sum_{i=0}^n (a_i \lambda_i)^2 t e^{-\lambda_i t} + \sum_{i \neq j} a_i \lambda_i a_j \lambda_j \left[\frac{e^{-\lambda_i t} - e^{-\lambda_j t}}{\lambda_j - \lambda_i} \right] \\
& + \left(1 - \sum_{j=0}^n a_j \right) \left(\sum_{i=0}^n a_i \lambda_i e^{-\lambda_i t} \right).
\end{aligned}$$

Assim, de (4.27) obtemos a fórmula (4.34).

O algoritmo 11, aproxima as convoluções usando as fórmulas (4.33) e (4.35).

Algoritmo 11 Subrotina cálculo da convolução para a função de densidade e cauda de $T_1 + T_2$

```

1: Dados:  $a, \lambda, t$ .
2:  $n = \text{length}(a)$ 
3:  $soma1 = 0$ 
4: Para  $i = 0 : n$  fazer:
5:      $soma1 = soma1 + [1 + a(i) * \lambda(i) * t] * a(i) * \exp(-\lambda(i) * t)$ 
6: Fim
7:  $soma2 = 0$ 
8: Para  $i = 0 : (n - 1)$  fazer:
9:     Para  $j = (i + 1) : n$  fazer:
10:         $soma2 = soma2 + a(i) * a(j) * \left[ \frac{\lambda(j) + \lambda(i)}{\lambda(j) - \lambda(i)} \right] * [\exp(-\lambda(i) * t) - \exp(-\lambda(j) * t)]$ 
11:    Fim
12: Fim
13:  $convbarFT = soma1 + soma2$ 
14:  $soma3 = 0$ 
15: Para  $i = 0 : n$  fazer:
16:      $soma3 = soma3 + [(a(i) * \lambda(i)) * (a(i) * \lambda(i))] * t * \exp(-\lambda(i) * t)$ 
17: Fim
18:  $soma4 = 0$ 
19: Para  $i = 0 : (n - 1)$  fazer:
20:     Para  $j = (i + 1) : n$  fazer:
21:         $soma4 = soma4 + a(i) * \lambda(i) * a(j) * \lambda(j) * \left[ \frac{\exp(-\lambda(i) * t) - \exp(-\lambda(j) * t)}{\lambda(j) - \lambda(i)} \right]$ 
22:    Fim
23: Fim
24:  $convfT = soma3 + 2 * soma4$ 

```

Para a variável aleatória $X = e^T$, calculamos a convolução de Mellin a partir de (3.39) e (3.40). Segue o algoritmo 12 abaixo.

Algoritmo 12 Subrotina cálculo da convolução para a função de densidade e cauda de $X_1 X_2$

```

1: Dados:  $a, \lambda$  e  $x$ .
2: Defina a subrotina convolução sendo do algoritmo 11.
3:  $[convbarFT, convfT] = \text{convolução}(a, \lambda, \log x)$ .
4:  $convbarFX = convbarFT$ 
5:  $convfX = \frac{convfT}{x}$ .

```

Com este resultado podemos desenvolver métodos numéricos de aproximação para a convolução de Mellin através de métodos numéricos de aproximação para a convolução de Fourier via mistura de exponenciais e depois aplicar o operador que transforma a convolução usual para a convolução de Mellin.

4.3 Métodos numéricos

A escolha de **bons** parâmetros α, β, p e r , determina uma boa aproximação via mistura de exponenciais de nossa distribuição de interesse. Apresentamos algumas simulações feita na linguagem Matlab. Para o caso de distribuições contínuas, a precisão de uma aproximação $\hat{F}$ de uma função de distribuição F é o máximo da diferença entre a verdadeira função de distribuição F e a sua aproximação $\hat{F}$, chamado de erro máximo absoluto, denotado como

$$\|F - \hat{F}\| = \sup_{t \geq 0} |F(t) - \hat{F}(t)|.$$

Geralmente, mais útil do que o erro máximo absoluto é a relação entre este e a grandeza que está afetada pelo erro. Ao quociente entre o erro absoluto e o módulo do valor exato, chama-se erro relativo de F .

$$\delta = \frac{\|F - \hat{F}\|}{\|F\|}. \quad (4.36)$$

Testamos as aproximações via mistura de exponenciais para três tipos de variáveis aleatórias, com diferentes graus do polinômio e três tipos de tamanho da amostra simuladas, para o cálculo da transformada de Laplace da fórmula (4.23), expressa em esperança, usando Monte Carlo.

Para garantir que a escolha de nossos parâmetros foram boas, para cada grau do polinômio e cada tamanho da amostra simulada, aproximamos 500 vezes uma mesma distribuição e calculamos o erro relativo (4.36) de cada repetição para obter o erro relativo médio das aproximações.

4.3.1 Aproximações via mistura de exponenciais

Recorde que, para $\alpha = 0$, obtemos a fórmula (4.23). A aproximação via mistura de exponenciais (4.10), para $t \geq 0$, foi feita através da série de polinômios de Jacobi deslocado sob $(0, 1]$ a partir da transformação $x = e^{-rt}$, com $r > 0$.

Os coeficientes λ_j , para cada $j = 0, \dots, n$, da fórmula (4.12) devem ser positivos. Portanto, como cada um depende de j, r e p , temos que $p > 0$.

Como consequência de (4.17), o teorema 4.1.2 vale para $p < (\beta + 1)/2$. Assim obtemos a restrição $0 < p < (\beta + 1)/2$.

O polinômio de Jacobi deslocado é reduzido para o polinômio de Legendre quando $\beta = 0$, desde que $\alpha = 0$. Os testes começaram com esse valor de β , caso o erro relativo obtido não for pequeno, o valor de β foi aumentado até encontrar um erro relativo pequeno. A partir daí, calculamos o erro relativo médio para garantir que tivemos uma boa aproximação.

Uniforme

Usamos mistura de exponenciais para aproximar a distribuição da v.a. $U \sim U(1, 2)$. Aqui, não

foi preciso computar por Monte Carlo a transformada de Laplace, gerando amostras da uniforme. Isto ocorreu porque podemos calcular analiticamente a sua respectiva transformada de Laplace. Assim, o erro relativo foi calculado para cada grau do polinômio de aproximação.

Tabela 4.1: Distribuição Uniforme

Grau do polinômio	Erro relativo
3	0.24143
5	0.16023
10	0.065960
15	0.053186
18	0.040506
19	0.037933
20	0.061932

Os parâmetros escolhidos para obter os valores da tabela foram $\alpha = 0$, $\beta = 0$, $r = 0.7$ e $p = 0.1$. A melhor aproximação foi obtida com o polinômio de grau 19 que mostra um erro relativo de 3,8%. A seguir apresentamos os gráficos:

(a) $n=3$ (b) $n=5$ (c) $n=19$

Figura 4.1: Cauda da Uniforme

(a) $n=3$ (b) $n=5$ (c) $n=19$

Figura 4.2: Distribuição da Uniforme

(a) $n=3$ (b) $n=5$ (c) $n=19$

Figura 4.3: Densidade da Uniforme

Pareto

Usamos mistura de exponenciais para aproximar a distribuição da v.a. Pareto(1,1). Aqui, usamos 3 tipos de graus do polinômio e 3 tipos de tamanhos da amostra, para o cálculo da transformada de Laplace da v.a. usando Monte Carlo. Assim, o erro relativo medio foi calculado para cada grau do polinômio de aproximação e cada tamanho da amostra.

Tabela 4.2: Distribuição de Pareto - 500 repetições

Grau do polinômio	Amostra em MC	Erro relativo médio	Desvio Padrão
2	2^{19}	0.2892	0.0029
2	2^{20}	0.2891	0.0020
2	2^{21}	0.2892	0.0014
3	2^{19}	0.3161	0.0109
3	2^{20}	0.3152	0.0083
3	2^{21}	0.3148	0.0052
5	2^{19}	0.6514	0.3568
5	2^{20}	0.5332	0.2318
5	2^{21}	0.4415	0.1389

Os parâmetros escolhidos para obter os valores da tabela foram $\alpha = 0$, $\beta = 0$, $r = 0.1$ e $p = 0.1$. A cauda da distribuição Pareto é mais pesada que a mistura de exponenciais. Portanto, vemos que não possuímos uma aproximação de boa qualidade. A seguir apresentamos os gráficos:

(a) $n=3$

(b) $n=5$

Figura 4.4: Cauda da Pareto

Figura 4.5: Distribuição da Pareto

Figura 4.6: Densidade da Pareto

Log-Normal

Usamos mistura de exponenciais para aproximar a distribuição da v.a. $\text{Log-Normal}(0, 1)$. Aqui, usamos 2 tipos de graus do polinômio e 3 tipos de tamanhos da amostra para o cálculo da transformada de Laplace da v.a. usando Monte Carlo. Assim, o erro relativo medio foi calculado para cada grau do polinômio de aproximação e cada tamanho da amostra.

Tabela 4.3: Distribuição Lognormal - 500 repetições

Grau do polinômio	Amostra em MC	Erro relativo médio	Desvio Padrão
3	2^{19}	0.0315	0.0208
3	2^{20}	0.0242	0.0155
3	2^{21}	0.0194	0.0114
5	2^{19}	0.5892	0.4436
5	2^{20}	0.4305	0.2993
5	2^{21}	0.2834	0.2109

Os parâmetros escolhidos para obter os valores da tabela foram $\alpha = 0$, $\beta = 0.5$, $r = 1$ e $p = 0.4$. A melhor aproximação foi obtida com o polinômio de grau 3 que mostra um erro relativo médio de 1,94%, fornecido pela baixa dispersão como podemos ver no seu respectivo desvio padrão igual a 1,14%. A seguir apresentamos os gráficos:

(a) n=3

(b) n=5

Figura 4.7: Cauda da Lognormal

(a) n=3

(b) n=5

Figura 4.8: Distribuição da Lognormal

Figura 4.9: Densidade da Lognormal

4.3.2 Aproximação da convolução

Geramos uma amostra normal truncada $[0, \infty)$. Isto foi feito porque a aproximação via mistura de exponenciais permite aproximações à distribuições com variáveis aleatórias não negativas.

Via mistura de exponenciais usamos a aproximação algébrica da convolução, usando as fórmulas (4.33) e (4.35) para duas funções, depois usamos o operador descrito por Collins [7] que transforma a convolução de Fourier para a convolução de Mellin e assim obter a aproximação numérica.

Usando este operador a transformação passa de uma aproximação da distribuição de soma de normais truncadas para uma distribuição do produto de log-normais truncadas. Isso é ilustrado na figura 4.10 a seguir.

Figura 4.10: Normal Truncada

Na figura 4.11 comparamos a aproximação a convolução de Mellin dada pelo método algébrico com a aproximação de discretização e condensação de Kaplan [11] usada por Andrade e Stern [3]. A aproximação é comparada usando a convolução de Mellin de log-normais truncadas geradas a partir de normais truncadas a valores não negativos.

Figura 4.11: Normal Truncada - Comparações das convoluções

4.4 Método direto para o cálculo das convoluções

Usando a estrutura algébrica de Hogg [10] para o cálculo das convoluções, percebemos que existe uma aproximação tão boa quanto o método de discretização e condensação, ver figura 4.11. Nosso cálculo da convolução de Mellin de log-normais truncadas a $[1, \infty)$ foi feita para o caso 2.

Se fizermos o cálculo da convolução para três ou mais distribuições, a fórmula algébrica fica cada vez mais complicada de representar por mistura de convoluções exponenciais. Isto ocorre porque acumulamos exponencialmente o número de operações ao aumentar o número de variáveis aleatórias para calcular a convolução de Mellin.

Além de aproximar as distribuições acumuladas, precisamos aproximar as densidades delas e calcular as convoluções algebricamente colocando mais restrições como visto na equação (4.27).

Assim, a teoria sugerida por Hogg [10] esclarece as propriedades algébricas de tipo de representação por mistura de exponenciais. Todavia, nesta forma as metas numéricas desenvolvidas têm alto custo computacional.

Sabemos que a convolução das distribuições de variáveis aleatórias independentes é também uma distribuição. Portanto, a ideia sugerida por Dufresne [8] é aproximar a distribuição diretamente via mistura de exponenciais como mencionaremos a seguir.

Recorde que, para $\alpha = 0$, obtemos a equação (4.23), ou seja

$$b_k = (2k + \beta + 1) \sum_{j=0}^k \frac{\rho_{kj}}{(j - p + \beta + 1)} \left[1 - E \left(e^{-(j-p+\beta+1)rT} \right) \right], \quad k = 0, \dots, n.$$

A aproximação para a distribuição de uma variável aleatória T depende da sua Transformada de Laplace, expressa em esperança matemática. Portanto, aproximar uma distribuição da variável aleatória $T_1 + \dots + T_N$, onde $T_1, \dots, T_N$ são v.a's independentes não negativas, também dependeria da sua Transformada de Laplace. Assim,

$$E\left(e^{-\omega(T_1+\dots+T_N)}\right) = \prod_{i=1}^N E\left(e^{-\omega T_i}\right). \quad (4.37)$$

Ou seja, geramos amostras para o cálculo da transformada de Laplace de cada v.a. T_i , $i = 1, \dots, N$, usando Monte Carlo. Depois, por serem elas independentes, calculamos a transformada de Laplace da v.a. $T_1 + \dots + T_N$ mediante o produto de cada transformada de Laplace da v.a. T_i , $i = 1, \dots, N$ calculada.

Para a variável aleatória $X = e^T$, usamos o operador $\mathcal{T}$ descrito por Collins [7] obtendo a convolução de Mellin

$$\bar{F}_{X_1 X_2 \dots X_N}(x) = [\mathcal{T}^{-1}(\bar{F}_{T_1+T_2+\dots+T_N})](x) = \bar{F}_{T_1+T_2+\dots+T_N}(\log(x))$$

onde

$$\bar{F}_T(t) = 1 - F_T(t), \quad t \geq 0.$$

Portanto, para fazer o cálculo direto via mistura de exponenciais, basta fazer uma mudança no algoritmo 8 usando a equação (4.37) como veremos a seguir.

Algoritmo 13 Subrotina cálculo de b e λ

```

1: Dados:  $n, \beta, p, r, N, chooseVarT$ .
2: Para  $k = 0 : n$  fazer:
3:      $x = 2 * k + \beta + 1$ 
4:      $y = 0$ 
5:     Para  $j = 0 : k$  fazer:
6:          $\theta = (\beta - p + j + 1) * r$ 
7:          $Espe = 1$ 
8:         Para  $i = 1 : N$  fazer:
9:             faça uma subrotina para escolher a  $i$ -ésima v.a.  $T$  usando como criterio
10:             $chooseVarT$ 
11:            faça uma subrotina para obter  $esp(\theta) = \mathbb{E}[e^{-\theta T}]$ 
12:             $Espe = Espe * esp(\theta)$ 
13:        Fim
14:         $y = y + \rho(k, j) * [1 - Espe] * \frac{r}{\theta}$ 
15:    Fim
16:     $b(k) = x * y$ 
17:     $\lambda(k) = (k + p) * r$ 
18: Fim.
```

Aproximamos via mistura de exponenciais a convolução seguindo os passos dos algoritmos mencionados acima trocando o algoritmo 8 pelo algoritmo 13.

Para o caso da convolução de Mellin de duas distribuições, fizemos a comparação entre a fórmula algébrica, usando a aproximação via mistura de convoluções exponenciais, e o método direto, usando o produto das Transformadas de Laplace, para a log-normal truncada em $[1, \infty)$, mencionada na figura 4.11.

Mostramos os gráficos.

Figura 4.12: Convolução de Mellin Alg Mix of Exp vs Direct Mix of Exp

Na figura 4.12, vemos que o cálculo direto da convolução via mistura de exponenciais usando o método direto, possui uma aproximação tão boa quanto o cálculo algébrico da convolução via mistura de convoluções exponenciais.

Nossa vantagem de usar o método direto para calcular a convolução de Mellin, é que ele pode ser generalizado para N variáveis aleatórias aumentando linearmente o número de operações. Portanto, é muito melhor comparado com o cálculo algébrico da convolução via mistura de convoluções exponenciais.

Capítulo 5

O Teste de independência condicionada

Estamos usando a aplicação do artigo de Andrade e Stern [3], onde é testada a independência condicionada das v.as $Y \in \{1, \dots, R\}, Z \in \{1, \dots, C\}$ dado $X \in \{1, \dots, N\}$. Aqui vamos aproximar numericamente as funções verdade definida no capítulo 1 para calcular a evidência definida como a convolução de Mellin das funções verdade vista na equação (1.4).

Usaremos a segunda proposta para o cálculo das convoluções comentada no capítulo 4. Assim, aproximaremos numericamente a convolução de Mellin das funções verdade e consequentemente encontraremos o valor da evidência do teste de independência condicionada.

5.1 O Modelo Multinomial

Queremos testar a independência condicionada das v.as $Y \in \{1, \dots, R\}, Z \in \{1, \dots, C\}$ dado $X \in \{1, \dots, N\}$.

O teste de independência condicionada $Y \perp Z | X$ pode ser escrito como a hipótese nula conjunta

$$H : [Y \perp Z | X = 1] \wedge [Y \perp Z | X = 2] \wedge \dots \wedge [Y \perp Z | X = N].$$

A hipótese H pode ser decomposta em hipóteses simples sendo

$$\begin{aligned} H_1 & : Y \perp Z | X = 1 \\ H_2 & : Y \perp Z | X = 2 \\ & \vdots \\ H_N & : Y \perp Z | X = N. \end{aligned}$$

Note que as hipóteses $H_1, \dots, H_N$ são condicionalmente independentes.

Para cada $k = 1, \dots, N$, com $X = k$, os valores tomados por Y e Z são considerados observações aleatórias estabelecido por alguma distribuição

$$\mathbb{P}(\{Y = i\}, \{Z = j\} | \{X = k\}), \quad i = 1, \dots, R, \quad j = 1, \dots, C.$$

Assim, definimos

$$\theta_{ijk} = \mathbb{P}(\{Y = i\}, \{Z = j\} | \{X = k\}), \quad i = 1, \dots, R, \quad j = 1, \dots, C, \quad k = 1, \dots, N.$$

Para $X = k$, $k = 1, \dots, N$, obtemos nosso parâmetro da seguinte maneira

Tabela 5.1: $\mathbf{X=k}$

v.as	Z=1	Z=2	$\dots$	Z=C	
Y=1	θ_{11k}	θ_{12k}	$\dots$	θ_{1Ck}	$\theta_{1\bullet k}$
Y=2	θ_{21k}	θ_{22k}	$\dots$	θ_{2Ck}	$\theta_{2\bullet k}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
Y=R	θ_{R1k}	θ_{R2k}	$\dots$	θ_{RCk}	$\theta_{R\bullet k}$
	$\theta_{\bullet 1k}$	$\theta_{\bullet 2k}$	$\dots$	$\theta_{\bullet Ck}$	$\theta_{\bullet\bullet k}$

onde

$$\theta_{\bullet jk} = \sum_{i=1}^R \theta_{ijk}, \quad \theta_{i\bullet k} = \sum_{j=1}^C \theta_{ijk} \quad \text{e} \quad \theta_{\bullet\bullet k} = \sum_{i,j}^{R,C} \theta_{ijk} = 1.$$

Vemos que para cada $i = 1, \dots, R$ e $j = 1, \dots, C$

$$\mathbb{P}(\{Y = i\} | \{X = k\}) = \sum_{j=1}^C \mathbb{P}(\{Y = i\}, \{Z = j\} | \{X = k\}) = \sum_{j=1}^C \theta_{ijk} = \theta_{i\bullet k}$$

e

$$\mathbb{P}(\{Z = j\} | \{X = k\}) = \sum_{i=1}^R \mathbb{P}(\{Y = i\}, \{Z = j\} | \{X = k\}) = \sum_{i=1}^R \theta_{ijk} = \theta_{\bullet jk}.$$

Dentro da nossa hipótese nula $H_k : Y \perp Z | X = k$, $k = 1, \dots, N$, temos que para cada $i = 1, \dots, R$ e $j = 1, \dots, C$

$$\mathbb{P}(\{Y = i\}, \{Z = j\} | \{X = k\}) = \mathbb{P}(\{Y = i\} | \{X = k\}) \mathbb{P}(\{Z = j\} | \{X = k\}),$$

ou seja

$$\theta_{ijk} = \theta_{i\bullet k} \theta_{\bullet jk},$$

o que é equivalente a

$$\begin{aligned} H_1 &: \theta_{ij1} = \theta_{i\bullet 1} \cdot \theta_{\bullet j1} \\ H_2 &: \theta_{ij2} = \theta_{i\bullet 2} \cdot \theta_{\bullet j2} \\ &\vdots \\ H_N &: \theta_{ijN} = \theta_{i\bullet N} \cdot \theta_{\bullet jN}, \quad \forall i, j. \end{aligned}$$

Sejam as observações

Tabela 5.2: $\mathbf{X}=\mathbf{k}$

v.as	Z=1	Z=2	$\cdots$	Z=C	
Y=1	η_{11k}	η_{12k}	$\cdots$	η_{1Ck}	$\eta_{1\bullet k}$
Y=2	η_{21k}	η_{22k}	$\cdots$	η_{2Ck}	$\eta_{2\bullet k}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
Y=R	η_{R1k}	η_{R2k}	$\cdots$	η_{RCk}	$\eta_{R\bullet k}$
	$\eta_{\bullet 1k}$	$\eta_{\bullet 2k}$	$\cdots$	$\eta_{\bullet Ck}$	$\eta_{\bullet\bullet k}$

onde

$$\eta_{\bullet jk} = \sum_{i=1}^R \eta_{ijk}, \quad \eta_{i\bullet k} = \sum_{j=1}^C \eta_{ijk} \quad \text{e} \quad \eta_{\bullet\bullet k} = \sum_{i,j}^{R,C} \eta_{ijk}.$$

Cada observação, η_{ijk} , acontece com probabilidade θ_{ijk} , $\forall i, j, k$. Assim, podemos estabelecer a nossa hipótese H_k o modelo multinomial definido como

$$P(\eta_k | \theta_k) = \eta_{\bullet\bullet k}! \prod_{i,j}^{R,C} \frac{\theta_{ijk}^{\eta_{ijk}}}{\eta_{ijk}!},$$

onde

$$\eta_k = \{\eta_{ijk}/i = 1, \dots, R; \quad j = 1, \dots, C\}, \quad k = 1, \dots, N.$$

é o vetor de observações, e

$$\theta_k = \{\theta_{ijk}/i = 1, \dots, R; \ j = 1, \dots, C\}, \quad k = 1, \dots, N.$$

é o vetor de parâmetros.

A distribuição a priori de θ_k conjugada do modelo multinomial é a distribuição Dirichlet definida como:

$$p_0(\theta_k|\alpha_k) = \Gamma\left(\sum_{i,j}^{R,C} \alpha_{ijk}\right) \prod_{i,j}^{R,C} \frac{\theta_{ijk}^{\alpha_{ijk}-1}}{\Gamma(\alpha_{ijk})},$$

cujo vetor de hiperparâmetros é dado por:

$$\alpha_k = \{\alpha_{ijk}/i = 1, \dots, R; \ j = 1, \dots, C\}, \quad k = 1, \dots, N.$$

Daqui, obtemos a distribuição a posteriori de θ_k , $k = 1, \dots, N$, como

$$p_n(\theta_k) \propto \prod_{i,j}^{R,C} \theta_{ijk}^{\alpha_{ijk} + \eta_{ijk} - 1}. \quad (5.1)$$

5.2 O exemplo

No caso do exemplo apresentado em Andrade e Stern [3], os dados do problema são:

Para $N = R = C = 3$, $\alpha_{ijk} = 1$, $\forall i, j, k$, e

Tabela 5.3: **X=1**

v.as	Z=1	Z=2	Z=3	Soma Y
Y=1	241	187	44	472
Y=2	139	130	30	299
Y=3	364	302	70	736
Soma Z	744	619	144	1507

Tabela 5.4: **X=2**

v.as	Z=1	Z=2	Z=3	Soma Y
Y=1	42	41	323	406
Y=2	39	41	341	421
Y=3	15	21	171	207
Soma Z	96	103	835	1534

Tabela 5.5: **X=3**

v.as	Z=1	Z=2	Z=3	Soma Y
Y=1	282	35	151	468
Y=2	131	37	79	247
Y=3	1055	143	546	1744
Soma Z	1468	215	776	2459

obtemos θ_k^* , $k = 1, 2, 3$, isto é, o parâmetro MAP sob a hipótese em teste

$$\begin{aligned}
\theta_1^* &= [0.15462810, 0.12864892, 0.02992802, 0.09795297, 0.081495820, \\
&\quad 0.01895864, 0.24111501, 0.20060509, 0.04666742] \\
\theta_2^* &= [0.03645492, 0.03911309, 0.31708189, 0.03780178, 0.040558160, \\
&\quad 0.32879673, 0.01858662, 0.01994190, 0.16166490] \\
\theta_3^* &= [0.11362002, 0.01664053, 0.06006072, 0.05996612, 0.008782504, \\
&\quad 0.03169871, 0.42340451, 0.06201088, 0.22381601].
\end{aligned}$$

O máximo da função surpresa a posteriori sob a hipótese H , $S(\theta_k)$, definida em (1.2) é obtida avaliando em θ_k^* , usando como valor de densidade de referência igual a 1, $r(\theta_k) = 1$. Assim

$$s_1^* = 1.082082 \cdot 10^{14}, \quad s_2^* = 9.275326 \cdot 10^{13}, \quad s_3^* = 1.049843 \cdot 10^{13}.$$

Com isto, obtemos a evidência da hipótese H_k definido em (1.3) tendo como resultado

$$ev(H_1) = 0.9890, \quad ev(H_2) = 0.9920, \quad ev(H_3) = 0.1100$$

Estes são os gráficos que representam as funções verdade, ou seja (W_1, W_2, W_3) ,

Figura 5.1: Distribuição da função verdade

Como podemos observar, as aproximações se comportam melhor na região das caudas.

Pelo comentado anteriormente, veja figura 5.1, a discretização e condensação tem comportamento melhor na região central, embora, podemos tirar vantagem com o método de mistura de exponenciais na região das caudas. Isto nos motivou a usar um método híbrido onde possamos ter um melhor comportamento na representação da função verdade.

Estes são os gráficos que representam as funções verdade hibridizadas, ou seja (W_1, W_2, W_3) .

Caso Horizontal hibridizada

(a) $n=3$ (b) $n=3$ (c) $n=3$

Figura 5.2: Distribuição da função verdade Híbrida Horizontal

Caso Vertical hibridizada

Figura 5.3: Distribuição da função verdade Híbrida Vertical

Os valores das evidências usando o método híbrido horizontal são

$$ev(H_1) = 0.9930, \quad ev(H_2) = 0.9840, \quad ev(H_3) = 0.1410$$

Os valores das evidências usando o método híbrido vertical são

$$ev(H_1) = 0.9920, \quad ev(H_2) = 0.9800, \quad ev(H_3) = 0.1040$$

Consequentemente, para cada função verdade aproximada via mistura de exponenciais, a obtenção dos coeficientes foi dado calculando a transformada de Laplace de cada variável aleatória que representa cada função verdade. Portanto, aproximamos via mistura de exponenciais a convolução das funções verdade calculando os coeficientes pela obtenção do produto das transformadas de Laplace de cada v.a. Assim hibridizamos o método e obtemos a convolução desejada.

Figura 5.4: Convolução de Mellin das Distribuições da função verdade

O valor da evidência usando a convolução de Mellin pelo método híbrido horizontal é dada por

$$ev(H) = 0.8667$$

O valor da evidência usando a convolução de Mellin pelo método híbrido vertical é dada por

$$ev(H) = 0.8960$$

o que indica uma alta probabilidade de não rejeitar nossa hipótese composta H e portanto para nossas observações dadas, a independência condicionada, $Y \perp Z | X$, é verdadeira.

Conclusões

Encontramos métodos eficientes de calcular $W = \bigotimes_{\{1 \leq j \leq N\}} W_j$ definida no capítulo 1. Estes métodos foram desenvolvidos utilizando:

1. Métodos de condensação, cujo comportamento é melhor na região central.
2. Aproximação à convolução de Mellin via mistura de exponenciais, desenvolvida usando a segunda proposta descrita no capítulo 4.

Os dois métodos, via discretização e condensação descrito por Kaplan [11] e via mistura de exponenciais descrito por Dufresne [8], precisam utilizar simulações para aproximar a convolução de Mellin.

O método de discretização e condensação utiliza as simulações, que representa alguma variável aleatória contínua, para transformar a variável aleatória contínua para uma variável aleatória discreta usando aritmética probabilística. A distribuição da variável aleatória contínua é aproximada pela distribuição da variável aleatória discreta calculada por dito método.

Os valores da distribuição aproximada via discretização e condensação está limitada pelo seu mínimo e máximo valor das simulações, portanto, não podemos calcular valores da distribuição fora desse intervalo.

O método de mistura de exponenciais utiliza as simulações para calcular os coeficientes de mistura através do compute da transformada de Laplace via Monte Carlo.

Os valores da distribuição aproximada via mistura de exponenciais pode ser calculada para um intervalo maior quanto quiser.

Como nem sempre a mistura de exponenciais aproxima melhor na região central, procuramos melhorar o método fazendo um método híbrido, onde fazemos discretização e condensação na região central e mistura de exponenciais na região das caudas. Se a mistura de exponenciais aproximar eficientemente na região central, podemos usar o método de mistura de exponenciais de forma única, mas isso dependerá do tipo de problema.

Na aplicação feita no capítulo 5 vimos que a melhor aproximação numérica foi obtida pelo método híbrido.

Usando a segunda proposta descrita no capítulo 4, foi possível generalizar a aproximação à convolução de Mellin via mistura de exponenciais para N funções.

Referências Bibliográficas

- [1] J. M. Almira, *Matemáticas para la recuperación de señales.*, Grupo Editorial Universitario (Granada), 2005. 21, 30, 31
- [2] D. Andrade, *Transformada de Laplace.* 21
- [3] P. M. Andrade, J. M. Stern, and C. A. Bragança, *Bayesian test of significance for conditional independence: The multinomial model*, Submitted (2013). 1, 9, 11, 19, 77, 83, 86
- [4] J. C. Barata, *Notas de aula: Curso de Física-Matemática*, http://denebola.if.usp.br/~jbarata/Notas_de_aula/, 2005. 37, 47
- [5] P. Billingsley, *Probability and measure*, Wiley India Pvt. Ltd., 2008. 5
- [6] W. Borges and J. M. Stern, *The rules of logic composition for the bayesian epistemic e-values*, *Logic Journal of the IGPL* v. 15 (2007), p. 401–420. iii, v, 1, 5, 7
- [7] D.H. Collins, *The relationship between fourier and mellin transforms, with applications to probability and stochastic processes*, (2011). iii, v, 2, 20, 21, 36, 39, 48, 66, 77, 79
- [8] D. Dufresne, *Fitting combinations of exponentials to probability distributions*, *Applied Stochastic Models in Business and Industry* v. 23 (2007), p. 23–48. iii, v, 2, 49, 50, 55, 58, 78, 93
- [9] G. González, *Series de fourier, transformadas de fourier y aplicaciones*, *Divulgaciones Matemáticas* v. 5 (1997), p. 43–60. 21, 31
- [10] D. Hogg, *An algebraic structure for the convolution of life distributions.*, Ph.D. thesis, Monterey, California. Naval Postgraduate School, 1982. iii, v, 1, 2, 41, 44, 49, 50, 62, 78
- [11] S. Kaplan and J. Lin, *An improved condensation procedure in discrete probability distribution calculations*, *Risk Analysis* v. 7 (1987), no. 1, p. 15–19. iii, v, 1, 16, 77, 93
- [12] J. B. Rodney, *Notas de aula: Introdução às equações diferenciais parciais*, www.mat.ufmg.br/~rodney, 2007. 21, 31
- [13] H.L. Royden, *Real analysis*, Macmillan New York, 1968. 6, 21, 23
- [14] R. A. Saenz, *Notas de aula: Análisis Real: Primer curso*, http://fejer.ucol.mx/cursos/2004-20/analisis_I/texto/texto.html, 2004. 21, 23
- [15] AN Shiryaev, *Probability. Number 95 in Graduate Texts in Mathematics*, 1996. 5
- [16] R. C. Williamson, *Probabilistic arithmetic*, Ph.D. thesis, Department of Electrical Engineering, University of Queensland, 1989. 11

- [17] R. C. Williamson and T. Downs, *Probabilistic arithmetic. I. Numerical Methods for Calculating Convolutions and Dependency Bounds*, vol. 4, 1990. [11](#)