
PROPAGAÇÃO DE INCERTEZAS EM ELETROMAGNETISMO

EDSON ALVES DA COSTA JÚNIOR

TESE DE DOUTORADO EM ENGENHARIA ELÉTRICA

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

PROPAGAÇÃO DE INCERTEZAS EM
ELETROMAGNETISMO

EDSON ALVES DA COSTA JÚNIOR

ORIENTADOR: LEONARDO RODRIGUES ARAÚJO XAVIER DE
MENEZES

TESE DE DOUTORADO EM
TELECOMUNICAÇÕES

PUBLICAÇÃO: PPGENE.TD - 040/09

BRASÍLIA/DF: MAIO - 2009.

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

PROPAGAÇÃO DE INCERTEZAS EM
ELETROMAGNETISMO

EDSON ALVES DA COSTA JÚNIOR

TESE DE DOUTORADO SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA ELÉTRICA DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM TELECOMUNICAÇÕES.

APROVADA POR:

Prof. Leonardo Rodrigues Araújo Xavier de Menezes, PhD. (ENE-UnB)
(Orientador)

Prof. Geovany Araújo Borges, PhD. (ENE-UnB)
(Examinador Interno)

Prof. Leonardo Guedes, PhD. (UFG)
(Examinador Externo)

BRASÍLIA/DF, 12 DE MAIO DE 2009.

FICHA CATALOGRÁFICA

COSTA JR, EDSON ALVES DA

Propagação de Incertezas em Eletromagnetismo. [Distrito Federal] 2009.
xvii, 103p., 297 mm (ENE/FT/UnB, Doutor, Telecomunicações, 2009).

Tese de Doutorado - Universidade de Brasília.

Faculdade de Tecnologia.

Departamento de Engenharia Elétrica.

- | | |
|---------------------------|-----------------------|
| 1. Eletromagnetismo | 2. Simulação numérica |
| 3. Processos estocásticos | 4. TLM |
| I. ENE/FT/UnB | II. Título (série) |

REFERÊNCIA BIBLIOGRÁFICA

COSTA JR., E. A. (2009). Propagação de incertezas em Eletromagnetismo. Tese de Doutorado em Telecomunicações, Publicação PPGENE.TD - 040/09, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 103p.

CESSÃO DE DIREITOS

NOME DO AUTOR: Edson Alves da Costa Júnior.

TÍTULO DA TESE DE DOUTORADO: Propagação de Incertezas em Eletromagnetismo.

GRAU / ANO: Doutor / 2009

É concedida à Universidade de Brasília permissão para reproduzir cópias desta tese de doutorado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta tese de doutorado pode ser reproduzida sem a autorização por escrito do autor.

Edson Alves da Costa Júnior
Departamento de Engenharia Elétrica - Universidade de Brasília
Brasília - DF - Brasil.

DEDICATÓRIA

Este trabalho é dedicado a Jesus Cristo, nosso Senhor, e a minha amada esposa Regiane Quezia, pelo inestimável carinho e apoio.

AGRADECIMENTOS

Agradeço a Deus por me cobrir de bênçãos, mesmo eu não merecendo.

Agradeço ao meu pai, Edson Alves da Costa, de quem herdei o nome, a teimosia e o gosto pelos números.

Agradeço a minha mãe, Maria Talva da Costa, cuja determinação e coragem fizeram com que ela cruzasse o oceano para honrar sua palavra e seus compromissos, deixando este exemplo e lição para nós.

Agradeço aos meus irmãos Patrícia, Leonardo, Eduardo e Camila, com os quais aprendi a compartilhar, rir, ler e conquistar tudo em conjunto.

Agradeço aos meus afilhados, Matheus Lênin e Júlia Vitória, dos quais tenho muito orgulho, e a toda minha família.

Agradeço ao professor Leonardo Rodrigues Araújo Xavier de Menezes, o orientador deste trabalho e bom amigo, que teve a coragem de me aceitar como seu aluno a despeito da minha formação prévia de matemático.

Agradeço ao professor Hemar Godinho, meu orientador de mestrado, que me incentivou a procurar o professor Leonardo, e ao professor Geovany Araújo Borges, pelo apoio nas dúvidas que surgiram ao longo do caminho e pelo incentivo em continuar em frente.

Agradeço, por fim, a Sebastião Gomes, padrinho e amigo, e a Geralda Rosário Gomes, que me acolheu como filho e que me concedeu a honra de desposar sua filha, minha amada esposa Regiane Quezia.

RESUMO

PROPAGAÇÃO DE INCERTEZAS EM ELETROMAGNETISMO

Autor: Edson Alves da Costa Júnior

Orientador: Leonardo Rodrigues Xavier Araújo de Menezes

Brasília, Maio de 2009

Este trabalho tem como objetivo o estudo da propagação de incertezas em eletromagnetismo. Entretanto, os resultados e conceitos apresentados aqui são aplicáveis em outros problemas de ordem prática que envolvam erros de medição, variáveis aleatórias e incertezas diversas. Após a introdução, são apresentados os conceitos elementares de probabilidade que fornecem o embasamento teórico do trabalho. Em seguida, o problema de se calcular a esperança de uma variável aleatória é abordado em duas frentes: através do uso de aproximações polinomiais de funções e através do paralelo com o problema de quadraturas. Por fim, são estudados os métodos para a determinação da função de densidade de probabilidade e proposto um método numérico para tal.

ABSTRACT

UNCERTAINTY'S PROPAGATION IN ELECTROMAGNETISM

Author: Edson Alves da Costa Júnior

Supervisor: Leonardo Rodrigues Araújo Xavier de Menezes

Brasília, May of 2009

The goal of this work is study the uncertainty's propagation problem in electromagnetism, but the developed results and concepts can be applied in general practical problems involving measure errors, random variables and several others uncertainties. Following the introduction we will begin with the basic concepts of probability theory, the foundation of this work. After it we will take the estimation of expected value by two routes: using Taylor's series and using Gaussian quadrature. Finally, we will study methods to find the probability density function of a random variable and propose a numeric approach to this problem.

SUMÁRIO

LISTA DE TABELAS	x
LISTA DE FIGURAS	xi
LISTA DE SÍMBOLOS E ABREVIACÕES	xii
1 INTRODUÇÃO	1
2 PROBABILIDADE: CONCEITOS ELEMENTARES	4
2.1 VARIÁVEIS ALEATÓRIAS	4
2.1.1 Espaço de probabilidade	4
2.1.2 Definição de variáveis aleatórias	6
2.2 CARACTERIZAÇÃO DAS VARIÁVEIS ALEATÓRIAS	8
2.2.1 Funções de probabilidade	8
2.2.2 Esperança Matemática	12
2.2.3 Momentos	17
2.3 VETORES E FUNÇÕES DE VARIÁVEIS ALEATÓRIAS	24
2.3.1 Vetores de variáveis aleatórias	24
2.3.2 Funções de variáveis aleatórias	27
3 APROXIMAÇÃO POLINOMIAL DAS FUNÇÕES DE VARIÁVEIS ALEATÓRIAS	31
3.1 SÉRIE DE TAYLOR	31
3.1.1 Definição	31
3.1.2 Esperança de uma série de Taylor	34
3.2 APLICAÇÃO: TLM	35
3.2.1 TLM unidimensional com um único nó	35
3.2.2 TLM unidimensional com múltiplos nós	39
3.2.3 Resultados Experimentais	44
4 ESTIMANDO A ESPERANÇA E A VARIÂNCIA	46
4.1 UNSCENTED TRANSFORM GUASSIANA	46

4.1.1	Introdução	46
4.1.2	Quadraturas	47
4.1.3	A quadratura gaussiana generalizada	51
4.1.4	Unscented Transform Gaussiana	56
4.2	APLICAÇÃO: LEI DE COULOMB	59
4.2.1	Esperança da Lei de Coulomb	60
4.2.2	Lei de Coulomb com Monte Carlo	62
4.2.3	Lei de Coulomb com Unscented Transform	63
4.2.4	Lei de Coulomb com Unscented Transform Gaussiana	64
4.3	APLICAÇÃO: LEI DE OHM	65
4.3.1	Esperança da Lei de Ohm	66
4.3.2	Lei de Ohm com Monte Carlo	67
4.3.3	Lei de Ohm com Unscented Transform	68
4.3.4	Lei de Ohm com Unscented Transform Gaussiana	69
5	ENCONTRANDO A FUNÇÃO DE DENSIDADE	71
5.1	MÉTODOS	71
5.1.1	Método Analítico	71
5.1.2	Método do Jacobiano	74
5.1.3	Método do Jacobiano Numérico	81
5.2	APLICAÇÃO: ENERGIA ARMAZENADA POR UM INDUTOR	86
5.2.1	Método do Jacobiano	87
5.2.2	Método do Jacobiano Numérico	88
5.3	APLICAÇÃO: CAPACITOR DE PLACAS PARALELAS	90
5.3.1	Método do Jacobiano	91
5.3.2	Método do Jacobiano Numérico	93
6	CONCLUSÃO	96
	REFERÊNCIAS BIBLIOGRÁFICAS	98
	APÊNDICES	100
A	TABELAS DE PONTOS-SIGMA E PESOS	101
A.1	Distribuição uniforme	101
A.2	Distribuição normal	102
A.3	Distribuição exponencial	103

LISTA DE TABELAS

1.1	Resultados do experimento da medição da corrente i	2
3.1	Tensões incidentes e refletidas sobre o nó 1.	44
3.2	Tensões incidentese refletidas sobre o nó 2.	45
4.1	Resultados para a força média segundo a Lei de Coulomb.	66
4.2	Resultados para a corrente média segundo a Lei de Ohm.	70
A.1	Pontos-sigma e pesos para $X \sim U(0, 1)$	101
A.2	Pontos-sigma e pesos para $X \sim N(0, 1)$	102
A.3	Pontos-sigma e pesos para $X \sim \text{Exp}(1)$	103

LISTA DE FIGURAS

3.1	TLM com um único nó	36
3.2	TLM com múltiplos nós	39
4.1	Força de repulsão	59
4.2	Força de atração	60
5.1	Região de integração de $f_{Z,W}(z, w)$	79
5.2	Função de densidade de $Z = X + Y$	81
5.3	Comparativo entre a f_{U_B} obtida analiticamente e através do Método do Jacobiano Numérico com 15 pontos	89
5.4	Comparativo entre a f_{U_B} obtida analiticamente e através do Método do Jacobiano Numérico com 25 pontos	89
5.5	Comparativo entre a f_{U_B} obtida analiticamente e através do Método do Jacobiano Numérico com 50 pontos	90
5.6	Capacitor de placas paralelas	91
5.7	Região de integração R	92
5.8	Comparativo entre a f_C obtida analiticamente e através do Método do Jacobiano Numérico com 15 pontos	94
5.9	Comparativo entre a f_C obtida analiticamente e através do Método do Jacobiano Numérico com 25 pontos	94
5.10	Comparativo entre a f_C obtida analiticamente e através do Método do Jacobiano Numérico com 50 pontos	95

LISTA DE SÍMBOLOS E ABREVIACÕES

Ω : espaço amostral de um experimento.

ω : resultado de um experimento.

A : evento aleatório.

A^c : conjunto complementar de A .

$P(A)$: probabilidade de A .

\mathbb{R} : conjunto dos números reais.

\mathbb{A} : álgebra de eventos.

\emptyset : conjunto vazio ou evento impossível.

(Ω, \mathbb{A}, P) : espaço de probabilidade.

X, Y : variáveis aleatórias.

F_X : função de distribuição acumulada da variável aleatória X .

p_X : função de probabilidade ou função de frequência de X .

f_X : função de densidade de probabilidade de X .

$U(a, b)$: distribuição uniforme de probabilidade no intervalo $[a, b]$.

$N(\mu, \sigma^2)$: distribuição normal de probabilidade com média μ e variância σ^2 .

$\text{Exp}(\lambda)$: distribuição exponencial com parâmetro λ .

$E\{X\}$: esperança da variável aleatória X .

$\text{Var}\{X\}$: variância da variável aleatória X .

μ_X : média da variável aleatória X .

σ_X : desvio-padrão da variável aleatória X .

m_X^k : momento de k -ésima ordem da variável aleatória X .

μ_X^k : momento central de k -ésima ordem da variável aleatória X .

$M_X(t)$: função geradora de momentos da variável aleatória X .

\tilde{X} : vetor aleatório ou variável aleatória N -dimensional

$F_{\tilde{X}}$: função de distribuição conjunta do vetor aleatório \tilde{X} .

$p_{\tilde{X}}$: função de frequência conjunta ou função de probabilidade conjunta de \tilde{X} .

$f_{\tilde{X}}$: função de densidade conjunta de \tilde{X} .

$\text{Cov}(X, Y)$: covariância entre X e Y .

$m_{\tilde{X}}^{(n_1, \dots, n_N)}$: momento de ordem $n = (n_1, n_2, \dots, n_N)$ do vetor aleatório \tilde{X} .

I : intervalo aberto da reta real.

g : função contínua.

$G(x)$: série de Taylor da função g .

$g^{(n)}$: derivada de n -ésima ordem de g .

$P_N(x)$: aproximação de Taylor de g ordem N .

$R_N(x)$: resto da aproximação de Taylor de ordem N .

$E_N(x)$: erro da aproximação de Taylor de ordem N .

R : raio de convergência da série de Taylor $G(x)$.

$\frac{\partial^{(n)} g}{\partial x_j^{(n)}}$: derivada parcial de n -ésima ordem de g em relação a x_j .

\tilde{x} : elemento do espaço \mathbb{R}^N .

TLM: *Transmission-Line Modeling Method*. Técnica de simulação.

Γ_1, Γ_2 : coeficientes da fronteira de um TLM.

${}_x\lambda_1, {}_x\lambda_2$: admitâncias de um nó TLM.

${}_xVL^i(t)$: voltagem incidente pela esquerda em um nó TLM x no instante de tempo t .

${}_xVL^r(t)$: voltagem refletida pela esquerda em um nó TLM x no instante de tempo t .

${}_xVR^i(t)$: voltagem incidente pela direita em um nó TLM x no instante de tempo t .

${}_xVR^r(t)$: voltagem refletida pela direita em um nó TLM x no instante de tempo t .

$w(x)$: função de peso.

x_i : i -ésimo ponto-sigma.

w_i : i -ésimo peso.

E : erro da quadratura.

M_r : r -ésimo momento da função de peso $w(x)$.

$\langle f, g \rangle$: produto interno de f por g .

$\{p_n\}$: seqüência de polinômios ortogonais.

$\pi_m(x)$: polinômio cujas raízes são os m pontos-sigma x_i .

α_i : coeficiente do termo x^i do polinômio $\pi_m(x)$.

N : *newton*, unidade de medida de força. Corresponde a força que faz um objeto de 1kg ser acelerado a 1m/s^2 .

F : força exercida em cargas puntiformes, em *newtons*.

k : constante eletrostática.

A : *ampère*, unidade de medida de intensidade de corrente elétrica.

C : *coulomb*, unidade de medida de carga. Corresponde a quantidade de carga elétrica carregada por uma corrente de 1 *ampère* durante um segundo.

e : carga elementar.

q_i : módulo da carga puntiforme, em *coulombs*.

r : distância entre cargas puntiformes, em metros.

Ω : *ohm*, unidade de medida de resistência elétrica.

R : resistência, em *ohms*.

V : *volt*, unidade de medida de tensão elétrica.

V : diferença de potencial elétrico, em *volts*.

i : corrente elétrica, em *ampères*.

UT : *Unscented Transform*. Técnica para o cálculo de momentos de variáveis aleatórias.

UTG : *Unscented Transform Gaussiana*. Técnica para o cálculo de momentos de variáveis aleatórias.

$f^{-1}(x)$: função inversa de $f(x)$.

$\chi(N)^2$: distribuição qui-quadrado com N graus de liberdade.

$J(\tilde{x}, \tilde{y})$: jacobiano de \tilde{x} por \tilde{y} .

$|x|$: módulo ou valor absoluto de x .

$\det A$: determinante da matriz A .

Δ_I^i : subintervalo do intervalo I .

Δ_I : partição do intervalo I .

$\{\delta_i\}_M$: conjunto de representes da partição Δ_I .

$\|\Delta_I^i\|$: comprimento do intervalo Δ^i .

\cup : união.

J : *joule*, unidade de medida de trabalho.

U_B : energia total armazenada em um campo magnético, em *joules*.

H : *henry*, unidade de medida de indutância.

L : indutância, em *henrys*.

\mathcal{E} : fem, em *volts*.

τ_L : constante de tempo indutiva.

F : *farad*, unidade de medida de capacitância.

C : capacitância, em *farads*.

ϵ_0 : constante de permissividade.

d : distância, em metros.

A : área, em metros quadrados.

1 INTRODUÇÃO

As ciências naturais pesquisam e desenvolvem fórmulas e métodos para o estudo dos fenômenos da natureza. Entretanto, se este estudo é baseado numa abordagem determinística, os erros e as incertezas que são comuns à tais fenômenos não serão refletidas ou mesmo consideradas.

Experimentos práticos são suscetíveis à imprecisão dos aparelhos utilizados, aos erros de medição dos instrumentos, às incertezas provenientes do ambiente onde os mesmos são realizados e a influência de variáveis desconhecidas ou mesmo ignoradas. Cada iteração do experimento é um *evento* à parte, com um certo *resultado* que é decorrente de todos os fatores já citados.

Em um sistema, os erros associados às variáveis de entrada alteram e influenciam tanto os valores quanto os erros associados às variáveis de saída. O ato de ignorar ou desprezar tais erros tem como consequência a produção de resultados imprecisos, que hoje são tratados através de aproximações, estimativas e margens de confiança baseadas em dados empíricos ou suposições.

As perguntas que surgem são: quais são os valores esperados para as variáveis de saída? Como determinar uma faixa de valores para os quais os resultados de um experimento sejam confiáveis? As respostas para estas questões podem contrariar a intuição em muitos casos, tornando necessária uma abordagem que seja fundamentada nos princípios da matemática e da probabilidade.

Por exemplo, pela fórmula da resistência, a corrente i é o quociente entre a diferença de potencial V pela resistência R . Porém, se consideramos erros nas medições tanto da diferença de potencial quanto da resistência, a corrente calculada também terá um erro associado.

Considere o seguinte experimento: sortear aleatoriamente 10.000 medições para a diferença de potencial V , cuja média seja $5V$ e que estejam distribuídas uniformemente em um determinado intervalo (isto é, cada ponto do intervalo tem a mesma probabilidade de ocorrer), sortear 10.000 medições para a resistência R , com média

1Ω , também uniformemente distribuídas e calcular a média e a variância das 10.000 correntes i computadas. A tabela 1.1 apresenta os resultados deste experimento para diferentes intervalos, onde m_i é a média de i e σ_i^2 é a variância de i (para as definições de média e variância, veja o capítulo 2).

Tabela 1.1: Resultados do experimento da medição da corrente i

V	R	m_i	σ_i^2
$5 \pm 0,2$	$1 \pm 0,2$	5,0661	0,3673
$5 \pm 0,5$	$1 \pm 0,2$	5,0745	0,4419
$5 \pm 0,2$	$1 \pm 0,5$	5,4936	3,1787

O experimento mostra que, ao contrário do que intuição aponta, a média do quociente não é o quociente das médias, pois nos três casos apresentados na tabela 1.1 a média de V e R eram as mesmas, sendo alterados apenas os erros associados, e a média de i foi diferente para cada situação.

Neste trabalho estudamos justamente esta influência dos erros nos resultados, isto é, a forma como os erros se propagam nos sistemas e nas variáveis que o compõem, analisando as interrelações entre os diversos erros e as expressões que nos permitem estimá-los. Em especial, este estudo será realizado visando problemas do eletromagnetismo, embora as técnicas utilizados para resolvê-los não se limitem a esta área de conhecimento.

O capítulo 2 apresenta os conceitos elementares de probabilidade que servirão de alicerce teórico para o desenvolvimento do trabalho. São apresentados os conceitos de variáveis aleatórias, funções de distribuição, valor esperado, variância, momentos, vetores e funções de variáveis aleatórias, que são imprescindíveis para a compreensão dos demais capítulos.

O capítulo 3 aborda o problema de se estimar o valor esperado de um vetor aleatório resultante da transformação de vetores aleatórios através do uso da série de Taylor desta transformação. Esta abordagem em seguida é aplicada ao problema do TLM, apresentando suas vantagens e desvantagens.

O capítulo 4 sugere outras formas de se resolver o problema do cálculo do valor esperado de uma variável aleatória. Será apresentada, dentre elas, a *Unscented Transform*, técnica proposta inicialmente em [1], para em seguida propôr uma melhoria deste

método, baseada na teoria de quadraturas, que o generalizada e amplia a precisão de seus resultados. Os diferente métodos serão comparados na prática através da resolução de problemas que envolvem a Lei de Coulomb e a Lei de Ohm.

Por fim, o capítulo 5 descreve métodos analíticos para a determinação da função de densidade de probabilidade de uma variável aleatória, e termina com a proposta de um método numérico para a resolução deste problema. Novamente com o intuito de comparar os resultados das diferentes técnicas, são utilizados exemplos envolvendo a teoria de indutores e de capacitores.

2 PROBABILIDADE: CONCEITOS ELEMENTARES

Este capítulo tem como objetivo apresentar os conceitos de probabilidade que são usados ao longo deste trabalho, com o intuito de dar ao leitor um breve resumo dos principais resultados relacionados à variáveis aleatórias, esperança, variância e funções de densidade de probabilidades.

Para um estudo mais completo e detalhado sobre probabilidade e variáveis aleatórias, consulte [2].

2.1 VARIÁVEIS ALEATÓRIAS

2.1.1 Espaço de probabilidade

Ao realizarmos, repetidas vezes, um dado experimento com condições, parâmetros e métodos bem definidos, obtemos um conjunto de resultados possíveis para tal experimento. A este conjunto chamamos *espaço amostral* do experimento.

Definição 2.1. O *espaço amostral* de um experimento é o conjunto Ω , tal que todo resultado possível do experimento corresponde a um, e somente um, ponto $\omega \in \Omega$, e um mesmo ω não pode corresponder a dois resultados distintos.

Supondo, que num dado experimento, nós desejamos observar um determinado subconjunto A do espaço amostral Ω . Se o resultado ω de uma das repetições do experimento pertencer a A , diremos que ω é favorável a A ; caso contrário, diremos que ω não é favorável a A . Daremos ao conjunto A o nome de *evento*.

Definição 2.2. Chamaremos *evento* qualquer subconjunto A do espaço amostral Ω de um experimento. O conjunto unitário $\{\omega\}$ é denominado evento *elementar*, \emptyset é chamado evento *impossível* e Ω é o evento *certo*.

O resultado ω de um experimento pode ou não ser favorável a um evento A ; pode acontecer dos resultados das diferentes iterações nunca serem favoráveis ou sempre serem favoráveis a A ; ou pode ainda acontecer de ω favorecer um evento B mais frequentemente do que favorece A . Vamos supor que possamos associar a cada evento $A \in \Omega$ um número real $P(A)$ que represente a chance de ocorrência de tal evento num experimento (entenderemos como ocorrência o caso de ω ser favorável a A). Aos eventos que pudermos fazer tal associação chamaremos *eventos aleatórios*; a este número denominaremos *probabilidade* de A . Doravante trataremos apenas de eventos aleatórios.

Definição 2.3. Seja $B \in \Omega$ um subconjunto de eventos aleatórios. Chamaremos de *função de probabilidade* P em B , ou simplesmente *probabilidade* em B , a uma função $P : B \rightarrow \mathbb{R}$ tal que

1. $P(A) \geq 0, \forall A \in B$.
2. $P(\Omega) = 1$.
3. Se quaisquer N eventos $A_1, A_2, \dots, A_N \in B$ são disjuntos (isto é, $A_i \cap A_j = \emptyset$ para $i \neq j$), então

$$P\left(\bigcup_{k=1}^N A_k\right) = \sum_{k=1}^N P(A_k). \quad (2.1)$$

Consideremos agora o subconjunto $\mathbb{A} \subset \Omega$ de todos os eventos aleatórios de Ω (isto é, dos eventos $A \in \Omega$ que possuem uma probabilidade $P(A)$ associada), o qual é denominado *classe dos eventos aleatórios* de Ω . Se uma classe de eventos aleatórios contiver o evento certo e for fechada para o complementar (isto é, o complementar de todos os seus elementos também pertencem a ela) e para a união (todas as possíveis uniões entre seus elementos estão contidas nela), a chamaremos *álgebra de eventos*.

Definição 2.4. Seja \mathbb{A} a classe de eventos aleatórios de Ω . Chamaremos \mathbb{A} de *álgebra de eventos* ou *álgebra de subconjuntos de Ω* se \mathbb{A} satisfizer as seguintes condições:

1. $\Omega \in \mathbb{A}$.
2. Se $A \in \mathbb{A}$, então $A^c \in \mathbb{A}$.

3. Se $A, B \in \mathbb{A}$, então $A \cup B \in \mathbb{A}$.

Considere \mathbb{A} uma álgebra de subconjuntos de Ω . Definiremos, para todo $A \in \mathbb{A}$,

$$P(A) = 1 - P(A^c). \quad (2.2)$$

Se $A, B \in \mathbb{A}$ são disjuntos, temos pela equação 2.1 que

$$P(A \cup B) = P(A) + P(B). \quad (2.3)$$

No caso de $A \cap B = C \neq \emptyset$, considere $A_1 = A - C$ e $B_1 = B - C$. Teremos $P(A) = P(A_1) + P(C)$, $P(B) = P(B_1) + P(C)$ (pois A_1, B_1 e C são disjuntos), e

$$\begin{aligned} P(A \cup B) &= P(A_1 \cup B_1 \cup C) \\ &= P(A_1) + P(B_1) + P(C) \\ &= P(A) - P(C) + P(B) - P(C) + P(C) \\ &= P(A) + P(B) - P(C) \end{aligned} \quad (2.4)$$

Deste modo,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \quad (2.5)$$

Somando os conceitos apresentados até o momento, podemos definir nosso experimento pelo espaço amostral Ω de seus possíveis resultados, a álgebra \mathbb{A} de eventos aleatórios e a função de probabilidade P que define a chance de ocorrência de cada um destes eventos. Esta abordagem pode ser formalizada matematicamente através do conceito de *espaço de probabilidade*, que será a base formal para os conceitos e para as idéias desenvolvidas ao longo de todo este trabalho.

Definição 2.5. Chamamos *espaço de probabilidade* (Ω, \mathbb{A}, P) a junção de três componentes: espaço amostral Ω não-nulo, álgebra de eventos \mathbb{A} e função de probabilidade P em \mathbb{A} .

2.1.2 Definição de variáveis aleatórias

Podemos, ao realizarmos um certo experimento, associar valores numéricos ao seus possíveis resultados. Estes valores podem ser utilizados para uma melhor identificação dos resultados possíveis ou podem mesmo ser a informação que se pretende obter a partir do experimento.

Considere o seguinte experimento: o lançamento simultâneo de dois dados de seis faces equilibrados. Teremos o espaço amostral

$$\Omega = \{(i, j) : 1 \leq i, j, \leq 6, i, j \in \mathbb{N}\} \quad (2.6)$$

e, se levarmos em conta a ordem dos dados na listagem dos resultados, (isto é, considerar (2,3) e (3,2) como resultados distintos), $P(i, j) = 1/36$, $\forall(i, j) \in \Omega$. Vamos agora supor que estejamos interessados na soma dos valores obtidos nos dados. Definiremos, então, a função $X : \Omega \rightarrow \mathbb{R}$, tal que $X(i, j) = i + j$. A função X é dependente de um evento aleatório do espaço amostral, então chamaremos X de *variável aleatória*.

Uma definição mais formal de variáveis aleatórias pode ser encontrada em [2], e será reproduzida abaixo. Para tal, considere o evento $[X \leq x] \stackrel{def}{=} \{\omega \in \Omega : X(\omega) \leq x\}$.

Definição 2.6. Uma *variável aleatória* X em um espaço de probabilidade (Ω, \mathbb{A}, P) é uma função real definida no espaço Ω tal que $[X \leq x]$ é evento aleatório para todo $x \in \mathbb{R}$, isto é, $X : \Omega \rightarrow \mathbb{R}$ é variável aleatória se $[X \leq x] \in \mathbb{A}$, $\forall x \in \mathbb{R}$.

Como o evento $[X \leq x]$ pertence a \mathbb{A} , podemos calcular sua probabilidade. Se tomarmos $x_1, x_2 \in \mathbb{R}$, com $x_1 \leq x_2$, teremos $P(X \leq x_1) \leq P(X \leq x_2)$. Deste modo, a medida que avançamos para a direita na reta real, a probabilidade evento $[X \leq x]$ aumenta, tendendo ao limite de 1 (por outro lado, se avançarmos para a esquerda a probabilidade tenderá a zero). Definiremos um função F que associará, para cada ponto x da reta real, a probabilidade do evento $[X \leq x]$ ocorrer. Devido as características citadas neste parágrafo, a denominaremos *função de distribuição acumulada*, ou simplesmente, *função de distribuição* de X .

Definição 2.7. Chamamos *função de distribuição acumulada* F_X da variável aleatória X a função $F_X : \mathbb{R} \rightarrow \mathbb{R}$ tal que

$$F_X(x) = P(X \leq x). \quad (2.7)$$

Uma função de distribuição F_X de uma variável aleatória X possui as seguintes propriedades, que refletem os resultados discutidos no parágrafo que antecede a Definição 2.7:

P1. Se $x \leq y$, então $F_X(x) \leq F_X(y)$.

P2. Se a seqüência x_n tende a x pela direita, então $F_X(x_n)$ tende a $F_X(x)$ pela direita.

P3. Se a seqüência x_n tende a $-\infty$ pela direita, então $F_X(x_n)$ tende a 0 pela direita; se a seqüência x_n tende a ∞ pela esquerda, então $F_X(x_n)$ tende a 1 pela esquerda.

A propriedade P1, a qual nos motivou à definição de função de distribuição acumulada, diz que F_X é não-decrescente; a propriedade P2, também já discutida, informa que F_X é contínua à direita; de acordo com a propriedade P3, podemos definir $F(-\infty) = 0$ e $F(\infty) = 1$.

Qualquer função F que satisfaça as propriedades P1, P2 e P3 é função de distribuição de probabilidade de alguma variável aleatória X , e uma mesma função F pode ser função de distribuição de variáveis aleatórias distintas. A demonstração destes dois importantes fatos, a qual é baseada em conceitos da Teoria da Medida, pode ser encontrada em [3].

2.2 CARACTERIZAÇÃO DAS VARIÁVEIS ALEATÓRIAS

Na seção anterior vimos, além da definição de variável aleatória, o conceito de função de distribuição, que caracteriza uma variável aleatória de acordo com probabilidade de ocorrência de certos tipos de evento. Nesta seção veremos outras maneiras de se caracterizar variáveis aleatórias.

2.2.1 Funções de probabilidade

As variáveis aleatórias, por serem definidas como funções, podem ser classificadas como *discretas*, *contínuas* ou *descontínuas*.

Definição 2.8. Uma variável aleatória X é *discreta* se o conjunto imagem $X(\Omega)$ é finito ou enumerável.

Sendo X uma variável aleatória discreta, podemos escrever $X(\Omega) = \{x_1, x_2, \dots\}$. Neste caso, é possível computar, para cada $x_i \in X(\Omega)$, a probabilidade de ocorrência do evento $[X = x_i]$. A função que associa os elementos do conjunto imagem de X com sua probabilidade é conhecida como *função de probabilidade* de X .

Definição 2.9. Seja X uma variável aleatória discreta. Denominamos *função de probabilidade* ou *função de frequência* de X a função $p_X : X(\Omega) \longrightarrow \mathbb{R}$ tal que

$$p_X(x_i) = P(X = x_i), \quad \forall i = 1, 2, \dots \quad (2.8)$$

Uma vez que o conjunto domínio de uma variável aleatória não é necessariamente um conjunto numérico (pois é um conjunto de eventos aleatórios), não podemos recorrer diretamente à definição matemática de continuidade usando limites. Podemos então definir a continuidade de uma variável aleatória X através da existência de uma função auxiliar que satisfaz uma determinada propriedade ([2], pg. 42):

Definição 2.10. A variável aleatória X é denominada *contínua* se existe uma função $f_X : X(\Omega) \longrightarrow \mathbb{R}$ tal que $f_X(x) \geq 0, \forall x \in X(\Omega)$ e

$$F_X(x) = \int_{-\infty}^x f_X(t)dt, \quad \forall x \in \mathbb{R}. \quad (2.9)$$

Neste caso, dizemos que f_X é *função de densidade de probabilidade* de X ou simplesmente *densidade* de X .

A função de densidade de probabilidade, assim como a função de frequência no caso discreto, são características que representam a variável aleatória, definindo o comportamento da mesma. O conhecimento de ambas é fundamental no estudo de problemas que envolvam variáveis aleatórias.

Assim como no caso da função de distribuição, determinadas funções podem ser funções de densidade de alguma variável aleatória. A proposição abaixo apresenta uma condição necessária e suficiente para que isto aconteça.

Proposição 2.11. Seja $I \subset \mathbb{R}$ e $f : I \longrightarrow \mathbb{R}$ uma função não-negativa, ou seja, $f(x) \geq 0, \forall x \in I$. A função f é função de densidade de probabilidade de alguma variável aleatória X se, e somente se,

$$\int_{-\infty}^{\infty} f(t)dt = 1. \quad (2.10)$$

Demonstração: Considere que f seja a função de densidade de X . Então temos que

$$F_X(x) = \int_{-\infty}^x f(t)dt. \quad (2.11)$$

Definimos anteriormente que $F_X(\infty) = 1$, daí

$$1 = F(\infty) = \int_{-\infty}^{\infty} f(t)dt. \quad (2.12)$$

Por outro lado, suponha agora que

$$\int_{-\infty}^{\infty} f(t)dt = 1 \quad (2.13)$$

e defina

$$F(x) = \int_{-\infty}^x f(t)dt. \quad (2.14)$$

Tome $x, y \in \mathbb{R}$ tais que $x \leq y$. Segue que

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t)dt \\ &\leq \int_{-\infty}^x f(t)dt + \int_x^y f(t)dt \\ &\leq \int_{-\infty}^y f(t)dt \\ &\leq F(y), \end{aligned} \quad (2.15)$$

logo vale a propriedade P1. Considera agora a sequência x_n que converge para x pela direita. Assim,

$$\lim_{x_n \rightarrow x^+} F(x_n) = \lim_{x_n \rightarrow x^+} \int_{-\infty}^{x_n} f(t)dt = \int_{-\infty}^x f(t)dt = F(x), \quad (2.16)$$

o que mostra a propriedade P2. Por último,

$$\lim_{x_n \rightarrow \infty^-} F(x_n) = \lim_{x_n \rightarrow \infty^-} \int_{-\infty}^{x_n} f(t)dt = \int_{-\infty}^{\infty} f(t)dt = 1, \quad (2.17)$$

por hipótese, e

$$\lim_{x_n \rightarrow -\infty^+} F(x_n) = \lim_{x_n \rightarrow -\infty^+} \int_{-\infty}^{x_n} f(t)dt = \int_{-\infty}^{-\infty} f(t)dt = 0, \quad (2.18)$$

o que verifica a propriedade P3. Logo, f é função de densidade de X , o que conclui a demonstração.

□

Exemplo 2.12. Como exemplo concreto de função de densidade de probabilidade, considere um intervalo $I = [a, b]$. Se para quaisquer $x, y \in I$ as probabilidades $P(X = x)$ e $P(X = y)$ forem iguais, dizemos neste caso que X é *uniformemente distribuído* em I , ou que X tem *distribuição uniforme de probabilidade* em I . A sua função de densidade é dada por

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{caso contrário.} \end{cases} \quad (2.19)$$

Para indicarmos este caso, a notação é $X \sim U(a, b)$. Observe que, conforme esperado, a função f_X é constante em I , nula fora do intervalo e satisfaz a proposição 2.11.

Exemplo 2.13. Um outro exemplo de função densidade é a função

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad x \in \mathbb{R}. \quad (2.20)$$

Se X é uma variável aleatória com esta função de densidade, dizemos que X tem *distribuição normal de probabilidade com média μ e variância σ^2* . Neste caso, escrevemos $X \sim N(\mu, \sigma^2)$.

Ao contrário da distribuição uniforme, a probabilidade de ocorrência de um valor x aumenta a medida que ele se aproxima da média μ , e o coeficiente σ (denominado *desvio-padrão*) mede a dispersão do conjunto dos valores mais prováveis em torno da média (quanto maior, mais dispersos os valores). Observe que neste exemplo a função f_X também satisfaz a proposição 2.11.

Exemplo 2.14. Considere um número real $\lambda > 0$. Dizemos que uma variável aleatória contínua X tem *distribuição exponencial com parâmetro λ* se tem função de distribuição

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{caso contrário.} \end{cases} \quad (2.21)$$

Assim como nos exemplos anteriores, a função f_X também satisfaz a proposição 2.11. A notação para este caso é $X \sim \text{Exp}(\lambda)$.

Nesta seção vimos que a tanto a função de distribuição quanto as funções de frequência e densidade caracterizam uma variável aleatória. Na duas próximas seções estudaremos grandezas que também revelam características e comportamentos das variáveis aleatórias.

2.2.2 Esperança Matemática

Ao realizarmos um experimento, mesmo levando em consideração os erros provenientes de medidas, aproximações, condições de temperatura e pressão e/ou quaisquer outros fatores, esperamos obter um certo resultado. Podemos formalizar esta expectativa, ou esperança, num coeficiente computado a partir das variáveis aleatórias envolvidas no experimento.

Antes da definição formal, apresentamos as idéias básicas com um exemplo.

Exemplo 2.15. Considere o seguinte experimento: o lançamento de um dado equilibrado. Temos o espaço amostral $\Omega = \{1, 2, 3, 4, 5, 6\}$, de modo que podemos definir a variável aleatória $X(w) = w$, com função de frequência $p_X(x_i) = 1/6, \forall x_i \in X(\Omega)$. Se repetirmos este experimento N vezes seguidas, o que se espera é obter cada resultado possível a mesma quantidade de vezes, de tal maneira que cada x_i ocorra $N/6$ vezes. A média aritmética dos resultados do experimento seria

$$\frac{1}{N} \sum_{i=1}^6 \frac{N}{6} x_i = \sum_{i=1}^6 \frac{1}{6} x_i = \sum_{i=1}^6 p_X(x_i) \cdot x_i. \quad (2.22)$$

Teríamos o mesmo somatório mesmo se o dado fosse desequilibrado (isto é, se algumas faces ocorressem com maior probabilidade do que as outras), de modo que ele representaria então a média ponderada dos resultados. O *valor esperado*, ou a *esperança matemática* do experimento seria justamente esta soma.

Definição 2.16. Seja X uma variável aleatória discreta, com função de frequência p_X e espaço amostral Ω . A *esperança matemática* ou *valor esperado* $E\{X\}$ de X é dada por

$$E\{X\} = \sum_{x_i \in X(\Omega)} p_X(x_i) \cdot x_i = \sum_{x_i \in X(\Omega)} P(X = x_i) \cdot x_i \quad (2.23)$$

Embora tenhamos usado variáveis discretas para apresentar o conceito de esperança, ele segue o mesmo princípio para variáveis contínuas, como podemos observar na definição abaixo.

Definição 2.17. Seja X uma variável aleatória contínua, com função de densidade de probabilidade f_X . A *esperança matemática* ou *valor esperado* $E\{X\}$ de X é dada

por

$$E\{X\} = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-\infty}^{\infty} x dF_X(x) \quad (2.24)$$

Na definição 2.16, a esperança pode ser computada tanto pela função de frequência quanto pela função de probabilidade. Já a definição 2.17 requer o conhecimento da função de densidade ou da derivada da função de distribuição. A proposição abaixo mostra a forma de se computar a esperança diretamente a partir da função de distribuição de probabilidade.

Proposição 2.18. Seja X uma variável aleatória contínua com função de distribuição de probabilidade F_X . Então

$$E\{X\} = \int_0^{\infty} (1 - F_X(x)) dx - \int_{-\infty}^0 F_X(x) dx \quad (2.25)$$

Demonstração (retirada de [2], pg. 112):

Vamos provar que

$$\int_0^{\infty} x dF_X(x) = \int_0^{\infty} (1 - F_X(x)) dx. \quad (2.26)$$

Sabendo que $d(xF_X(x)) = x dF_X(x) + F_X(x) dx$ e integrando por partes, temos para todo $b > 0$ que

$$\int_0^b x dF_X(x) = bF_X(b) - \int_0^b F_X(x) dx = \int_0^b [F_X(b) - F_X(x)] dx \quad (2.27)$$

Como $F_X(b) \leq 1$ e $1 - F_X(x) \geq 0$, temos

$$\int_0^b x dF_X(x) = \int_0^b [F_X(b) - F_X(x)] dx \leq \int_0^{\infty} [1 - F_X(x)] dx, \quad \forall b > 0, \quad (2.28)$$

de modo que

$$\int_0^{\infty} x dF_X(x) = \lim_{b \rightarrow \infty} \int_0^b x dF_X(x) \leq \int_0^{\infty} [1 - F_X(x)] dx. \quad (2.29)$$

Por outro lado, seja $\lambda > 0$. Se $b > \lambda$, então

$$\begin{aligned}
\int_0^b [F_X(b) - F_X(x)]dx &\geq \int_0^\lambda [F_X(b) - F_X(x)]dx \\
&= \int_0^\lambda [F_X(b) - 1]dx + \int_0^\lambda [1 - F_X(x)]dx \\
&= \lambda[F_X(b) - 1] + \int_0^\lambda [1 - F_X(x)]dx
\end{aligned} \tag{2.30}$$

e portanto,

$$\begin{aligned}
\int_0^\infty x dF_X(x) &= \lim_{b \rightarrow \infty} \int_0^b x dF_X(x) \\
&= \lim_{b \rightarrow \infty} \int_0^b [F_X(b) - F_X(x)]dx \\
&\geq \int_0^\lambda [1 - F_X(x)]dx + \lim_{b \rightarrow \infty} \lambda[F_X(b) - 1] \\
&= \int_0^\lambda [1 - F_X(x)]dx.
\end{aligned} \tag{2.31}$$

Já que isto vale para todo $\lambda > 0$, temos

$$\int_0^\infty x dF_X(x) \geq \lim_{\lambda \rightarrow \infty} \int_0^\lambda [1 - F_X(x)]dx = \int_0^\infty [1 - F_X(x)]dx. \tag{2.32}$$

Combinando as equações (2.29) e (2.32), temos

$$\int_0^\infty [1 - F_X(x)]dx \leq \int_0^\infty x dF_X(x) \leq \int_0^\infty [1 - F_X(x)]dx, \tag{2.33}$$

ou seja,

$$\int_0^\infty x dF_X(x) = \int_0^\infty [1 - F_X(x)]dx, \tag{2.34}$$

Falta apenas mostrar que

$$\int_{-\infty}^0 x dF_X(x) = - \int_{-\infty}^0 F_X(x)dx, \tag{2.35}$$

mas como a demonstração é análoga ao caso anterior, a mesma será deixada à cargo do leitor.

□

A próxima proposição apresenta as propriedades da esperança matemática.

Proposição 2.19. Sejam X e Y variáveis aleatórias em um espaço amostral Ω e $a, b, c \in \mathbb{R}$. Valem as seguintes afirmações:

E1. Se $X(\omega) = c, \forall \omega \in \Omega$, então $E\{X\} = c$.

E2. Se $X(\omega) \leq Y(\omega), \forall \omega \in \Omega$, então $E\{X\} \leq E\{Y\}$.

E3. $E\{X\}$ é um operador linear, isto é, $E\{aX + b\} = aE\{X\} + b$.

Demonstração: (E1.) Neste caso, X é uma variável aleatória discreta com $X(\Omega) = \{c\}$. Logo

$$E\{X\} = \sum_{x_i \in X(\Omega)} P(X = x_i) \cdot x_i = P(X = c) \cdot c = 1 \cdot c = c. \quad (2.36)$$

(E2.) Tome $z \in \mathbb{R}$. Se $Y \leq z$, então $X \leq z$, de modo que $[Y \leq z] \subset [X \leq z]$. Logo, $F_Y(z) \leq F_X(z)$ (pela propriedade P1) e $1 - F_Y(z) \geq 1 - F_X(z)$. De acordo com a proposição 2.18,

$$\begin{aligned} E\{Y\} &= \int_0^\infty (1 - F_Y(z))dz - \int_{-\infty}^0 F_Y(z)dz \\ &\geq \int_0^\infty (1 - F_X(z))dz - \int_{-\infty}^0 F_X(z)dz \\ &= E\{X\}. \end{aligned} \quad (2.37)$$

(E3.) (Demonstração retirada de [2], pg. 116) Se $a = 0$, então

$$E\{aX + b\} = E\{b\} = b = aE\{X\} + b. \quad (2.38)$$

Se $a > 0$, então

$$F_{aX+b} = P(aX + b \leq x) = P\left(X \leq \frac{x-b}{a}\right) = F_X\left(\frac{x-b}{a}\right), \quad (2.39)$$

logo, de acordo com a proposição 2.18, e fazendo $y = (x - b)/a$, temos

$$\begin{aligned}
 E\{aX + b\} &= \int_0^\infty (1 - F_{aX+b}(x))dx - \int_{-\infty}^0 F_{aX+b}(x)dx \\
 &= \int_0^\infty \left(1 - F_X\left(\frac{x-b}{a}\right)\right) dx - \int_{-\infty}^0 F_X\left(\frac{x-b}{a}\right) dx \\
 &= a \int_{-b/a}^\infty (1 - F_X(y))dy - \int_{-\infty}^{-b/a} F_X(y)dy \\
 &= a \int_0^\infty (1 - F_X(y))dy - a \int_{-\infty}^0 F_X(y)dy \\
 &\quad + a \int_{-b/a}^0 (1 - F_X(y))dy + a \int_{-b/a}^0 F_X(y)dy \\
 &= aE\{X\} + a \int_{-b/a}^0 dy \\
 &= aE\{X\} + b.
 \end{aligned} \tag{2.40}$$

O caso $a < 0$ é análogo.

□

Exemplo 2.20. Seja $X \sim U(a, b)$ (ver exemplo 2.12). De acordo com a definição 2.17, temos

$$\begin{aligned}
 E\{X\} &= \int_{-\infty}^\infty x f_X(x) dx \\
 &= \int_a^b x \cdot \frac{1}{b-a} dx \\
 &= \frac{1}{b-a} \cdot \frac{x^2}{2} \Big|_a^b \\
 &= \frac{b^2 - a^2}{2(b-a)} \\
 &= \frac{(b-a)(b+a)}{2(b-a)} \\
 &= \frac{b+a}{2}
 \end{aligned} \tag{2.41}$$

Logo, a esperança de uma variável aleatória X uniformemente distribuída no intervalo $[a, b]$ é igual a $(b + a)/2$.

Exemplo 2.21. Seja $X \sim N(\mu, \sigma^2)$ (ver exemplo 2.13). De acordo com a definição 2.17, temos

$$\begin{aligned}
 E\{X\} &= \int_{-\infty}^{\infty} x f_X(x) dx \\
 &= \int_{-\infty}^{\infty} x \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx \\
 &= \int_{-\infty}^{\infty} (u + \mu) \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-u^2/2\sigma^2} du \\
 &= \int_{-\infty}^{\infty} u \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-u^2/2\sigma^2} du + \mu \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-u^2/2\sigma^2} du \\
 &= 0 + \mu \cdot 1 = \mu
 \end{aligned} \tag{2.42}$$

Logo, o parâmetro μ da distribuição normal de probabilidades é, de fato, a esperança de X .

Exemplo 2.22. Seja $X \sim \text{Exp}(\lambda)$ (ver exemplo 2.14). De acordo com a definição 2.17, e integrando por partes, temos

$$\begin{aligned}
 E\{X\} &= \int_{-\infty}^{\infty} x f_X(x) dx \\
 &= \int_0^{\infty} x \cdot \lambda e^{-\lambda x} dx \\
 &= -x e^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx \\
 &= 0 - \frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} \\
 &= \frac{1}{\lambda}.
 \end{aligned} \tag{2.43}$$

Logo, a esperança de uma variável aleatória X com distribuição exponencial de parâmetro λ é igual a $1/\lambda$.

2.2.3 Momentos

Momentos são grandezas associadas a uma variável aleatória X que caracterizam a mesma, dando, dentre outras, informações sobre seu valor esperado (a esperança matemática é um momento), sua medida de dispersão e a forma da sua função de densidade de probabilidade. A esperança matemática foi introduzida anteriormente porque os momentos podem ser definidos como esperanças de certas funções de X , como podemos ver na definição abaixo.

Definição 2.23. Definimos como *k*-ésimo momento da variável aleatória X em torno de b o valor

$$m_X^k(b) = E\{(X - b)^k\}, \quad (2.44)$$

onde $k \in \mathbb{N}$ e $b \in \mathbb{R}$.

Se $b = 0$, chamamos $m_X^k = m_X^k(0) = E\{X^k\}$ de *k*-ésimo momento de X ou *momento de ordem k de X*; se $b = E\{X\}$, denominamos $\mu_X^k = m_X^k(E\{X\}) = E\{(X - E\{X\})^k\}$ de *k*-ésimo momento central de X .

Os momentos de baixa ordem ($k = 1, 2, 3, 4$) são de grande interesse e utilidade. Se X é uma variável aleatória, o primeiro momento de m_X de X é a esperança matemática, pois

$$m_X = m_X^1 = E\{X^1\} = E\{X\}, \quad (2.45)$$

enquanto o primeiro momento central μ_X de X é zero, pois pelas propriedades E1 e E3,

$$\mu_X = E\{X - E\{X\}\} = E\{X\} - E\{E\{X\}\} = E\{X\} - E\{X\} = 0. \quad (2.46)$$

O segundo momento central é denominado *variância* de X , e é uma medida de dispersão estatística, no sentido que mensura a distância entre os resultados obtidos em um experimento e o valor esperado.

Definição 2.24. A *variância* $\text{Var}\{X\}$ (ou σ_X^2) de X é a denominação para o segundo momento central da variável aleatória X . O coeficiente σ_X , definido como

$$\sigma_X = \sqrt{\sigma_X^2} \quad (2.47)$$

é denominado o *desvio-padrão* de X .

O desvio-padrão pode ser relacionado como a largura do intervalo de confiança de uma variável aleatória, uma vez que em geral os seus valores tendem a se encontrar no intervalo $(m_X - \sigma_X, m_X + \sigma_X)$ com uma maior probabilidade do que fora deste intervalo.

A variância pode ser computada a partir da definição do segundo momento central da seguinte forma

$$\begin{aligned}
\text{Var}\{X\} &\stackrel{def}{=} E\{(X - E\{X\})^2\} \\
&= E\{X^2 - 2XE\{X\} + (E\{X\})^2\} \\
&= E\{X^2\} - 2E\{X\}E\{X\} + (E\{X\})^2 \\
&= E\{X^2\} - (E\{X\})^2.
\end{aligned} \tag{2.48}$$

Dois propriedades da variância serão apresentadas na proposição abaixo.

Proposição 2.25. Seja X uma variável aleatória definida em Ω e $a, b, c \in \mathbb{R}$. Então valem as seguintes afirmações.

V1. Se $X(\omega) = c, \forall \omega \in \Omega$, então $\text{Var}\{X\} = 0$.

V2. $\text{Var}\{aX + b\} = a^2\text{Var}\{X\}$.

Demonstração: (V1.) Por hipótese, $X(\omega) = c, \forall \omega \in \Omega$, logo

$$\text{Var}\{X\} = E\{(X - c)^2\} = E\{0\} = 0. \tag{2.49}$$

(V2.) Como visto na propriedade E3 da proposição 2.19, $E\{aX + b\} = aE\{X\} + b$. Então

$$\begin{aligned}
\text{Var}\{aX + b\} &= E\{(aX + b - aE\{X\} - b)^2\} \\
&= E\{a^2(X - E\{X\})^2\} \\
&= a^2E\{(X - E\{X\})^2\} \\
&= a^2\text{Var}\{X\}.
\end{aligned} \tag{2.50}$$

□

Exemplo 2.26. Seja X uma variável aleatória com distribuição uniforme em $[a, b]$. Temos, conforme a equação 2.48 e o exemplo 2.20,

$$\begin{aligned}
 \text{Var}\{X\} &= E\{X^2\} - (E\{X\})^2 \\
 &= \int_a^b x^2 \cdot \frac{1}{b-a} dx - \left(\frac{b+a}{2}\right)^2 \\
 &= \frac{1}{b-a} \cdot \frac{x^3}{3} \Big|_a^b - \left(\frac{b^2+2ab+a^2}{4}\right) \\
 &= \frac{1}{b-a} \cdot \left(\frac{b^3-a^3}{3}\right) - \left(\frac{b^2+2ab+a^2}{4}\right) \\
 &= \left(\frac{b^2+ab+a^2}{3}\right) - \left(\frac{b^2+2ab+a^2}{4}\right) \\
 &= \frac{(b-a)^2}{12}. \tag{2.51}
 \end{aligned}$$

Exemplo 2.27. Seja $X \sim N(\mu, \sigma^2)$. Temos, conforme a equação 2.48 e o exemplo 2.21,

$$\begin{aligned}
 \text{Var}\{X\} &= E\{X^2\} - (E\{X\})^2 \\
 &= \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx - \mu^2 \\
 &= \int_{-\infty}^{\infty} (u+\mu)^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-u^2/2\sigma^2} du - \mu^2 \\
 &= \int_{-\infty}^{\infty} u^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-u^2/2\sigma^2} du + 2 \int_{-\infty}^{\infty} u \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-u^2/2\sigma^2} du \\
 &\quad + \mu^2 \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-u^2/2\sigma^2} du - \mu^2 \\
 &= u \cdot \frac{\sigma}{\sqrt{2\pi}} e^{-u^2/2\sigma^2} \Big|_{-\infty}^{\infty} + \sigma^2 \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-u^2/2\sigma^2} du + 2 \cdot 0 + \mu^2 \cdot 1 - \mu^2 \\
 &= 0 + \sigma^2 + \mu^2 - \mu^2 \\
 &= \sigma^2, \tag{2.52}
 \end{aligned}$$

de modo que o parâmetro σ^2 da distribuição normal é, de fato, sua variância.

Exemplo 2.28. Seja $X \sim \text{Exp}(\lambda)$. Temos, conforme a equação 2.48 e o exemplo 2.22,

$$\begin{aligned}
 \text{Var}\{X\} &= E\{X^2\} - (E\{X\})^2 \\
 &= \int_0^{\infty} x^2 \cdot \lambda e^{-\lambda x} dx - \frac{1}{\lambda^2} \\
 &= -x^2 \cdot e^{-\lambda x} \Big|_0^{\infty} + 2 \int_0^{\infty} x \cdot e^{-\lambda x} dx - \frac{1}{\lambda^2} \\
 &= 0 + \frac{-2x}{\lambda} e^{-\lambda x} \Big|_0^{\infty} + \frac{2}{\lambda} \int_0^{\infty} e^{-\lambda x} dx - \frac{1}{\lambda^2} \\
 &= 0 + \frac{2}{\lambda^2} - \frac{1}{\lambda^2} \\
 &= \frac{1}{\lambda^2}.
 \end{aligned} \tag{2.53}$$

Logo, a variância de uma variável aleatória X com distribuição exponencial com parâmetro λ é igual a $1/\lambda^2$.

Uma alternativa para o cálculo dos momentos de n -ésima ordem é o uso da *função geradora de momentos*, definida abaixo.

Definição 2.29. Seja X uma variável aleatória. Denominamos *função geradora de momentos* a função

$$M_X(t) = E\{e^{tX}\}. \tag{2.54}$$

A função geradora de momentos recebe este nome devido ao fato de que, para cada k inteiro positivo,

$$\lim_{t \rightarrow 0} [M^{(k)}(t)] = \lim_{t \rightarrow 0} E\{x^k e^{tx}\} = E\{x^k\} = m_X^k \tag{2.55}$$

quando o limite existir. Em outras palavras, é possível computar os momentos de n -ésima ordem através das derivadas da função geradora de momentos.

Exemplo 2.30. Seja $X \sim U(a, b)$. Para $t \neq 0$, a função geradora de momentos de X é igual a

$$\begin{aligned}
 M_X(t) &= E\{e^{tX}\} \\
 &= \int_b^a \frac{1}{b-a} \cdot e^{xt} dx \\
 &= \frac{1}{b-a} \cdot \frac{1}{t} e^{xt} \Big|_a^b \\
 &= \frac{1}{t(b-a)} (e^{tb} - e^{ta}).
 \end{aligned} \tag{2.56}$$

Para $t = 0$, temos

$$M_X(0) = E\{e^0\} = E\{1\} = 1. \tag{2.57}$$

Os momentos de X podem ser computados através dos limites das n -ésimas derivadas de $M_X(t)$ quando t tende a zero. A fórmula para o cálculo dos momentos de $X \sim U(a, b)$ pode ser computada desta maneira, e é dada por

$$m_X^k = \frac{1}{k+1} \sum_{i=0}^k a^i b^{k-i}. \tag{2.58}$$

Exemplo 2.31. Seja $X \sim N(\mu, \sigma^2)$. A função geradora de momentos de X é igual a

$$\begin{aligned}
 M_X(t) &= E\{e^{tX}\} \\
 &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \cdot e^{xt} dx \\
 &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{xt - [(x-\mu)^2/2\sigma^2]} dx \\
 &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x^2 - 2\sigma^2 tx + 2\mu x + \mu^2)/2\sigma^2} dx \\
 &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x^2 - 2x(\mu + \sigma^2 t) + \mu^2)/2\sigma^2} dx \\
 &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-((x - \mu - \sigma^2 t)^2 - 2\mu\sigma^2 t - \sigma^4 t^2)/2\sigma^2} dx \\
 &= \frac{1}{\sigma\sqrt{2\pi}} e^{\mu t + \sigma^2 t^2/2} \int_{-\infty}^{\infty} e^{-(x - \mu - \sigma^2 t)^2/2\sigma^2} dx \\
 &= \frac{\sigma\sqrt{2}}{\sigma\sqrt{2\pi}} e^{\mu t + \sigma^2 t^2/2} \int_{-\infty}^{\infty} e^{-u^2 \sigma^2} dx \\
 &= \frac{\sqrt{\pi}}{\sqrt{\pi}} e^{\mu t + \sigma^2 t^2/2} \\
 &= e^{\mu t + \sigma^2 t^2/2}.
 \end{aligned} \tag{2.59}$$

A fórmula para o cálculo dos k -ésimos momentos de uma variável X com distribuição normal de média μ e variância σ^2 é dada por

$$m_X^k = \sum_{i=0}^{\lfloor k/2 \rfloor} \binom{k}{k-2i} \frac{(2i)!}{i!2^i} \mu^{k-2i} \sigma^{2i}, \quad (2.60)$$

onde $\lfloor x \rfloor$ é o menor inteiro maior que x .

Exemplo 2.32. Seja $X \sim \text{Exp}(\lambda)$. Para $t < \lambda$, a função geradora de momentos de X é dada por

$$\begin{aligned} M_X(t) &= E\{e^{tX}\} \\ &= \int_0^{\infty} \lambda e^{-\lambda x} \cdot e^{xt} dx \\ &= \lambda \int_0^{\infty} e^{(t-\lambda)x} dx \\ &= \lambda \cdot \frac{1}{t-\lambda} \cdot e^{(t-\lambda)x} \Big|_0^{\infty} \\ &= -\frac{\lambda}{t-\lambda} \\ &= \frac{\lambda}{\lambda-t} \\ &= \frac{1}{1-t/\lambda}. \end{aligned} \quad (2.61)$$

Os k -ésimos momentos da variável X de distribuição exponencial com parâmetro λ podem ser computados através da fórmula

$$m_X^k = \frac{k!}{\lambda^k}. \quad (2.62)$$

Os momentos de n -ésima ordem serão úteis no cálculo da esperança no caso de funções de variáveis aleatórias, seja quando estas forem aproximadas por polinômios, seja quando a esperança for computada por métodos numéricos. Funções de variáveis aleatórias serão apresentadas na próxima seção; aproximações polinômiais no capítulo 3 e uma técnica numérica para a computação da esperança no capítulo 4.

2.3 VETORES E FUNÇÕES DE VARIÁVEIS ALEATÓRIAS

2.3.1 Vetores de variáveis aleatórias

Até então, tratamos de experimentos que envolviam uma única variável aleatória. Contudo, na prática, os experimentos envolvem várias variáveis aleatórias simultaneamente. Se todas elas estiverem definidas num mesmo espaço de probabilidades, podemos organizá-las num vetor aleatório.

Definição 2.33. Considere o espaço de probabilidades (Ω, \mathbb{A}, P) e as variáveis aleatórias X_i , $i = 1, 2, \dots, N$. Se todas as variáveis X_i estiverem definidas neste espaço, chamamos ao vetor

$$\tilde{X} = (X_1, X_2, \dots, X_N) \quad (2.63)$$

de *vetor aleatório* ou *variável aleatória N -dimensional*.

Como cada variável tem sua própria função de distribuição de probabilidades, podemos definir uma função de distribuição para o vetor como um todo, denominada *distribuição conjunta*.

Definição 2.34. Seja $\tilde{X} = (X_1, X_2, \dots, X_N)$ um vetor aleatório. A função $F_{\tilde{X}} = F_{X_1, X_2, \dots, X_N}$, definida como

$$F_{\tilde{X}}(\tilde{x}) = F_{\tilde{X}}(x_1, x_2, \dots, x_N) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_N \leq x_N), \quad (2.64)$$

onde $(x_1, x_2, \dots, x_N) \in \mathbb{R}^N$, é denominada *função de distribuição conjunta* das variáveis aleatórias X_1, X_2, \dots, X_N .

As funções de distribuição conjunta tem as propriedades apresentadas na proposição abaixo. A demonstração das mesmas será deixada a critério do leitor.

Definição 2.35. Seja $\tilde{X} = (X_1, X_2, \dots, X_N)$ um vetor aleatório e $F_{\tilde{X}}$ a função de distribuição conjunta de \tilde{X} . Valem as seguintes afirmativas:

D1. $F_{\tilde{X}}(x_1, \dots, x_N)$ é não-decrescente para cada variável X_i , isto é, se $x < y$, então para cada $i = 1, 2, \dots, N$,

$$F_{\tilde{X}}(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_N) \leq F_{\tilde{X}}(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_N). \quad (2.65)$$

D2. $F_{\tilde{X}}(x_1, \dots, x_N)$ é contínua para cada variável X_i , isto é, se $x_n \rightarrow x$, então para cada $i = 1, 2, \dots, N$,

$$\lim_{x_n \rightarrow x} F_{\tilde{X}}(x_1, \dots, x_{i-1}, x_n, x_{i+1}, \dots, x_N) = F_{\tilde{X}}(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_N). \quad (2.66)$$

D3. Para cada $i = 1, 2, \dots, N$,

$$\lim_{x_i \rightarrow -\infty} F_{\tilde{X}}(x_1, x_2, \dots, x_N) = 0. \quad (2.67)$$

Também temos que

$$\lim_{\forall i, x_i \rightarrow \infty} F_{\tilde{X}}(x_1, x_2, \dots, x_N) = 1. \quad (2.68)$$

D4. Seja $I_k = (a_k, b_k]$, $a_k, b_k \in \mathbb{R}$, $k \in \mathbb{N}$ intervalos semi-abertos e

$$\begin{aligned} \Delta_{I_k} g(x_1, x_2, \dots, x_N) &= g(x_1, x_2, \dots, x_{k-1}, b_k, x_{k+1}, \dots, x_N) - \\ &g(x_1, x_2, \dots, x_{k-1}, a_k, x_{k+1}, \dots, x_N), \end{aligned} \quad (2.69)$$

onde $g : \mathbb{R}^N \rightarrow \mathbb{R}$. Então

$$\Delta_{I_1} \Delta_{I_2} \dots \Delta_{I_N} F(x_1, x_2, \dots, x_N) \geq 0, \quad \forall I_k. \quad (2.70)$$

Uma função $F : \mathbb{R}^N \rightarrow \mathbb{R}$ que satisfizer as propriedades D1, D2, D3 e D4 será função de distribuição conjunto de algum vetor aleatório \tilde{X} .

Uma vez conhecida da função de distribuição conjunta do vetor aleatório \tilde{X} , é possível obter a função de distribuição F_{X_i} de cada variável aleatória $X_i \in \tilde{X}$.

Definição 2.36. Seja \tilde{X} um vetor aleatório com função de distribuição conjunta $F_{\tilde{X}}$. Então a *função de distribuição marginal* de X_i é dada por

$$F_{X_i} = \lim_{x_i \rightarrow \infty} F_{\tilde{X}}(x_1, x_2, \dots, x_i, \dots, x_N) \stackrel{def}{=} F_{\tilde{X}}(x_1, x_2, \dots, x_{i-1}, \infty, x_{i+1}, \dots, x_N), \quad (2.71)$$

onde $X_i \in \tilde{X}$.

Os vetores aleatórios também podem ser classificados como contínuos ou discretos, e terem funções de frequência ou de densidade de probabilidade, de modo análogo ao caso unidimensional.

Definição 2.37. Seja \tilde{X} um vetor aleatório definido no espaço de probabilidade (Ω, \mathbb{A}, P) e $F_{\tilde{X}}$ a sua função de distribuição conjunta.

- (a) Se as variáveis X_i , $i = 1, 2, \dots, N$ são discretas, então \tilde{X} é denominado *vetor aleatório discreto* e para todo $A \in \mathbb{A}$,

$$p_{\tilde{X}}(A) = \sum_{\tilde{x} \in X(A)} P(\tilde{X} = \tilde{x}). \quad (2.72)$$

é a *função de frequência conjunta* ou *função de probabilidade conjunta* de \tilde{X} .

- (b) Se existe uma função $f_{\tilde{X}}(x_1, x_2, \dots, x_N) \geq 0$ tal que, $\forall(x_1, x_2, \dots, x_N) \in \mathbb{R}^N$,

$$F_{\tilde{X}}(x_1, x_2, \dots, x_N) = \int_{-\infty}^{x_N} \dots \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f_{\tilde{X}}(t_1, t_2, \dots, t_N) dt_1 dt_2 \dots dt_N, \quad (2.73)$$

então chamamos \tilde{X} de *vetor aleatório contínuo* com *função de densidade conjunta* $f_{\tilde{X}}$.

Num espaço de probabilidade (Ω, \mathbb{A}, P) , dois eventos aleatórios $A_1, A_2 \in \mathbb{A}$ são ditos *independentes* se $P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$. Podemos expandir este conceito para vetores de variáveis aleatórias.

Definição 2.38. Seja \tilde{X} um vetor aleatório definido no espaço de probabilidade (Ω, \mathbb{A}, P) . As variáveis X_i , $i = 1, 2, \dots, N$ são ditas *independentes* se

$$P(X_1 \in A_1, X_2 \in A_2, \dots, X_N \in A_N) = \prod_{i=1}^N P(X_i \in A_i) \quad (2.74)$$

para todos $A = (A_1, A_2, \dots, A_N) \in \mathbb{A}$.

2.3.2 Funções de variáveis aleatórias

Seja $g : \mathbb{R} \rightarrow \mathbb{R}$ uma função contínua. Se aplicarmos g em uma variável aleatória X , obtemos uma nova variável aleatória $Y = g(X)$. As perguntas que surgem naturalmente são: qual é a distribuição F_Y de Y ? Qual é a esperança $E\{Y\}$ e a variância $\text{Var}\{Y\}$ desta nova variável?

De modo geral, dado um vetor aleatório $\tilde{X} = (X_1, X_2, \dots, X_N)$ e a função contínua $g : \mathbb{R}^N \rightarrow \mathbb{R}^M$, obtemos a vetor aleatório $\tilde{Y} = (Y_1, Y_2, \dots, Y_M)$, onde

$$Y_i = g_i(X_1, X_2, \dots, X_N). \quad (2.75)$$

Nesta seção veremos a abordagem teórica para cada uma das questões anteriores. Começemos pela mais simples, que é determinar a função de probabilidade de \tilde{Y} .

Definição 2.39. Seja \tilde{X} um vetor aleatório discreto e $\tilde{Y} = g(\tilde{X})$. A *função de probabilidade* de \tilde{Y} é dada por

$$p_{\tilde{Y}}(\tilde{y}_j) = \sum_{i: g(\tilde{x}_i) = \tilde{y}_j} P_{\tilde{X}}(\tilde{x}_i) \quad (2.76)$$

Como \tilde{X} é discreto, \tilde{Y} também o será, de modo que o conjunto dos \tilde{y}_j tais que $g(\tilde{x}_i) = \tilde{y}_j$ é enumerável. Assim, na prática, é possível determinar $p_{\tilde{Y}}$ realizando o mapeamento entre os \tilde{x}_i e \tilde{y}_j . Já o caso onde \tilde{X} é contínuo é de difícil solução, tanto teórica quanto prática. No capítulo 5 estudaremos as abordagens para esta questão.

Em relação à esperança, temos os seguintes resultados, cuja demonstração depende de conceitos de Teoria da Medida, o que foge ao escopo deste trabalho. Se o leitor desejar detalhes sobre a Teoria da Medida, consulte [4].

Proposição 2.40. Seja $\tilde{X} = (X_1, X_2, \dots, X_N)$ um vetor aleatório e g uma função contínua.

(a) Se \tilde{X} é discreto, então

$$E\{g(\tilde{X})\} = \sum_i g(\tilde{x}_i) p_{\tilde{X}}(\tilde{x}_i). \quad (2.77)$$

(b) Se \tilde{X} é contínuo com densidade $f_{\tilde{X}}(x_1, x_2, \dots, x_N)$, então

$$E\{g(\tilde{X})\} = \int \dots \int g(x_1, x_2, \dots, x_N) f_{\tilde{X}}(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N. \quad (2.78)$$

No capítulo 3 e no capítulo 4 veremos diferentes formas de abordar o problema de se determinar a esperança e a variância de um vetor $\tilde{Y} = g(\tilde{X})$.

A proposição abaixo nos diz que se X_1, X_2, \dots, X_N são independentes, então a esperança do produto é igual ao produto das esperanças.

Proposição 2.41. Seja $\tilde{X} = (X_1, X_2, \dots, X_N)$ um vetor aleatório definido num espaço de probabilidade (Ω, \mathbb{A}, P) tal que X_1, X_2, \dots, X_N são independentes e

$$g(\tilde{X}) = g(X_1, X_2, \dots, X_N) = \prod_i^N X_i. \quad (2.79)$$

Então

$$E\{g(\tilde{X})\} = \prod_i^N E\{X_i\}. \quad (2.80)$$

Demonstração: Por indução. No caso $N = 2$, temos $f_{\tilde{X}}(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$, pois X_1 e X_2 são independentes (consequência da definição 2.26). Portanto,

$$\begin{aligned} E\{g(\tilde{X})\} &= E\{g(X_1, X_2)\} \\ &= \int \int g(x_1, x_2) f_{\tilde{X}}(x_1, x_2) dx_1 dx_2 \\ &= \int \int x_1 \cdot x_2 \cdot f_{X_1}(x_1) \cdot f_{X_2}(x_2) dx_1 dx_2 \\ &= \int \left[\int x_1 f_{X_1}(x_1) dx_1 \right] x_2 f_{X_2}(x_2) dx_2 \\ &= \int E\{X_1\} x_2 f_{X_2}(x_2) dx_2 \\ &= E\{X_1\} \cdot E\{X_2\}. \end{aligned} \quad (2.81)$$

Suponha que o resultado é válido para $N - 1$. Então, considerando a variável aleatória $Y_{N-1} = g(X_1, X_2, \dots, X_{N-1})$, temos que

$$\begin{aligned} E\{g(\tilde{X})\} &= E\{g(Y_{N-1}, X_N)\} \\ &= E\{Y_{N-1}\} \cdot E\{X_N\} \\ &= E\{X_1\} \cdot E\{X_2\} \cdots E\{X_N\}, \end{aligned} \quad (2.82)$$

o que completa a indução e a nossa demonstração.

□

Atente para o fato de que a recíproca nem sempre é válida, isto é, se X e Y são variáveis aleatórias tais que $E\{XY\} = E\{X\} \cdot E\{Y\}$, X e Y não são necessariamente independentes. Em geral, $E\{XY\} \neq E\{X\} \cdot E\{Y\}$, e a diferença entre estes valores é denominada *covariância* entre X e Y .

Definição 2.42. Sejam X e Y duas variáveis aleatórias definidas num mesmo espaço de probabilidades. Chamamos *covariância* entre X e Y ao valor

$$\text{Cov}(X, Y) = E\{(X - E\{X\})(Y - E\{Y\})\}. \quad (2.83)$$

Observe que

$$\begin{aligned} \text{Cov}(X, Y) &= E\{(X - E\{X\})(Y - E\{Y\})\} \\ &= E\{(XY - YE\{X\} - XE\{Y\} + E\{X\} \cdot E\{Y\})\} \\ &= E\{XY\} - E\{X\} \cdot E\{Y\}, \end{aligned} \quad (2.84)$$

o que motivou a caracterização inicial da covariância como a diferença entre $E\{XY\}$ e $E\{X\} \cdot E\{Y\}$.

Se $\text{Cov}(X, Y) = 0$, dizemos que X e Y são *não-correlacionadas*. Observe que variáveis independentes são sempre não-correlacionadas, mas variáveis não-correlacionadas não são necessariamente independentes.

A proposição a seguir mostra como computar a variância da soma de variáveis não-correlacionadas.

Proposição 2.43. Seja $\tilde{X} = (X_1, X_2, \dots, X_N)$ um vetor aleatório tal que as variáveis X_i , $i = 1, 2, \dots, N$ são variáveis não-correlacionadas e

$$g(\tilde{X}) = g(X_1, X_2, \dots, X_N) = X_1 + X_2 + \dots + X_N. \quad (2.85)$$

Então

$$\text{Var}\{g(\tilde{X})\} = \sum_{i=1}^N \text{Var}\{X_i\}. \quad (2.86)$$

Demonstração: Sabendo que $E\{X + Y\} = E\{X\} + E\{Y\}$, temos que

$$\begin{aligned} \text{Var}\{\tilde{X}\} &= \text{Var}\{X_1 + X_2 + \dots + X_N\} \\ &= E\{(X_1 + X_2 + \dots + X_N - E\{(X_1 + X_2 + \dots + X_N)\})^2\} \\ &= E\{((X_1 - E\{X_1\}) + (X_2 - E\{X_2\}) + \dots + (X_N - E\{X_N\}))^2\} \\ &= E\left\{\sum_{i=1}^N (X_i - E\{X_i\})^2 + 2 \sum_{i < j} (X_i - E\{X_i\})(X_j - E\{X_j\})\right\} \\ &= \sum_{i=1}^N \text{Var}\{X_i\} + 2 \sum_{i < j} \text{Cov}(X_i, X_j) \\ &= \sum_{i=1}^N \text{Var}\{X_i\}, \end{aligned} \quad (2.87)$$

pois, por hipótese, as variáveis X_i , $i = 1, 2, \dots, N$ são não-correlacionadas.

□

A covariância é, de fato, um exemplo de momento de variáveis aleatórias multidimensionais.

Definição 2.44. Seja $\tilde{X} = (X_1, X_2, \dots, X_N)$ um vetor aleatório. Denominamos *momento de ordem* $n = (n_1, n_2, \dots, n_N)$ de \tilde{X} em torno de $a = (a_1, a_2, \dots, a_N)$ ao valor

$$m_{\tilde{X}}^{(n_1, \dots, n_N)}(\tilde{a}) = E\{(X_1 - a_1)^{n_1} \dots (X_N - a_N)^{n_N}\}. \quad (2.88)$$

Se $m = (m_{X_1}, m_{X_2}, \dots, m_{X_N})$, dizemos que $\mu_{\tilde{X}}^{(n_1, \dots, n_N)} = m_{\tilde{X}}^{(n_1, \dots, n_N)}(m)$ é o *momento central de \tilde{X}* de ordem $n = (n_1, n_2, \dots, n_N)$.

Os demais capítulos deste trabalho tem por objetivo mostrar os métodos para a determinação e/ou aproximação dos valores que caracterizam uma variável aleatória $Y = g(X)$. O capítulo 3 aborda o problema de usar uma aproximação polinomial de g para computar tais valores; no capítulo 4 temos um estudo sobre as formas de se calcular a esperança e a variância de Y , em especial nos casos onde g não pode ser determinada diretamente; por fim, no capítulo 5, discutiremos as formas de se obter a função de distribuição de probabilidade de Y .

3 APROXIMAÇÃO POLINOMIAL DAS FUNÇÕES DE VARIÁVEIS ALEATÓRIAS

Neste capítulo nos estudaremos o problema de se determinar a esperança de uma variável aleatória $Y = g(X)$ substituindo a função g por sua série de Taylor. A vantagem desta abordagem é que, sendo a expansão de Taylor um polinômio, além de termos um método uniforme de realizar tal cálculo, independentemente de qual seja a função g , polinômios são mais fáceis de se implementar computacionalmente do que funções analíticas, o que auxilia o processo de simulação de experimentos em computador.

Em seguida apresentaremos, a título de ilustração, a aplicação desta abordagem no TLM (*Transmission Line Modeling Method*) unidimensional, o qual consiste numa poderosa técnica de simulação eletromagnética, usada na análise de propagação de ondas e antenas (ver [5]). Este caso foi estudado em [6] e [7].

3.1 SÉRIE DE TAYLOR

3.1.1 Definição

Começaremos com o caso unidimensional, antes de generalizarmos para funções no espaço \mathbb{R}^N . Seja $g : \mathbb{R} \rightarrow \mathbb{R}$ uma função contínua. Para que g tenha uma série de Taylor associada, é necessário que g seja *infinitamente diferenciável* num intervalo aberto $I \subset \mathbb{R}$. A definição de funções infinitamente diferenciáveis e da *série de Taylor*, dadas a seguir, foram retiradas de [8].

Definição 3.1. Seja g uma função definida num intervalo aberto $I \subset \mathbb{R}$ e $g^{(n)}$ a sua derivada de n -ésima ordem (para fins de notação, faremos $g^{(0)} = g$). A função g é *infinitamente diferenciável* em I se g tem derivadas $g^{(n)}$ de todas as ordens em I .

Definição 3.2. Seja a função g infinitamente diferenciável em um intervalo aberto $I \subset \mathbb{R}$ e seja a um número em I . Então, a *série de Taylor de g em a* é a série de

potências

$$G(x) = \sum_{k=0}^{\infty} \frac{g^{(k)}(a)}{k!} (x-a)^k. \quad (3.1)$$

Se $a = 0$, a série de Taylor de g em a é denominada *série de Maclaurin* de g em a .

Embora seja uma série infinita, na prática g é expandida apenas até um certo índice N , ignorando-se os termos subsequêntes. Costuma-se, então, denominar os termos expandidos de *aproximação de Taylor de g de ordem N* e os termos restantes de *resto da aproximação de Taylor de g de ordem N* .

Definição 3.3. Seja $g : I \subset \mathbb{R} \rightarrow \mathbb{R}$ uma função infinitamente diferenciável e seja g a série de Taylor de g em a . Chamamos *aproximação de Taylor de g de ordem N* ao polinômio

$$P_N(x) = \sum_{k=0}^N \frac{g^{(k)}(a)}{k!} (x-a)^k \quad (3.2)$$

e a série

$$R_N(x) = \sum_{k=N+1}^{\infty} \frac{g^{(k)}(a)}{k!} (x-a)^k \quad (3.3)$$

de *resto da aproximação de Taylor de g de ordem N* .

Observe que, para todo $N \in \mathbb{N}$,

$$G(x) = P_N(x) + R_N(x). \quad (3.4)$$

O grande interesse de se obter a série de Taylor de uma função g é que, sob determinadas circunstâncias, ela coincide pontualmente com a função, de modo que podemos computar os valores $g(x)$ calculando os valores de $G(x)$, o que é computacionalmente mais fácil devido ao fato de $G(x)$ ser um polinômio. O teorema abaixo discrimina as condições nas quais $G = g$. A demonstração pode ser encontrada em [8], pg. 671.

Definição 3.4. Seja $g : \mathbb{R} \rightarrow \mathbb{R}$ uma função infinitamente diferenciável em $I \subset \mathbb{R}$ e $a \in I$. Se existem $M, R \in \mathbb{R}$, $M, R > 0$ tais que a desigualdade

$$|g^{(N)}(x)| \leq M \quad (3.5)$$

se verifica para todos $x \in (a - R, a + R)$ e todos $N \in \mathbb{N}$, então

$$g(x) = G(x), \quad \forall x \in (a - R, a + R) \quad (3.6)$$

e R é denominado *raio de convergência* da série $G(x)$.

Como dito anteriormente, na prática tomamos $g(x) \approx P_N(x)$ e, ao fazermos isso, cometemos um erro igual a

$$E_N(x) = g(x) - P_N(x) = R_N(x). \quad (3.7)$$

Ainda resta a dificuldade de calcular uma soma infinita. A proposição a seguir fornece uma fórmula para a computação do resto $R_N(x)$, de modo que possamos estimar $E_N(x)$. A demonstração desta proposição pode ser encontrada em [9].

Proposição 3.5. Seja g uma função infinitamente diferenciável em $I \subset \mathbb{R}$, $N \in \mathbb{N}$ e $a \in I$. Então, para cada $x \in I$, $x \neq a$, existe um certo $c \in J$ (onde $J = (a, x)$, se $x > a$, ou $J = (x, a)$, se $x < a$) tal que

$$R_N(x) = \frac{g^{N+1}(c)}{(N+1)!} (x-a)^{N+1}, \quad (3.8)$$

e

$$|E_N(x)| \leq |R_N(x)|, \quad \forall x \in I. \quad (3.9)$$

Para ilustrar a aplicação da série de Taylor, considere a função $g(x) = \cos x$. Como as derivadas n -ésimas de $g(x)$ são ou seno ou cosseno, a menos de sinal, podemos ver que $|g^{(n)}(x)| \leq 1$, $\forall x \in (-\infty, \infty)$, o que nos dá um raio de convergência infinito, isto é, a série de Taylor $G(x)$ de $g(x)$ converge para todos os números reais.

A aproximação de segunda ordem $P_2(x)$ de $g(x)$ em torno de $a = 0$ é igual a

$$\begin{aligned} P_2(x) &= g(a) + g'(a)(x-a) + \frac{g^{(2)}(a)}{2!}(x-a)^2 \\ &= \cos 0 - \sin 0 \cdot (x-0) - \frac{(\cos 0)^2}{2}(x-0)^2 \\ &= 1 - \frac{x^2}{2}. \end{aligned} \quad (3.10)$$

Ao usar esta aproximação, estaremos cometendo um erro menor do que

$$\begin{aligned}
 |E_2(x)| &\leq |R_2(x)| \\
 &\leq \left| \frac{g^{(3)}(c)}{3!} (x-a)^3 \right| \\
 &\leq \frac{|x^3|}{6}
 \end{aligned} \tag{3.11}$$

pois, como dito anteriormente, $a = 0$ e $|g^{(n)}(x)| \leq 1$, $\forall x \in \mathbb{R}$.

Se usássemos $P_2(x)$ para computar o valor de $g(x) = \cos x$ para $x = 1/2$ obteríamos $P_2(1/2) = 0,875$, cometendo um erro menor do que $|E_2(1/2)| \leq 0,020$ (o que já nos garante a correção da primeira casa decimal). Como curiosidade, o valor correto seria $0,877582562\dots$.

Retornando aos conceitos, a definição abaixo nos mostra como obter a série de Taylor no caso multidimensional. Embora a notação seja um pouco mais rebuscada, a idéia permanece a mesma: obter uma série de potências, cujos coeficientes dependem das derivadas parciais das variáveis envolvidas, que aproxime o valor da função em torno de um determinado ponto do espaço.

Proposição 3.6. Seja $g : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ uma função infinitamente diferenciável em $I \subset \mathbb{R}^N$ e $\tilde{a} = (a_1, a_2, \dots, a_N) \in I$. A *série de Taylor de g em \tilde{a}* é dada por

$$G(x_1, \dots, x_N) = \sum_{k_1=0}^{\infty} \cdots \sum_{k_N=0}^{\infty} \frac{\partial^{k_1}}{\partial x_1^{k_1}} \cdots \frac{\partial^{k_N}}{\partial x_N^{k_N}} \frac{g(a_1, \dots, a_N)}{k_1! \dots k_N!} (x_1 - a_1)^{k_1} \dots (x_N - a_N)^{k_N}, \tag{3.12}$$

onde $\frac{\partial^0 g}{\partial x_j^0} = g$, $\forall j = 1, 2, \dots, N$. Para $M \in \mathbb{N}$,

$$P(x_1, \dots, x_N) = \sum_{k_1=0}^M \cdots \sum_{k_N=0}^M \frac{\partial^{k_1}}{\partial x_1^{k_1}} \cdots \frac{\partial^{k_N}}{\partial x_N^{k_N}} \frac{g(a_1, \dots, a_N)}{k_1! \dots k_N!} (x_1 - a_1)^{k_1} \dots (x_N - a_N)^{k_N}. \tag{3.13}$$

3.1.2 Esperança de uma série de Taylor

Nesta seção computaremos a esperança de uma série de Taylor. Para facilitar a notação, considere $\tilde{x} = (x_1, x_2, \dots, x_N)$ e $\tilde{a} = (a_1, a_2, \dots, a_N)$ elementos de \mathbb{R}^N .

Seja $g : \mathbb{R}^N \longrightarrow \mathbb{R}^M$ uma função infinitamente derivável em $I \subset \mathbb{R}^N$ e \tilde{a} um elemento de I . Usando a equação 3.12 e a propriedade E3 da proposição 2.19, a esperança de g

será dada por

$$\begin{aligned}
E\{G(\tilde{x})\} &= E \left\{ \sum_{k_1=0}^{\infty} \cdots \sum_{k_N=0}^{\infty} \frac{\partial^{k_1}}{\partial x_1^{k_1}} \cdots \frac{\partial^{k_N}}{\partial x_1^{k_N}} \frac{g(\tilde{a})}{k_1! \cdots k_N!} (x_1 - a_1)^{k_1} \cdots (x_N - a_N)^{k_N} \right\} \\
&= \sum_{k_1=0}^{\infty} \cdots \sum_{k_N=0}^{\infty} \frac{\partial^{k_1}}{\partial x_1^{k_1}} \cdots \frac{\partial^{k_N}}{\partial x_1^{k_N}} \frac{g(\tilde{a})}{k_1! \cdots k_N!} E\{(x_1 - a_1)^{k_1} \cdots (x_N - a_N)^{k_N}\} \\
&= \sum_{k_1=0}^{\infty} \cdots \sum_{k_N=0}^{\infty} \frac{\partial^{k_1}}{\partial x_1^{k_1}} \cdots \frac{\partial^{k_N}}{\partial x_1^{k_N}} \frac{g(\tilde{a})}{k_1! \cdots k_N!} m_{\tilde{X}}^{(k_1, \dots, k_N)}(\tilde{a}). \tag{3.14}
\end{aligned}$$

Deste modo, podemos entender a esperança de um vetor aleatório $\tilde{Y} = g(\tilde{X})$, gerado a partir de uma transformação g do vetor \tilde{X} , como a combinação linear dos momentos de \tilde{X} , cujos coeficientes dependem das derivadas parciais de todas as ordens de g .

Nem sempre é possível ter todos os momentos necessários ou computar todas as derivadas parciais, mas ainda assim este método é útil em diversas situações, como por exemplo no caso em que as transformações são polinomiais, pois nesta situação a aproximação $P_N(x)$ é exata quando N é maior do que o grau da transformação. Na próxima seção teremos a aplicação desta técnica no problema TLM unidimensional.

3.2 APLICAÇÃO: TLM

Vamos agora aplicar a série de Taylor no problema do TLM unidimensional. Vamos, a título de exemplo, apresentar dois casos: o primeiro, com um único nó; o segundo, com múltiplos nós.

Como visto, para se determinar a esperança a partir da série de Taylor, devemos conhecer os momentos de N -ésima ordem das variáveis aleatórias e também os coeficientes da série. A seguir mostraremos expressões recursivas para o cálculo destes coeficientes.

3.2.1 TLM unidimensional com um único nó

Considere um nó, cujas admitâncias são dadas por λ_1 e λ_2 , rodeado por duas fronteiras com coeficientes Γ_1 e Γ_2 . A figura 3.1 ilustra esta situação.

Este primeiro caso exemplificará o processo de se computar os coeficientes do polinômio de Taylor, então nós consideraremos as seguintes simplificações: sejam Γ_1 e Γ_2 variáveis aleatórias independentes e λ_1 e λ_2 constantes tais que $\lambda_1 = \lambda_2$. Suponha, para fins de

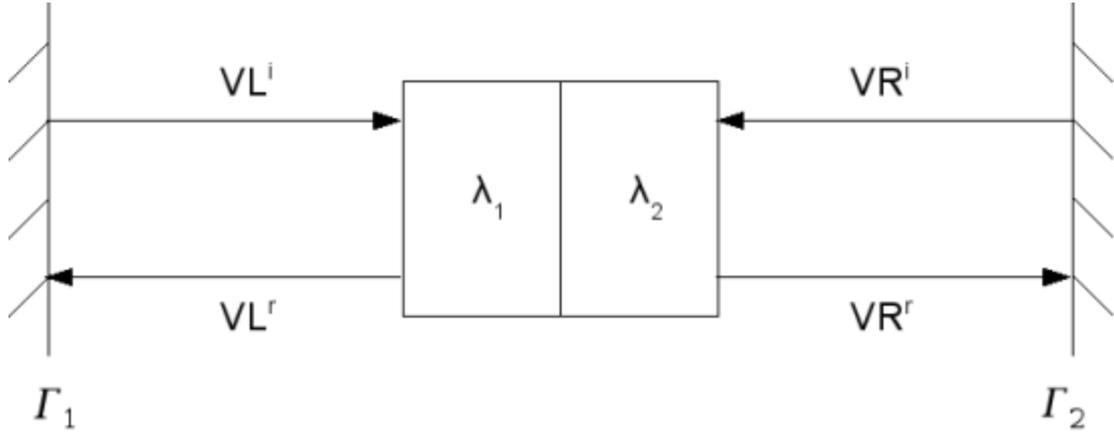


Figura 3.1: TLM com um único nó

aproximação apenas (pois a afirmação nem sempre é verdadeira), que $\mu_{\Gamma_1, \Gamma_2}^{(k_1, k_2)} = 0$ se $k_1 \cdot k_2 \neq 0$. Nestas condições, para cada instante t , temos as quatro variáveis aleatórias abaixo:

$$\begin{cases} VL^i(t) = \Gamma_1 VL^r(t-1) \\ VR^i(t) = \Gamma_2 VR^r(t-1) \\ VL^r(t) = VR^i(t) \\ VR^r(t) = VL^i(t) \end{cases}$$

Os valores iniciais do problema são

$$\begin{cases} VL^i(0) = C = VR^r(0) \\ VR^i(t) = 0 = VL^r(0), \end{cases}$$

onde C é uma constante.

O problema consiste em determinar os valores esperados das variáveis $VL^i(t)$, $VL^r(t)$, $VR^i(t)$ e $VR^r(t)$ para cada instante t . Isto será feito usando a série de Taylor calculada previamente (ver seção 3.1), valendo-se do fato de que os coeficientes da série num dado instante podem ser computados recursivamente a partir dos coeficientes do instante anterior, o que será mostrado a seguir.

Num dado instante t , temos

$$VR^r(t) = VL^i(t) = \Gamma_1 VL^r(t-1). \quad (3.15)$$

Calculando as derivadas parciais em relação a Γ_2 ,

$$\begin{aligned}\frac{\partial VR^r(t)}{\partial \Gamma_2} &= \Gamma_1 \frac{\partial VL^r(t-1)}{\partial \Gamma_2} \\ \frac{\partial^2 VR^r(t)}{\partial \Gamma_2^2} &= \Gamma_1 \frac{\partial^2 VL^r(t-1)}{\partial \Gamma_2^2} \\ &\vdots\end{aligned}$$

Em geral,

$$\frac{\partial^{(n)} VR^r(t)}{\partial \Gamma_2^{(n)}} = \Gamma_1 \frac{\partial^{(n)} VL^r(t-1)}{\partial \Gamma_2^{(n)}}. \quad (3.16)$$

Calculando as derivadas parcial em relação a Γ_1 , temos

$$\begin{aligned}\frac{\partial VR^r(t)}{\partial \Gamma_1} &= VL^r(t-1) + \Gamma_1 \frac{\partial VL^r(t-1)}{\partial \Gamma_1} \\ \frac{\partial^2 VR^r(t)}{\partial \Gamma_1^2} &= \frac{\partial}{\partial \Gamma_1} \left(\frac{\partial VL^r(t-1)}{\partial \Gamma_1} \right) \\ &= \frac{\partial VL^r(t-1)}{\partial \Gamma_1} + \frac{\partial VL^r(t-1)}{\partial \Gamma_1} + \Gamma_1 \frac{\partial^2 VL^r(t-1)}{\partial \Gamma_1^2} \\ &= 2 \cdot \frac{\partial VL^r(t-1)}{\partial \Gamma_1} + \Gamma_1 \frac{\partial^2 VL^r(t-1)}{\partial \Gamma_1^2} \\ \frac{\partial^3 VR^r(t)}{\partial \Gamma_1^3} &= \frac{\partial}{\partial \Gamma_1} \left(\frac{\partial^2 VL^r(t-1)}{\partial \Gamma_1^2} \right) \\ &= 2 \cdot \frac{\partial^2 VL^r(t-1)}{\partial \Gamma_1^2} + \frac{\partial^2 VL^r(t-1)}{\partial \Gamma_1^2} + \Gamma_1 \frac{\partial^3 VL^r(t-1)}{\partial \Gamma_1^3} \\ &= 3 \cdot \frac{\partial^2 VL^r(t-1)}{\partial \Gamma_1^2} + \Gamma_1 \frac{\partial^3 VL^r(t-1)}{\partial \Gamma_1^3} \\ &\vdots\end{aligned}$$

Em geral,

$$\frac{\partial^{(n)} VR^r(t)}{\partial \Gamma_1^{(n)}} = n \cdot \frac{\partial^{(n-1)} VL^r(t-1)}{\partial \Gamma_1^{(n-1)}} + \Gamma_1 \frac{\partial^{(n)} VL^r(t-1)}{\partial \Gamma_1^{(n)}} \quad (3.17)$$

Por outro lado,

$$VL^r(t) = VR^i(t) = \Gamma_2 VR^r(t-1). \quad (3.18)$$

De modo análogo, é fácil ver que

$$\frac{\partial^{(n)} VL^r(t)}{\partial \Gamma_1^{(n)}} = \Gamma_2 \frac{\partial^{(n)} VR^r(t-1)}{\partial \Gamma_1^{(n)}}. \quad (3.19)$$

e

$$\frac{\partial^{(n)} VL^r(t)}{\partial \Gamma_2^{(n)}} = n \cdot \frac{\partial^{(n-1)} VR^r(t-1)}{\partial \Gamma_2^{(n-1)}} + \Gamma_2 \frac{\partial^{(n)} VR^r(t-1)}{\partial \Gamma_2^{(n)}}. \quad (3.20)$$

Com as expressões (3.14), (3.16), (3.17), (3.19) e (3.20), é possível escrever um algoritmo que compute, para cada instante t , a esperança das variáveis aleatórias $VL^i(t)$, $VL^r(t)$, $VR^i(t)$ e $VR^r(t)$.

Algoritmo 3.7. *Cálculo das esperanças para o TLM unidimensional com um nó.*

ENTRADA: As variáveis aleatórias Γ_1 e Γ_2 , valor inicial C e o instante t .

SAÍDA: As esperanças das variáveis aleatórias $VL^i(t)$, $VL^r(t)$, $VR^i(t)$ e $VR^r(t)$.

1. Compute os momentos centrais $\mu_{\Gamma_1}^i$ e $\mu_{\Gamma_2}^i$, $\forall i \in \mathbb{N}$.
2. Inicialize os vetores vr_{r1} , vr_{r2} , vl_{r1} e vl_{r2} .
3. Para $i = 0$, faça
 - (a) Calcule $vr_{r1}[i] = \frac{\partial^{(i)}VR^r(t)}{\partial\Gamma_1^{(i)}}$
 - (b) Se $vr_{r1}[i] = 0$, vá ao passo 4; caso contrário, faça $i = i + 1$ e retorne ao passo 3a.
4. Para $i = 0$, faça
 - (a) Calcule $vr_{r2}[i] = \frac{\partial^{(i)}VR^r(t)}{\partial\Gamma_2^{(i)}}$
 - (b) Se $vr_{r2}[i] = 0$, vá ao passo 5; caso contrário, faça $i = i + 1$ e retorne ao passo 4a.
5. Para $i = 0$, faça
 - (a) Calcule $vl_{r1}[i] = \frac{\partial^{(i)}VL^r(t)}{\partial\Gamma_1^{(i)}}$
 - (b) Se $vl_{r1}[i] = 0$, vá ao passo 6; caso contrário, faça $i = i + 1$ e retorne ao passo 5a.
6. Para $i = 0$, faça
 - (a) Calcule $vl_{r2}[i] = \frac{\partial^{(i)}VL^r(t)}{\partial\Gamma_2^{(i)}}$
 - (b) Se $vl_{r2}[i] = 0$, vá ao passo 7; caso contrário, faça $i = i + 1$ e retorne ao passo 6a.
7. Determine as esperanças $E\{VL^i(t)\}$, $E\{VL^r(t)\}$, $E\{VR^i(t)\}$ e $E\{VR^r(t)\}$ a partir dos vetores vr_{r1} , vr_{r2} , vl_{r1} e vl_{r2} e da expressão (3.14).

3.2.2 TLM unidimensional com múltiplos nós

Nesta seção partiremos para o caso geral do TLM unidimensional. Considere a seqüência de N nós rodeados por duas fronteiras cujos coeficientes são as variáveis aleatórias Γ_1 e Γ_2 . Considere que a posição de cada nó seja representada pela variável inteira x e que cada nó tenha admitâncias ${}_x\lambda_1$ e ${}_x\lambda_2$, as quais também são variáveis aleatórias. A figura 3.2 mostra um exemplo visual destas condições.

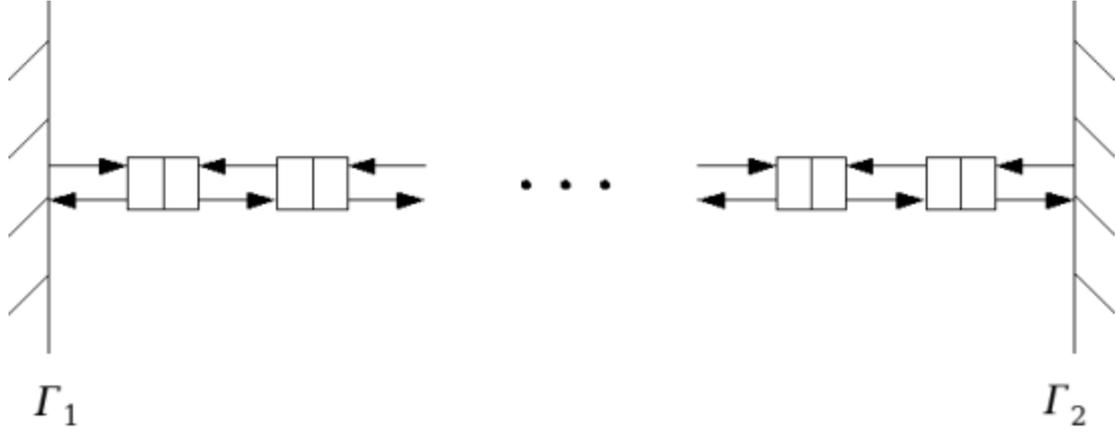


Figura 3.2: TLM com múltiplos nós

Neste caso temos, para cada instante t e para cada nó x ,

$${}_xVL^i(t) = \begin{cases} {}_{x-1}VR^r(t-1), & \text{se } x \neq 1 \\ \Gamma_1 \cdot {}_xVL^r(t-1), & \text{se } x = 1. \end{cases}$$

$${}_xVR^i(t) = \begin{cases} {}_{x-1}VL^r(t-1), & \text{se } x \neq N \\ \Gamma_2 \cdot {}_xVR^r(t-1), & \text{se } x = N \end{cases}$$

$$\begin{cases} {}_xVL^r(t) = {}_xS_{11} \cdot {}_xVL^i(t) + {}_xS_{12} \cdot {}_xVR^i(t) \\ {}_xVR^r(t) = {}_xS_{21} \cdot {}_xVL^i(t) + {}_xS_{22} \cdot {}_xVR^i(t) \end{cases}$$

$$\begin{cases} {}_xS_{11} = \frac{{}_x\lambda_2 - {}_x\lambda_1}{{}_x\lambda_1 + {}_x\lambda_2} \\ {}_xS_{12} = \frac{2 \cdot {}_x\lambda_1}{{}_x\lambda_1 + {}_x\lambda_2} \\ {}_xS_{21} = \frac{2 \cdot {}_x\lambda_2}{{}_x\lambda_1 + {}_x\lambda_2} \\ {}_xS_{22} = \frac{{}_x\lambda_1 - {}_x\lambda_2}{{}_x\lambda_1 + {}_x\lambda_2} \end{cases}$$

O único valor inicial não-nulo do problema é ${}_1VL^i(0) = C$, onde C é uma constante.

Novamente vamos encontrar fórmulas recursivas para os coeficientes da série de Taylor, e usar o mesmo algoritmo do caso de um único nó (algoritmo 3.6). A diferença é que, para este caso, temos 4 variáveis aleatórias para cada nó (${}_xVL^i(t)$, ${}_xVL^r(t)$, ${}_xVR^i(t)$ e ${}_xVR^r(t)$), totalizando $4N$ variáveis aleatórias.

3.2.2.1 Coeficientes das voltagens refletivas

Num dado instante t temos

$${}_xVL^r(t) = s_{11} \cdot {}_xVL^i(t) + s_{12} \cdot {}_xVR^i(t). \quad (3.21)$$

Calculando as derivadas parciais em relação a Γ_1 , temos

$$\begin{aligned} \frac{\partial[{}_xVL^r(t)]}{\partial\Gamma_1} &= s_{11} \cdot \frac{\partial[{}_xVL^i(t)]}{\partial\Gamma_1} + s_{12} \cdot \frac{\partial[{}_xVR^i(t)]}{\partial\Gamma_1} \\ \frac{\partial^2[{}_xVL^r(t)]}{\partial\Gamma_1^2} &= s_{11} \cdot \frac{\partial^2[{}_xVL^i(t)]}{\partial\Gamma_1^2} + s_{12} \cdot \frac{\partial^2[{}_xVR^i(t)]}{\partial\Gamma_1^2} \end{aligned}$$

Em geral,

$$\frac{\partial^{(n)}[{}_xVL^r(t)]}{\partial\Gamma_1^{(n)}} = s_{11} \cdot \frac{\partial^{(n)}[{}_xVL^i(t)]}{\partial\Gamma_1^{(n)}} + s_{12} \cdot \frac{\partial^{(n)}[{}_xVR^i(t)]}{\partial\Gamma_1^{(n)}} \quad (3.22)$$

Calculando as derivadas parciais em relação a Γ_2 , temos

$$\begin{aligned} \frac{\partial[{}_xVL^r(t)]}{\partial\Gamma_2} &= s_{11} \cdot \frac{\partial[{}_xVL^i(t)]}{\partial\Gamma_2} + s_{12} \cdot \frac{\partial[{}_xVR^i(t)]}{\partial\Gamma_2} \\ \frac{\partial^2[{}_xVL^r(t)]}{\partial\Gamma_2^2} &= s_{11} \cdot \frac{\partial^2[{}_xVL^i(t)]}{\partial\Gamma_2^2} + s_{12} \cdot \frac{\partial^2[{}_xVR^i(t)]}{\partial\Gamma_2^2} \end{aligned}$$

Em geral,

$$\frac{\partial^{(n)}[{}_xVL^r(t)]}{\partial\Gamma_2^{(n)}} = s_{11} \cdot \frac{\partial^{(n)}[{}_xVL^i(t)]}{\partial\Gamma_2^{(n)}} + s_{12} \cdot \frac{\partial^{(n)}[{}_xVR^i(t)]}{\partial\Gamma_2^{(n)}} \quad (3.23)$$

Calculando as derivadas parciais em relação a λ_1 , temos para cada x ,

$$\begin{aligned}
\frac{\partial [{}_x VL^r(t)]}{\partial \lambda_1} &= \frac{\partial [{}_x s_{11}]}{\partial \lambda_1} \cdot {}_x VL^i(t) + {}_x s_{11} \cdot \frac{\partial [{}_x VL^i(t)]}{\partial \lambda_1} + \\
&\quad \frac{\partial [{}_x s_{12}]}{\partial \lambda_1} \cdot {}_x VR^i(t) + {}_x s_{12} \cdot \frac{\partial [{}_x VR^i(t)]}{\partial \lambda_1} \\
\frac{\partial^2 [{}_x VL^r(t)]}{\partial \lambda_1^2} &= \frac{\partial^2 [{}_x s_{11}]}{\partial \lambda_1^2} \cdot {}_x VL^i(t) + \frac{\partial [{}_x s_{11}]}{\partial \lambda_1} \cdot \frac{\partial [{}_x VL^i(t)]}{\partial \lambda_1} + \\
&\quad {}_x s_{11} \cdot \frac{\partial^2 [{}_x VL^i(t)]}{\partial \lambda_1^2} + \frac{\partial [{}_x s_{11}]}{\partial \lambda_1} \cdot \frac{\partial [{}_x VL^i(t)]}{\partial \lambda_1} + \\
&\quad \frac{\partial^2 [{}_x s_{12}]}{\partial \lambda_1^2} \cdot {}_x VR^i(t) + \frac{\partial [{}_x s_{12}]}{\partial \lambda_1} \cdot \frac{\partial [{}_x VR^i(t)]}{\partial \lambda_1} + \\
&\quad {}_x s_{12} \cdot \frac{\partial^2 [{}_x VR^i(t)]}{\partial \lambda_1^2} + \frac{\partial [{}_x s_{12}]}{\partial \lambda_1} \cdot \frac{\partial [{}_x VR^i(t)]}{\partial \lambda_1} \\
\frac{\partial^2 [{}_x VL^r(t)]}{\partial \lambda_1^2} &= \frac{\partial^2 [{}_x s_{11}]}{\partial \lambda_1^2} \cdot {}_x VL^i(t) + 2 \cdot \frac{\partial [{}_x s_{11}]}{\partial \lambda_1} \cdot \frac{\partial [{}_x VL^i(t)]}{\partial \lambda_1} + \\
&\quad {}_x s_{11} \cdot \frac{\partial^2 [{}_x VL^i(t)]}{\partial \lambda_1^2} + \frac{\partial^2 [{}_x s_{12}]}{\partial \lambda_1^2} \cdot {}_x VR^i(t) + \\
&\quad 2 \cdot \frac{\partial [{}_x s_{12}]}{\partial \lambda_1} \cdot \frac{\partial [{}_x VR^i(t)]}{\partial \lambda_1} + {}_x s_{12} \cdot \frac{\partial^2 [{}_x VR^i(t)]}{\partial \lambda_1^2} \\
&\quad \vdots
\end{aligned}$$

No caso geral,

$$\frac{\partial_{(n)} [{}_x VL^r(t)]}{\partial \lambda_1^{(n)}} = \sum_{j=0}^n \binom{n}{j} \left[\frac{\partial^{(j)} [{}_x s_{11}]}{\partial \lambda_1^{(j)}} \cdot \frac{\partial^{(n-j)} [{}_x VL^i(t)]}{\partial \lambda_1^{(n-j)}} + \frac{\partial^{(j)} [{}_x s_{12}]}{\partial \lambda_1^{(j)}} \cdot \frac{\partial^{(n-j)} [{}_x VR^i(t)]}{\partial \lambda_1^{(n-j)}} \right]. \quad (3.24)$$

De modo análogo, derivando em relação a λ_2 , temos

$$\frac{\partial_{(n)} [{}_x VL^r(t)]}{\partial \lambda_2^{(n)}} = \sum_{j=0}^n \binom{n}{j} \left[\frac{\partial^{(j)} [{}_x s_{11}]}{\partial \lambda_2^{(j)}} \cdot \frac{\partial^{(n-j)} [{}_x VL^i(t)]}{\partial \lambda_2^{(n-j)}} + \frac{\partial^{(j)} [{}_x s_{12}]}{\partial \lambda_2^{(j)}} \cdot \frac{\partial^{(n-j)} [{}_x VR^i(t)]}{\partial \lambda_2^{(n-j)}} \right]. \quad (3.25)$$

Continuando, num dado instante t temos

$${}_x VR^r(t) = s_{21} \cdot {}_x VL^i(t) + s_{22} \cdot {}_x VR^i(t). \quad (3.26)$$

Os desenvolvimentos das derivadas são análogos aos já realizados para ${}_x VL^r(t)$. Teremos então

$$\frac{\partial^{(n)} [{}_x VR^r(t)]}{\partial \Gamma_1^{(n)}} = s_{21} \cdot \frac{\partial^{(n)} [{}_x VL^i(t)]}{\partial \Gamma_1^{(n)}} + s_{22} \cdot \frac{\partial^{(n)} [{}_x VR^i(t)]}{\partial \Gamma_1^{(n)}} \quad (3.27)$$

$$\frac{\partial^{(n)} [{}_x VR^r(t)]}{\partial \Gamma_2^{(n)}} = s_{21} \cdot \frac{\partial^{(n)} [{}_x VL^i(t)]}{\partial \Gamma_2^{(n)}} + s_{22} \cdot \frac{\partial^{(n)} [{}_x VR^i(t)]}{\partial \Gamma_2^{(n)}} \quad (3.28)$$

$$\frac{\partial_{(n)}[{}_xVR^r(t)]}{\partial\lambda_1^{(n)}} = \sum_{j=0}^n \binom{n}{j} \left[\frac{\partial^{(j)}[{}_xs_{21}]}{\partial\lambda_1^{(j)}} \cdot \frac{\partial^{(n-j)}[{}_xVL^i(t)]}{\partial\lambda_1^{(n-j)}} + \frac{\partial^{(j)}[{}_xs_{22}]}{\partial\lambda_1^{(j)}} \cdot \frac{\partial^{(n-j)}[{}_xVR^i(t)]}{\partial\lambda_1^{(n-j)}} \right]. \quad (3.29)$$

$$\frac{\partial_{(n)}[{}_xVR^r(t)]}{\partial\lambda_2^{(n)}} = \sum_{j=0}^n \binom{n}{j} \left[\frac{\partial^{(j)}[{}_xs_{21}]}{\partial\lambda_2^{(j)}} \cdot \frac{\partial^{(n-j)}[{}_xVL^i(t)]}{\partial\lambda_2^{(n-j)}} + \frac{\partial^{(j)}[{}_xs_{22}]}{\partial\lambda_2^{(j)}} \cdot \frac{\partial^{(n-j)}[{}_xVR^i(t)]}{\partial\lambda_2^{(n-j)}} \right]. \quad (3.30)$$

3.2.2.2 Coeficientes das voltagens incidentes

Para um dado instante t e $x \neq 1$, temos

$${}_xVL^i(t) = {}_{x-1}VR^r(t-1). \quad (3.31)$$

Neste caso, todas as derivadas são computadas diretamente, de modo que

$$\frac{\partial^{(n)}[{}_xVL^i(t)]}{\partial\Gamma_1^{(n)}} = \frac{\partial^{(n)}[{}_{x-1}VR^r(t-1)]}{\partial\Gamma_1^{(n)}}, \quad (3.32)$$

$$\frac{\partial^{(n)}[{}_xVL^i(t)]}{\partial\Gamma_2^{(n)}} = \frac{\partial^{(n)}[{}_{x-1}VR^r(t-1)]}{\partial\Gamma_2^{(n)}}, \quad (3.33)$$

$$\frac{\partial^{(n)}[{}_xVL^i(t)]}{\partial\lambda_1^{(n)}} = \frac{\partial^{(n)}[{}_{x-1}VR^r(t-1)]}{\partial\lambda_1^{(n)}}, \quad (3.34)$$

e

$$\frac{\partial^{(n)}[{}_xVL^i(t)]}{\partial\lambda_2^{(n)}} = \frac{\partial^{(n)}[{}_{x-1}VR^r(t-1)]}{\partial\lambda_2^{(n)}}. \quad (3.35)$$

Quando $x = 1$, temos

$${}_xVL^i(t) = \Gamma_1 \cdot {}_xVL^r(t-1). \quad (3.36)$$

Calculando as derivadas em relação a Γ_2 , temos

$$\begin{aligned} \frac{\partial[{}_xVL^i(t)]}{\partial\Gamma_2} &= \Gamma_1 \cdot \frac{\partial[{}_xVL^r(t-1)]}{\partial\Gamma_2} \\ \frac{\partial^2[{}_xVL^i(t)]}{\partial\Gamma_2^2} &= \Gamma_1 \cdot \frac{\partial^2[{}_xVL^r(t-1)]}{\partial\Gamma_2^2} \\ &\vdots \end{aligned}$$

De modo geral,

$$\frac{\partial^{(n)}[{}_xVL^i(t)]}{\partial\Gamma_2^{(n)}} = \Gamma_1 \cdot \frac{\partial^{(n)}[{}_xVL^r(t-1)]}{\partial\Gamma_2^{(n)}} \quad (3.37)$$

Calculando as derivadas em relação a Γ_1 , temos que

$$\begin{aligned}
\frac{\partial [{}_x VL^i(t)]}{\partial \Gamma_1} &= {}_x VL^r(t-1) + \Gamma_1 \cdot \frac{\partial [{}_x VL^r(t-1)]}{\partial \Gamma_1} \\
\frac{\partial^2 [{}_x VL^i(t)]}{\partial \Gamma_1^2} &= \frac{\partial}{\partial \Gamma_1} \left(\frac{\partial [{}_x VL^i(t)]}{\partial \Gamma_1} \right) \\
&= \frac{\partial [{}_x VL^r(t-1)]}{\partial \Gamma_1} + \frac{\partial [{}_x VL^r(t-1)]}{\partial \Gamma_1} + \Gamma_1 \cdot \frac{\partial^2 [{}_x VL^r(t-1)]}{\partial \Gamma_1^2} \\
&= 2 \cdot \frac{\partial [{}_x VL^r(t-1)]}{\partial \Gamma_1} + \Gamma_1 \cdot \frac{\partial^2 [{}_x VL^r(t-1)]}{\partial \Gamma_1^2} \\
&\vdots
\end{aligned}$$

Em geral,

$$\frac{\partial^{(n)} [{}_x VL^i(t)]}{\partial \Gamma_1^{(n)}} = n \cdot \frac{\partial^{(n-1)} [{}_x VL^r(t-1)]}{\partial \Gamma_1^{(n-1)}} + \Gamma_1 \cdot \frac{\partial^{(n)} [{}_x VL^r(t-1)]}{\partial \Gamma_1^{(n)}}. \quad (3.38)$$

As derivadas em relação a λ_1 e λ_2 são diretas, e correspondem a

$$\frac{\partial^{(n)} [{}_x VL^i(t)]}{\partial \lambda_1^{(n)}} = \Gamma_1 \cdot \frac{\partial^{(n)} [{}_x VL^r(t-1)]}{\partial \lambda_1^{(n)}}. \quad (3.39)$$

e

$$\frac{\partial^{(n)} [{}_x VL^i(t)]}{\partial \lambda_2^{(n)}} = \Gamma_1 \cdot \frac{\partial^{(n)} [{}_x VL^r(t-1)]}{\partial \lambda_2^{(n)}}. \quad (3.40)$$

Por fim, quando $x \neq N$, temos

$${}_x VR^i(t) = {}_{x-1} VL^r(t-1). \quad (3.41)$$

Analogamente aos casos anteriores,

$$\frac{\partial^{(n)} [{}_x VR^i(t)]}{\partial \Gamma_1^{(n)}} = \frac{\partial^{(n)} [{}_{x-1} VL^r(t-1)]}{\partial \Gamma_1^{(n)}}, \quad (3.42)$$

$$\frac{\partial^{(n)} [{}_x VR^i(t)]}{\partial \Gamma_2^{(n)}} = \frac{\partial^{(n)} [{}_{x-1} VL^r(t-1)]}{\partial \Gamma_2^{(n)}}, \quad (3.43)$$

$$\frac{\partial^{(n)} [{}_x VR^i(t)]}{\partial \lambda_1^{(n)}} = \frac{\partial^{(n)} [{}_{x-1} VL^r(t-1)]}{\partial \lambda_1^{(n)}}, \quad (3.44)$$

e

$$\frac{\partial^{(n)} [{}_x VR^i(t)]}{\partial \lambda_2^{(n)}} = \frac{\partial^{(n)} [{}_{x-1} VL^r(t-1)]}{\partial \lambda_2^{(n)}}. \quad (3.45)$$

Quando $x = N$, temos

$${}_xVR^i(t) = \Gamma_2 \cdot {}_xVR^r(t-1). \quad (3.46)$$

Segue, por analogia, que

$$\frac{\partial^{(n)}[{}_xVR^i(t)]}{\partial \Gamma_1^{(n)}} = \Gamma_2 \cdot \frac{\partial^{(n)}[{}_xVR^r(t-1)]}{\partial \Gamma_1^{(n)}} \quad (3.47)$$

$$\frac{\partial^{(n)}[{}_xVR^i(t)]}{\partial \Gamma_2^{(n)}} = n \cdot \frac{\partial^{(n-1)}[{}_xVR^r(t-1)]}{\partial \Gamma_2^{(n-1)}} + \Gamma_2 \cdot \frac{\partial^{(n)}[{}_xVR^r(t-1)]}{\partial \Gamma_2^{(n)}}. \quad (3.48)$$

$$\frac{\partial^{(n)}[{}_xVR^i(t)]}{\partial \lambda_1^{(n)}} = \Gamma_2 \cdot \frac{\partial^{(n)}[{}_xVR^r(t-1)]}{\partial \lambda_1^{(n)}}. \quad (3.49)$$

$$\frac{\partial^{(n)}[{}_xVR^i(t)]}{\partial \lambda_2^{(n)}} = \Gamma_2 \cdot \frac{\partial^{(n)}[{}_xVR^r(t-1)]}{\partial \lambda_2^{(n)}}. \quad (3.50)$$

3.2.3 Resultados Experimentais

Para ilustrar o uso desta abordagem, foi realizado um experimento de TLM unidimensional com as seguintes condições: considere tanto a fronteira esquerda Γ_1 quanto a fronteira direita Γ_2 duas variáveis aleatórias com distribuição normal de probabilidade com média 0,5 e variância 0,0016 e dois nós com admitâncias ${}_1\lambda_1 = 0,3$, ${}_1\lambda_2 = 0,8$, ${}_2\lambda_1 = 0,4$ e ${}_2\lambda_2 = 0,6$ constantes. Foram observados os 10 primeiros instantes t de propagação do impulso inicial.

Tabela 3.1: Tensões incidentes e refletidas sobre o nó 1.

t	Monte Carlo				Taylor			
	${}_1VL^i(t)$	${}_1VR^i(t)$	${}_1VL^r(t)$	${}_1VR^r(t)$	${}_1VL^i(t)$	${}_1VR^i(t)$	${}_1VL^r(t)$	${}_1VR^r(t)$
1	1,0000	0	0,4545	1,4545	1,0000	0	0,4545	1,4545
2	0,2273	0	0,1033	0,3306	0,2273	0	0,1033	0,3306
3	0,0517	0,2909	0,1822	-0,0571	0,0520	0,2909	0,1823	-0,0566
4	0,0911	0,7643	0,4583	-0,2149	0,0914	0,7643	0,4584	-0,2145
5	0,2291	0,0774	0,1464	0,2981	0,2294	0,0771	0,1463	0,2987
6	0,0732	-0,0793	-0,0100	0,1425	0,0736	-0,0789	-0,0096	0,1429
7	-0,0050	-0,0399	-0,0240	0,0109	-0,0045	-0,0396	-0,0236	0,0115
8	-0,0120	0,1815	0,0936	-0,1000	-0,0117	0,1820	0,0939	-0,0997
9	0,0468	0,0553	0,0514	0,0429	0,0470	0,0554	0,0516	0,0431
10	0,0257	-0,0201	0,0007	0,0465	0,0259	-0,0198	0,0010	0,0466

Para fins de comparação, foi realizado um teste de Monte Carlo para 10.000 amostras, usando os mesmos parâmetros. Os resultados obtidos estão organizados nas tabelas 3.1 e 3.2.

Como podemos observar, o método de Taylor traz resultados muito próximos aos obtidos com o Método de Monte Carlo, com a vantagem que não foi necessário a repetição do experimento N vezes. Contudo, embora a hipótese de que $\mu_{\Gamma_1, \Gamma_2}^{(k_1, k_2)} = 0$ para $k_1 \cdot k_2 \neq 0$ levou a pequenas discrepâncias, ela serviu para mostrar que a maior relevância está concentrada nos momentos relacionados as variáveis individuais (isto é, os momentos $\mu_{\Gamma_1, \Gamma_2}^{(k_1, k_2)}$ tais $k_1 \cdot k_2 = 0$), e não nos cruzados (os demais).

Tabela 3.2: Tensões incidentese refletidas sobre o nó 2.

t	Monte Carlo				Taylor			
	${}_2VL^i(t)$	${}_2VR^i(t)$	${}_2VL^r(t)$	${}_2VR^r(t)$	${}_2VL^i(t)$	${}_2VR^i(t)$	${}_2VL^r(t)$	${}_2VR^r(t)$
1	0	0	0	0	0	0	0	0
2	1,4545	0	0,2909	1,7455	1,4545	0	0,2909	1,7455
3	0,3306	0,8727	0,7643	0,2221	0,3306	0,8727	0,7643	0,2221
4	-0,0571	0,1111	0,0774	-0,0907	-0,0566	0,1105	0,0771	-0,0900
5	-0,2149	-0,0454	-0,0793	-0,2488	-0,2145	-0,0450	-0,0789	-0,2484
6	0,2981	-0,1244	-0,0399	0,3826	0,2987	-0,1242	-0,0396	0,3833
7	0,1425	0,1913	0,1815	0,1327	0,1429	0,1917	0,1820	0,1331
8	0,0109	0,0664	0,0553	-0,0002	0,0115	0,0664	0,0554	0,0005
9	-0,1000	-0,0001	-0,0201	-0,1200	-0,0997	0,0002	-0,0198	-0,1197
10	0,0429	-0,0600	-0,0394	0,0635	0,0431	-0,0599	-0,0393	0,0637

A grande desvantagem de usar a aproximação de Taylor para o cálculo da esperança reside no fato que é necessário o conhecimento prévio da transformação g , e que a mesma seja analítica num dado intervalo I . Além disso, o número de momentos a serem conhecidos está diretamente ligado ao grau de precisão da aproximação, e no caso multidimensional, o número de momentos a ser computado cresce exponencialmente.

No próximo capítulo apresentamos um método que supera estas duas dificuldades, onde não é necessário conhecimento da transformação (apenas de seus resultados para determinados pontos de entrada) e a precisão da aproximação será duas vezes maior para o mesmo número de momentos utilizados na aproximação de Taylor.

4 ESTIMANDO A ESPERANÇA E A VARIÂNCIA

O cálculo dos momentos das variáveis aleatórias, e em especial a determinação da esperança e da variância, é de fundamental importância para o estudo do comportamento e das características das mesmas, sobretudo nos casos em que estas variáveis são originadas a partir de transformações não-lineares de outras variáveis aleatórias.

No capítulo anterior, vimos uma maneira de se determinar a esperança, mas para tanto era suposto o conhecimento da transformação a qual as variáveis foram submetidas e os momentos destas variáveis. Contudo, nem sempre é possível conhecer todas estas informações. Neste capítulo propomos um novo método para tais cálculos onde não será imposto o conhecimento prévio da transformação.

Na primeira parte, apresentamos o método; na segunda, mostramos aplicações deste método, fazendo comparativos com o método de Taylor e também com o método de Monte Carlo, que é o único que não exige conhecimento prévio de nenhum parâmetro, exibindo os resultados obtidos com cada método.

4.1 UNSCENTED TRANSFORM GAUSSIANA

4.1.1 Introdução

Dada uma variável aleatória X com função de densidade de probabilidade f_X , a esperança $E\{X\}$ de X é dada, segundo a expressão 2.24, por

$$E\{X\} = \int_{-\infty}^{\infty} x f_X(x) dx. \quad (4.1)$$

Matematicamente falando, a esperança é a soma ponderada de todos os possíveis valores assumidos pela variável X vezes a probabilidade de acontecimento dos mesmos. Desta forma, podemos entender os valores x como os pesos dos valores $f_X(x)$ na referida soma.

A idéia da *Unscented Transform* (UT), técnica proposta inicialmente em [1], é determinar um conjunto finito de pontos, denominados pontos-sigma (*sigma points* no original), que aproximam esta soma quando ponderados pelos pesos devidos.

Interpretando o problema de uma segunda forma, se a esperança $E\{X\}$ é a área A sob a curva $xf_X(x)$, o problema da *Unscented Transform* é determinar um conjunto finito de retângulos R_i , cuja base sejam os pesos e as alturas os valores de f_X avaliada nos pontos-sigma, tais que

$$\bigcup_i R_i \approx A. \quad (4.2)$$

Tanto a proposição original da *Unscented Transform* quanto suas variantes (ver [10] e [11]) apresentam formulações determinísticas para a escolha tanto dos pontos-sigma quanto dos pesos a partir de determinadas condições. Embora estas aproximações tenham mostrados resultados eficientes e úteis, elas podem sofrer distorções no caso geral.

A proposta deste capítulo é apresentar um método para a escolha de tais parâmetros a partir da interpretação do problema da esperança como uma quadratura (ver [12]). Desta maneira, podemos determinar os pontos-sigma e os pesos para quaisquer variáveis aleatórias independentemente de suas distribuições de probabilidade e/ou quaisquer outras condições.

4.1.2 Quadraturas

Um conceito elementar e fundamental no estudo das quadraturas é o da *função peso*, definido abaixo.

Definição 4.1. Seja $w(x)$ uma função definida em um intervalo $I = [a, b]$. Dizemos que $w(x)$ é uma *função peso* em I se $w(x) \geq 0$, $\forall x \in I$ e

$$\int_a^b w(x)dx = 1. \quad (4.3)$$

Dada uma função $f(x)$ definida em I , podemos computar a integral

$$S = \int_a^b w(x)f(x)dx. \quad (4.4)$$

Em geral, um problema de quadratura consiste em se determinar um conjunto de finito de abcissas x_i e de pesos w_i tais que a integral S possa ser aproximada soma das áreas dos retângulos de base w_i e altura $f(x_i)$. Uma definição mais formal é apresentada a seguir.

Definição 4.2. Seja $f(x)$ uma função definida em um intervalo $I = [a, b]$ e $w(x)$ uma função peso em I . Definimos como *quadratura* o problema de se determinar um conjunto de m abscissas x_i (denominadas *pontos-sigma*) e de m pesos w_i tais que

$$\int_a^b w(x)f(x)dx \approx \sum_{i=1}^m w_i f(x_i) \quad (4.5)$$

Por constituir uma aproximação, a quadratura tem um erro associado, dado pela diferença entre o real valor da integral e o valor da aproximação.

Definição 4.3. Chamamos de *erro* E da quadratura ao valor

$$E = \int_a^b w(x)f(x)dx - \sum_{i=1}^m w_i f(x_i). \quad (4.6)$$

Além do erro, a qualidade da aproximação de uma quadratura pode ser mensurado através de um coeficiente de *precisão*.

Definição 4.4. Dizemos que uma quadratura tem *precisão de, no mínimo, N* se temos $E = 0$ para, no mínimo, $f(x) = x^r$, com $r = 0, 1, \dots, N$.

A precisão está relacionada diretamente com a série de Taylor de $f(x)$, pois a definição acima nos diz que a quadratura é *exata* (isto é, tem erro zero) para as aproximação de Taylor $P_i(x)$ de $f(x)$ (ver definição 3.3), para $i = 0, 1, \dots, N$.

A proposição abaixo apresenta condições suficientes para que uma quadratura tenha precisão de, no mínimo, N .

Proposição 4.5. Se definimos o r -ésimo momento M_r de $w(x)$ em I como

$$M_r = \int_a^b x^r w(x)dx, \quad (4.7)$$

então a quadratura

$$\int_a^b w(x)f(x)dx \approx \sum_{i=1}^m w_i f(x_i) \quad (4.8)$$

tem grau de precisão de, no mínimo, N se valem as $N + 1$ equações

$$\sum_{i=1}^m w_k x_k^r = M_r, \quad (4.9)$$

com $r = 0, 1, \dots, N$.

Demonstração: Tome um r inteiro tal que $r \in [0, N]$ e faça $f(x) = x^r$. De acordo com a definição 4.3, e usando a hipóteses dadas pela equação 4.9, temos

$$\begin{aligned} E &= \int_a^b w(x)f(x)dx - \sum_{i=1}^m w_i f(x_i) \\ &= \int_a^b w(x)x^r dx - \sum_{i=1}^m w_i x_i^r \\ &= M_r - \sum_{i=1}^m w_i x_i^r \\ &= 0. \end{aligned} \quad (4.10)$$

Como a conclusão acima vale para todo $r = 0, 1, \dots, N$, logo $f(x)$ tem precisão de, no mínimo, N .

□

Outro conceito importante no estudo de quadraturas é o de *polinômios ortogonais*, que por sua vez depende da definição de um *produto interno*, dada abaixo.

Definição 4.6. Seja um intervalo $I = [a, b] \in \mathbb{R}$ e uma função de peso $w(x)$ em I . Definimos, para quaisquer polinômios f e g , o *produto interno em I de f por g no espaço vetorial dos polinômios* como

$$\langle f, g \rangle = \int_a^b f(x)g(x)w(x)dx. \quad (4.11)$$

Definição 4.7. Considere uma sequência de polinômios

$$\{p_n\} = p_0, p_1, p_2, \dots \quad (4.12)$$

onde p_i tem grau i . Dizemos que $\{p_n\}$ é uma *seqüência de polinômios ortogonais* se, para $i \neq j$, $\langle p_i, p_j \rangle = 0$.

A característica mais relevante de uma sequência de polinômios ortogonais é que ela constitui uma base para o espaço vetorial dos polinômios, isto é, qualquer polinômio $g(x)$ de grau n pode ser escrito como uma combinação linear dos $n + 1$ primeiros termos da sequência, conforme a proposição abaixo.

Proposição 4.8. Seja $\{p_n\}$ uma sequência de polinômios ortogonais em I . Então, para qualquer polinômio $g(x)$ de grau n , existem $n + 1$ coeficientes c_i tais que

$$g(x) = \sum_{i=0}^n c_i p_i(x). \quad (4.13)$$

Demonstração: A demonstração é feita por indução matemática. Se o grau de $g(x)$ é igual a 1, então $g(x) = c$, onde $c \in \mathbb{R}$. Deste modo

$$g(x) = c = \left(\frac{c}{p_0} \right) \cdot p_0 = c_0 p_0. \quad (4.14)$$

Suponha agora que a hipótese é verdadeira para todos os polinômios de grau $n - 1$. Se $g(x) = a_n x^n + \dots + a_0$ tem grau n e $p_n = b_n x^n + \dots + b_0$, então para $c_n = a_n/b_n$,

$$\bar{g}(x) = g(x) - c_n p_n \quad (4.15)$$

tem grau $n - 1$. Daí, por hipótese, existem n coeficientes c_i tais que

$$\bar{g}(x) = \sum_{i=0}^{n-1} c_i p_i(x). \quad (4.16)$$

Como $g(x) = c_n p_n + \bar{g}(x)$, então

$$g(x) = \sum_{i=0}^n c_i p_i(x), \quad (4.17)$$

completando assim a nossa indução.

□

Outros importantes fatos a respeito dos polinômios ortogonais são dados na próxima proposição. A demonstração dos mesmos foge ao escopo deste trabalho. Os interessados podem consultar [13] ou [14] para maiores detalhes.

Proposição 4.9. Seja $\{p_n\}$ uma sequência de polinômios ortogonais em I . Valem as seguintes afirmações:

1. Cada polinômio p_i é ortogonal a qualquer polinômio $g(x)$ de grau menor que i .
2. Considere $p_i = k_i x^i + k'_i x^{i-1} + \dots \in \{p_n\}$. A fórmula de recorrência dos termos da sequência é dada por

$$p_{n+1} = (a_n x + b_n)p_n - c_n p_{n-1}, \quad \forall n > 0, \quad (4.18)$$

onde

$$a_n = \frac{k_{n+1}}{k_n}, \quad b_n = a_n \left(\frac{k'_{n+1}}{k_{n+1}} - \frac{k'_n}{k_n} \right), \quad c_n = a_n \left(\frac{k_{n-1} h_n}{k_n h_{n-1}} \right) \quad (4.19)$$

e $h_i = \langle p_i, p_i \rangle$.

3. Cada polinômio p_i tem todas suas i raízes reais, distintas e estritamente dentro do intervalo I .
4. As raízes de cada polinômio p_i estão estritamente entre as raízes do polinômio p_{i+1} .

Estes fatos serão úteis na análise e no estudo da solução da quadratura que será proposta na próxima seção.

4.1.3 A quadratura gaussiana generalizada

Fazendo o paralelo com as quadraturas gaussianas (ver [12]), podemos trabalhar sem restrições na expressão da quadratura (tanto no intervalo quanto na função $f(x)$, mantendo apenas a condição da função peso $w(x)$) e determinar como meta de precisão de, no mínimo, $2m - 1$. Denominaremos, para este trabalho, estas condições como *quadratura gaussiana generalizada*.

Definição 4.10. Seja $f(x)$ uma função definida em um intervalo $I = [a, b]$ e $w(x)$ uma função peso em I . Definimos como *quadratura gaussiana generalizada em I com precisão de, no mínimo, $2m - 1$* o problema de se determinar um conjunto de m pontos-sigma x_i e de m pesos w_i tais que

$$\int_a^b w(x) f(x) dx \approx \sum_{i=1}^m w_i f(x_i) \quad (4.20)$$

O termo *generalizada* se deve ao fato de que, ao contrário da quadratura gaussiana que é definida no intervalo $(-\infty, \infty)$, não impusemos nenhuma restrição ao intervalo I .

Para mostrar a solução da quadratura gaussiana generalizada, começaremos com o caso mais simples, com $m = 2$. A solução geral é idêntica e será descrita logo após.

De acordo com as condições apresentadas na seção anterior, para obtermos precisão igual a $2m - 1$ precisamos solucionar o seguinte sistema de equações, dado pela equação 4.9:

$$\begin{aligned}w_1 + w_2 &= M_0 \\w_1x_1 + w_2x_2 &= M_1 \\w_1x_1^2 + w_2x_2^2 &= M_2 \\w_1x_1^3 + w_2x_2^3 &= M_3\end{aligned}\tag{4.21}$$

Deste modo, temos quatro equações com quatro variáveis, a saber x_1, x_2, w_1 e w_2 . Observe que, embora o sistema seja linear em w_1 e w_2 , ele não o é em x_1 e x_2 .

Para resolver o sistema 4.21, considere x_1 e x_2 como sendo os zeros do polinômio $\pi_2(x)$ dado por

$$\pi_2(x) = (x - x_1)(x - x_2) = x^2 + \alpha_1x + \alpha_2.\tag{4.22}$$

Se obtivermos os coeficientes α_1 e α_2 , poderemos encontrar os valores de x_1 e x_2 desejados. Multiplicando a primeira equação por α_2 , a segunda por α_1 e a terceira por 1, teremos

$$\begin{aligned}\alpha_2w_1 + \alpha_2w_2 &= \alpha_2M_0 \\ \alpha_1w_1x_1 + \alpha_1w_2x_2 &= \alpha_1M_1 \\ w_1x_1^2 + w_2x_2^2 &= M_2\end{aligned}\tag{4.23}$$

Somando as equações, teremos

$$(x_1^2 + \alpha_1x_1 + \alpha_2)w_1 + (x_2^2 + \alpha_1x_2 + \alpha_2)w_2 = M_2 + \alpha_1M_1 + \alpha_2M_0,\tag{4.24}$$

isto é,

$$\pi_2(x_1)w_1 + \pi_2(x_2)w_2 = M_2 + \alpha_1M_1 + \alpha_2M_0.\tag{4.25}$$

Como x_1 e x_2 são zeros de $\pi_2(x)$, a equação acima se resume a

$$M_2 + \alpha_1 M_1 + \alpha_2 M_0 = 0. \quad (4.26)$$

Fazendo o mesmo processo, multiplicando a segunda equação por α_2 , a terceira por α_1 e a quarta por 1, teremos

$$\pi_2(x_1)x_1w_1 + \pi_2(x_2)x_2w_2 = M_3 + \alpha_1 M_2 + \alpha_2 M_1, \quad (4.27)$$

o que nos leva a expressão

$$M_3 + \alpha_1 M_2 + \alpha_2 M_1 = 0. \quad (4.28)$$

Procedendo desta maneira, encontramos o sistema

$$\begin{aligned} M_2 + \alpha_1 M_1 + \alpha_2 M_0 &= 0 \\ M_3 + \alpha_1 M_2 + \alpha_2 M_1 &= 0, \end{aligned} \quad (4.29)$$

o qual é linear em α_1 e α_2 . O sistema terá solução única se $M_1^2 \neq M_2 M_0$, o que é garantido desde que $w(x) \geq 0$ no intervalo I (que equivale a restrição determinada para a função peso). Logo teremos

$$\alpha_1 = \frac{M_3 M_0 - M_2 M_1}{M_1^2 - M_2 M_0} \quad (4.30)$$

e

$$\alpha_2 = \frac{M_2^2 - M_3 M_1}{M_1^2 - M_2 M_0}. \quad (4.31)$$

Como x_1 e x_2 são zeros de $\pi_2(x)$, temos também que

$$x_1 = \frac{-\alpha_1 + \sqrt{\alpha_1^2 - 4\alpha_2}}{2} \quad (4.32)$$

e

$$x_2 = \frac{-\alpha_1 - \sqrt{\alpha_1^2 - 4\alpha_2}}{2}, \quad (4.33)$$

as quais são as abscissas que deveríamos determinar. Para encontrar os pesos, basta solucionar o sistema

$$\begin{aligned} w_1 + w_2 &= M_0 \\ w_1 x_1 + w_2 x_2 &= M_1, \end{aligned} \quad (4.34)$$

o que nos dá

$$w_1 = \frac{M_1 - x_2 M_0}{x_1 - x_2} \quad (4.35)$$

e

$$w_2 = \frac{M_1 - x_1 M_0}{x_2 - x_1}. \quad (4.36)$$

O algoritmo abaixo mostra a solução para o caso geral.

Algoritmo 4.11. *Solução da Quadratura Gaussiana Generalizada.*

ENTRADA: O intervalo I , a função peso $w(x)$ e a precisão $2m - 1$.

SAÍDA: Os m pesos w_i e os m pontos-sigma x_i .

1. Resolva o sistema linear de m equações dadas por

$$\sum_{i=0}^m \alpha_i M_{N+m-i} = 0, \quad (4.37)$$

onde $N = 0, 1, \dots, m - 1$, $\alpha_0 = 1$ e

$$M_r = \int_a^b x^r w(x) dx. \quad (4.38)$$

2. Encontrados os valores α_i , determine as raízes x_1, x_2, \dots, x_m do polinômio

$$\pi_m(x) = \sum_{i=0}^m \alpha_i x^{m-i}. \quad (4.39)$$

3. Determine os pesos w_1, w_2, \dots, w_m através do sistema linear de m equações dadas pela expressão

$$\sum_{i=1}^m w_i x_i^r = M_r, \quad (4.40)$$

com $r = 0, 1, \dots, m - 1$.

Na solução de uma quadratura gaussiana generalizada com precisão de, no mínimo, $2m - 1$ em I , o algoritmo 4.11 determina os pontos-sigma x_i através do cálculo das raízes do polinômio $\pi_m(x)$, definido na equação 4.39. Este polinômio tem um significado e uma importância maior, pois a sequência de polinômios $\pi_0(x), \pi_1(x), \dots, \pi_m(x)$ é uma sequência de polinômios ortogonais em I .

Proposição 4.12. Dada uma quadratura gaussiana generalizada com precisão de, no mínimo, $2m - 1$ no intervalo I , a sequência de polinômios $\pi_0(x), \pi_1(x), \dots, \pi_m(x)$ é ortogonal em I .

Demonstração: Provaremos, de fato, que $\pi_m(x)$ é ortogonal a qualquer polinômio de grau menor do que m (veja o item 1 da proposição 4.9), e, por indução, que a sequência $\pi_0(x), \pi_1(x), \dots, \pi_m(x)$ é ortogonal em I .

Considere $g(x)$ um polinômio de grau $m - 1$ e $h(x) = g(x)\pi_m(x)$. Deste modo, $h(x)$ tem grau $2m - 1$ e $h(x_i) = 0$ para cada x_i tal que $\pi_m(x_i) = 0$. Como temos, por hipótese, precisão de no mínimo $2m - 1$ (ver definição 4.4), segue que

$$\begin{aligned} \int_a^b w(x)g(x)\pi_m(x)dx &= \int_a^b w(x)h(x)dx \\ &= \sum_{i=0}^m w_i h(x_i) \\ &= 0, \end{aligned} \tag{4.41}$$

de modo que $\pi_m(x)$ é ortogonal (ver 4.7) a todos polinômios de grau inferior a n . Em particular, $\pi_m(x)$ é ortogonal a $\pi_{m-1}(x), \pi_{m-2}(x), \dots, \pi_0(x)$. Se repetirmos o raciocínio para π_{m-1} , e depois para π_{m-2} e assim sucessivamente, teremos a sequência de polinômios ortogonais $\pi_0(x), \pi_1(x), \dots, \pi_m(x)$

□

A proposição 4.12, quando combinada com as afirmações da proposição 4.9, tem importantes consequências:

- O primeiro passo do algoritmo 4.11 tem como objetivo determinar os coeficientes do polinômio $\pi_m(x)$ e consiste na solução de um sistema linear de m equações. Como o sistema cresce de acordo com o grau de precisão, este passo pode ser substituído, após o cálculo dos casos $m = 0, 1, 2$, pelo algoritmo de recorrência apresentado no item 2 da proposição 4.9, quando este for computacionalmente mais viável.
- O segundo passo do algoritmo 4.11 consiste no cálculo das raízes do polinômio $\pi_m(x)$. O item 3 da proposição 4.9 garante a existência destas raízes e que as

mesmas são reais; já o item 4 delimita os intervalos onde estas raízes podem ser encontradas, o que é uma importante etapa nos algoritmos utilizados na determinação de zeros de polinômios (ver [15], capítulo 2).

Nas condições da quadratura gaussiana generalizada em I com precisão de, no mínimo, $2m - 1$, o erro E pode ser aproximado, segundo [12], pág. 380, através da expressão

$$E = \frac{f^{(2m)}(\xi)}{(2m)!} \int_a^b w(x) [\pi_m(x)]^2 dx, \quad (4.42)$$

onde $\xi \in (a, b)$.

Observe que, naturalmente, o erro depende do intervalo $I = [a, b]$, da $2m$ -ésima derivada da função $f(x)$, da função peso $w(x)$ e do polinômio $\pi(x)$, definido na equação 4.39, cujas raízes são os pontos-sigma x_i .

4.1.4 Unscented Transform Gaussiana

A *Unscented Transform Gaussiana* consiste na adaptação da quadratura gaussiana generalizada para o caso de funções de variáveis aleatórias.

4.1.4.1 Caso unidimensional

Considere uma variável aleatória X com função de densidade de probabilidade $p_X(x)$ e uma transformação $g(x)$ tal que $Y = g(X)$. Sabemos, segundo o teorema 2.28, que

$$E\{Y\} = \int_{-\infty}^{\infty} p_X(x)g(x)dx. \quad (4.43)$$

Como $p_X(x)$ é uma função de densidade de probabilidade, ela atende os requisitos de uma função peso, isto é, $p_X(x) \geq 0, \forall x$ e

$$\int_{-\infty}^{\infty} p_X(x)dx = 1. \quad (4.44)$$

Desta maneira, para se determinar a esperança da variável Y com um grau de precisão de, no mínimo, $2m - 1$, devemos encontrar as m abcissas x_1, x_2, \dots, x_m e os m pesos w_1, w_2, \dots, w_m tais que

$$E\{Y\} = \int_{-\infty}^{\infty} p_X(x)g(x)dx \approx \sum_{i=0}^m w_i g(x_i), \quad (4.45)$$

ou seja, o problema da esperança de Y é uma quadratura gaussiana generalizada.

Observe também que, de acordo com a definição 2.23 e a proposição 4.5, que $M_r = m_X^r$. Assim, a função geradora de momentos (ver definição 2.29) pode ser usada no cálculo dos coeficientes M_r necessários no passo 1 do algoritmo 4.11.

Podemos também computar a variância de Y , usando a relação

$$\text{Var}\{Y\} = E\{(Y - E\{Y\})^2\}. \quad (4.46)$$

Uma vez que o cálculo dos parâmetros x_i e w_i independe da função que transforma a variável aleatória, levando em consideração apenas a função peso em I , se defirmos $E\{Y\} = \bar{y}$ e $h(x) = (g(x) - \bar{y})^2$, temos

$$\begin{aligned} \text{Var}\{Y\} &= E\{(Y - E\{Y\})^2\} \\ &= E\{h(x)\} \\ &= \int_{-\infty}^{\infty} p_X(x)h(x)dx \\ &\approx \sum_{i=0}^m w_i h(x_i) \\ &\approx \sum_{i=0}^m w_i (g(x_i) - \bar{y})^2 \end{aligned} \quad (4.47)$$

de modo que uma vez computadas as abcissas e os pesos, podemos também determinar a variância. Como o cálculo da variância utiliza uma função cujo grau é o dobro do grau da função original, a precisão da aproximação diminui pela metade, isto é, se para o cálculo da esperança temos precisão de, no mínimo, $2m - 1$, para variância teremos precisão de, no mínimo, $m - 1$.

De fato, podemos computar o k -ésimo momento central da variável Y através da expressão

$$M_k \approx \sum_{i=0}^m w_i (y_i - \bar{y})^k, \quad (4.48)$$

cuja precisão será de, no mínimo, $[(2m - 1)/k]$, onde $[z]$ é o maior número inteiro menor ou igual a z .

4.1.4.2 Caso multidimensional

Nesta seção apresentamos o caso bidimensional. A abordagem para as demais dimensões é análoga e será deixada a cargo do leitor.

Considere duas variáveis aleatórias independentes X e Y e uma transformação g tal que $Z = g(X, Y)$. Neste caso, usando a *unscented transform gaussiana*, teremos:

$$\begin{aligned}
 E\{Z\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{X,Y}(x, y)g(x, y)dx dy \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_X(x)p_Y(y)g(x, y)dx dy \\
 &= \int_{-\infty}^{\infty} p_Y(y) \left(\int_{-\infty}^{\infty} p_X(x)g(x, y)dx \right) dy \\
 &\approx \int_{-\infty}^{\infty} p_Y(y) \left(\sum_{i=0}^m w_i^X g(x_i, y) \right) dy \\
 &\approx \sum_{i=0}^m w_i^X \left(\int_{-\infty}^{\infty} p_Y(y)g(x_i, y) dy \right) \\
 &\approx \sum_{i=0}^m w_i^X \left(\sum_{j=0}^m w_j^Y g(x_i, y_j) \right) \\
 &\approx \sum_{j=0}^m \sum_{i=0}^m w_i^X w_j^Y g(x_i, y_j)
 \end{aligned} \tag{4.49}$$

Assim como no caso unidimensional, podemos calcular os outros momentos centrais de Z usando raciocínio análogo.

Finalmente, para o caso mais geral, onde X e Y não são independentes, defina o momento

$$M_{i,j} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^i y^j p_{X,Y}(x, y)dx dy. \tag{4.50}$$

Para termos precisão $2m - 1$, precisamos resolver o sistema de $4m^2$ equações dadas por

$$\sum_{j=1}^m \sum_{i=1}^m w_{i,j} x_i^k y_j^t = M_{k,t}, \tag{4.51}$$

com $k, t = 0, 1, \dots, 2m - 1$. Definindo o polinômio $\pi_m(x, y)$ como

$$\pi_m(x, y) = \left(\prod_{j=1}^m (y - y_j) \right) \left(\prod_{i=1}^m (x - x_i) \right) \quad (4.52)$$

e usando uma estratégia semelhante a do caso unidimensional, teremos um sistema linear de m^2 equações da forma

$$\sum_{j=1}^m \sum_{i=1}^m \alpha_{i,j} M_{k+m-i, t+m-j} = 0, \quad (4.53)$$

onde $k, t = 0, 1, \dots, m - 1$ e $\alpha_{i,j}$ são os coeficientes do termo $x^i y^j$ do polinômio $\pi_m(x, y)$. Conhecidos os coeficientes de $\pi_m(x, y)$, pode-se determinar os pares (x_i, y_j) através do cálculo dos seus zeros e, a partir deles, computar os valores dos pesos $w_{i,j}$.

Como pode se observar, o aumento da dimensão e/ou do grau de precisão acarreta num aumento da complexidade das operações, sobretudo nos cálculos dos momentos e dos zeros do polinômio associado.

4.2 APLICAÇÃO: LEI DE COULOMB

A **Lei de Coulomb**, que fora deduzida por Charles Augustus Coulomb em 1785, diz respeito às forças exercidas sobre duas cargas puntiformes separadas por uma determinada distância. Se ambas cargas são positivas ou ambas são negativas, a força é de **repulsão**, de modo que as cargas tendem a se afastar (ver Figura 4.1); se uma é negativa e a outra é positiva, ou vice-versa, a força é de **atração**, isto é, as cargas tendem a se aproximar (ver Figura 4.2).

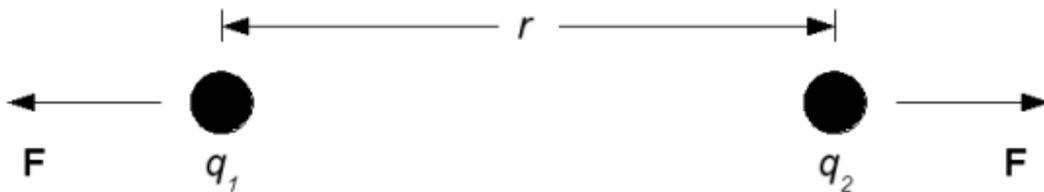


Figura 4.1: Força de repulsão

Segundo Coulomb, o módulo da força exercida em cada carga é dada pela expressão

$$F = k \frac{q_1 q_2}{r^2}, \quad (4.54)$$

onde q_1 e q_2 são os módulos das cargas (em *coulombs*), r é a distância entre as cargas (em metros) e k é a *constante eletrostática*, dada por

$$k = 8,99 \times 10^9 N \cdot m^2 / C^2. \quad (4.55)$$

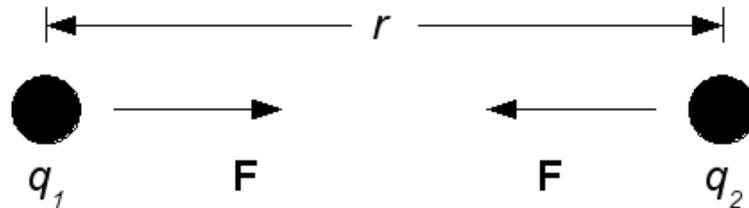


Figura 4.2: Força de atração

Para o exercício desta equação, [16] propõe, na página 8, o seguinte exercício: “A distância média r entre o elétron e o próton no núcleo do átomo de hidrogênio é de $5,3 \times 10^{-11}$ m. Qual é o módulo da força eletrostática média que atua entre essas duas partículas?”

Acontece que a solução proposta pelo livro, que consiste na aplicação direta da Lei de Coulomb, é imprecisa. Exporemos primeiramente este fato, e em seguida apresentaremos soluções alternativas para este problema, baseadas nos conceitos e técnicas desenvolvidas ao longo deste trabalho.

Adotaremos, para a *carga elementar* e , o valor

$$e = 1,6 \times 10^{-19} C. \quad (4.56)$$

4.2.1 Esperança da Lei de Coulomb

A solução proposta pelo livro, reproduzida aqui, é a seguinte:

$$\begin{aligned} F &= k \frac{q^2}{r^2} \\ &= \frac{(8,99 \times 10^9 N \cdot m^2 / C^2) \cdot (1,6 \times 10^{-19} C)^2}{5,3 \times 10^{-11} m} \\ &= 8,2 \times 10^{-8} N. \end{aligned} \quad (4.57)$$

No problemas não temos a distância exata entre as partículas, e sim a distância **média**, de forma que a força a ser calculada também não é a exata, e sim a média. Em termos

de probabilidade, o problema poderia ser reformulado da seguinte forma: *qual é a esperança da força, conhecida a esperança da distância?*

De fato, esta abordagem é a que acontece na prática, pois devido as imprecisões dos aparelhos de medição, a distância não pode ser computada diretamente, mas sim medida com uma certa precisão e uma determinada margem de erro. A Lei de Coulomb é verdadeira no caso ideal, onde todas as medições são feitas com perfeição; havendo erros e/ou aproximações, devemos visualizar a questão sob a ótica da probabilidade.

Considere então que a distância r é medida como $5,3 \times 10^{-11}$ m, podendo este valor variar igualmente para baixo ou para cima em até δ metros. Em outras palavras, considere $r \sim U(a, b)$, onde $a = r - \delta$ e $b = r + \delta$. Segundo a proposição 2.40(b), temos

$$\begin{aligned}
 E\{F\} &= \int_a^b \left(\frac{1}{b-a} \right) \cdot \left(k \frac{q^2}{r^2} \right) dr \\
 &= \frac{kq^2}{b-a} \int_a^b \frac{1}{r^2} dr \\
 &= \frac{kq^2}{b-a} \cdot \left(-\frac{1}{r} \Big|_a^b \right) \\
 &= \frac{kq^2}{b-a} \cdot \left(\frac{1}{a} - \frac{1}{b} \right) \\
 &= \frac{kq^2}{b-a} \cdot \left(\frac{b-a}{ab} \right) \\
 &= \frac{kq^2}{ab}
 \end{aligned} \tag{4.58}$$

Considerando, a título de exemplo, um erro de medição $\delta = 0,05 \times 10^{-11}$, a força média seria, segundo a equação 4.58, igual a

$$\bar{F} = 8,193822874944370 \times 10^{-8}, \tag{4.59}$$

que corresponde ao resultado correto do problema.

Usando a Lei de Coulomb diretamente para obtermos a força média, teríamos

$$F = \mathbf{8,193093627625485} \times 10^{-8}, \tag{4.60}$$

de modo que, em relação ao valor dado em 4.59, estaríamos corretos apenas nas 3 primeiras casas.

A solução obtida em 4.59 é a melhor solução para o problema, mas ela depende da resolução analítica a esperança de F . Em outros casos, nem sempre é possível computar a esperança diretamente, seja pela dificuldade da integração, seja pelo fato de não termos a expressão analítica da transformação da variável aleatória. Nas próximas subseções resolvemos este problema novamente usando outras técnicas, mostrando as vantagens e desvantagens de cada uma, e comparando os resultados obtidos com o resultado ideal.

4.2.2 Lei de Coulomb com Monte Carlo

O método de **Monte Carlo** consiste na repetição de N iterações de um experimento, no registro dos resultados obtidos em cada passo e, ao final do processo, no cálculo da média e de outros coeficientes associados ao problema.

A principal vantagem do método de Monte Carlo é que ele não impõe nenhuma restrição ao problema e nem exige conhecimento prévio de nenhum parâmetro. Podemos, através dele, calcular a média de uma variável aleatória resultante da transformação de uma variável aleatória (a qual não precisa ser inteiramente determinada) por um processo, que pode ou não ser conhecido. É necessária apenas a repetição do experimento e o registro dos resultados obtidos.

Porém as inúmeras repetições constituem a maior fraqueza do método. Isto porque estas repetições tendem a ser muito custosas, seja em recursos materiais e financeiros, seja em tempo de realização. Em geral usa-se a simulação computacional para reduzir os custos materiais e financeiros, mas ainda assim há o custo de processamento e de tempo.

Outra importante desvantagem do método é a lentidão da convergência dos resultados. Em alguns casos é necessário um número elevado de amostras para a obtenção de um resultado com precisão aceitável, e a própria mensuração da precisão do resultado consiste um problema à parte.

Dadas estas características, o método de Monte Carlo é, em geral, utilizado ou como guia e validação para as outras técnicas ou como alternativa única, quando a situação não permitir a utilização de métodos alternativos.

Voltando ao nosso problema, realizamos 1.000.000 de repetições de nosso experimento, que consiste no cálculo da Lei de Coulomb usando uma distância r_i escolhida uniforme-

mente no intervalo $[5, 25 \times 10^{-11}; 5, 35 \times 10^{-11}]$. Ao final das repetições calculamos a média \bar{F}_{MC} das forças obtidas.

O resultado da aplicação do método de Monte Carlo ao nosso problema foi

$$\bar{F}_{MC} = \mathbf{8,193844465022471} \times 10^{-8}, \quad (4.61)$$

que concorda com o resultado 4.59 nas 4 primeiras casas decimais.

Para uma melhor precisão, é necessário aumentar a quantidade de repetições, e esta quantidade aumenta exponencialmente a medida que aumentamos a precisão. Na próxima seção mostramos um método que, além de computacionalmente mais eficaz, também oferece resultados com maior precisão.

4.2.3 Lei de Coulomb com Unscented Transform

Como dito anteriormente, a *Unscented Transform* é uma técnica que visa a aproximação do valor da esperança e dos demais momentos através do cálculo de pontos-sigma e de pesos apropriados. Em [1] temos a formulação para a escolha dos pontos-sigma e dos pesos apresentada a seguir.

Seja X uma variável aleatória com média μ e variância σ^2 . Como X é unidimensional, teremos $2 \cdot (1) + 1 = 3$ pontos-sigma dados pelas expressões

$$x_0 = \mu, \quad x_1 = \mu + \sigma\sqrt{3}, \quad x_2 = \mu - \sigma\sqrt{3}, \quad (4.62)$$

e os pesos dados por

$$w_0 = 2/3, \quad w_1 = w_2 = 1/6. \quad (4.63)$$

No nosso problema, como $r \sim U(5, 25 \times 10^{-11}; 5, 35 \times 10^{-11})$, temos (veja exemplos 2.20 e 2.26)

$$\mu = \frac{5, 25 \times 10^{-11} + 5, 35 \times 10^{-11}}{2} = 5, 3 \times 10^{-11} \quad (4.64)$$

e

$$\sigma^2 = \frac{(5, 35 \times 10^{-11} - 5, 25 \times 10^{-11})^2}{12} \approx 8, 33 \times 10^{-26}. \quad (4.65)$$

Os pontos-sigma são dados por

$$\begin{aligned} r_0 &= 5,3 \times 10^{-11} \\ r_1 &= 5,3 \times 10^{-11} + \sqrt{3 \cdot 8,33 \times 10^{-26}} = 5,35 \times 10^{-11} \\ r_2 &= 5,3 \times 10^{-11} - \sqrt{3 \cdot 8,33 \times 10^{-26}} = 5,25 \times 10^{-11} \end{aligned} \quad (4.66)$$

O resultado \bar{F}_{UT} obtido foi igual a

$$\bar{F}_{UT} = \sum_{i=0}^2 w_i \cdot \left(k \frac{q^2}{r_i^2} \right) = 8,193822918216722 \times 10^{-8}, \quad (4.67)$$

o qual, em relação ao resultado 4.59, está correto nas 6 primeiras casas decimais.

Embora mais precisa e computacionalmente muito mais barata que o método de Monte Carlo, a *Unscented Transform* utiliza uma aproximação da solução ideal do problema de quadraturas. De fato, a escolha dos pontos-sigma e dos pesos extrapola os pesos e os pontos-sigma de uma distribuição uniforme para as demais distribuições (compare a escolha feita aqui com a entrada $N = 3$ da tabela A.2). Conforme vimos na seção 4.1.4, a *Unscented Transform Gaussiana* é baseada nesta solução ótima para cada distribuição de probabilidade, e a solução utilizando esta técnica é apresentada a seguir.

4.2.4 Lei de Coulomb com Unscented Transform Gaussiana

Utilizando o algoritmo 4.11 com os seguinte parâmetros: precisão $2m - 1 = 5$, intervalo $I = [5,25 \times 10^{-11}; 5,35 \times 10^{-11}]$ e função peso $w(x) = 1/(0,1 \times 10^{-11})$ (ver exemplo 2.12), obtemos os pontos-sigma

$$\begin{aligned} r_0 &= 5,3000006688046 \times 10^{-11} \\ r_1 &= 5,3387300974361 \times 10^{-11} \\ r_2 &= 5,2612704379213 \times 10^{-11} \end{aligned} \quad (4.68)$$

e pesos

$$\begin{aligned} w_0 &= 0,44444433816305 \\ w_1 &= 0,27777207364006 \\ w_2 &= 0,27778358819689. \end{aligned} \quad (4.69)$$

O resultado final \bar{F}_{UTG} obtido é igual a

$$\bar{F}_{UTG} = \sum_{i=0}^2 w_i \cdot \left(k \frac{q^2}{r_i^2} \right) = 8,193822874943365 \times 10^{-8}, \quad (4.70)$$

que concorda com o resultado 4.59 nas 11 primeiras casas decimais.

Como podemos observar, a *Unscented Transform Gaussiana* (UTG) nos forneceu, neste experimento, um resultado com quase o dobro de precisão em relação ao resultado obtido com a *Unscented Transform* (UT) usando o mesmo número de pesos e pontos-sigma, mostrando as vantagens de utilizar a solução ideal do problema de quadratura.

Embora o cálculo destes parâmetros seja computacionalmente mais caro no caso da UTG, o ganho de precisão é notável em relação a UT. Além disso, como os pesos e os pontos-sigma independem da função que transforma a variável aleatória, estes podem ser tabelados em relação aos parâmetros da função de distribuição de probabilidade, sendo desnecessário o cálculo dos mesmos a cada novo experimento que use uma variável aleatória com distribuição idêntica. O apêndice A contém tabela dos pontos-sigma e dos pesos para as distribuições mais comuns.

É possível determinar a precisão obtida no caso da UTG através da equação 4.42. Aplicando à mesma ao nosso caso, para $a = 5,25 \times 10^{-11}$ e $b = 5,35 \times 10^{-11}$, temos

$$\begin{aligned}
 |E| &\leq \left| \frac{f^{(2m)}(\xi)}{(2m)!} \right| \int_a^b w(x) [\pi_m(x)]^2 dx \\
 &\leq \left| \frac{F^6(a)}{6!(b-a)} \right| \int_a^b [\pi_3(x)]^2 dx \\
 &\leq \left| \frac{7kq^2}{a^6(b-a)} \right| \int_a^b [\pi_3(x)]^2 dx \\
 &\leq 0,9963 \times 10^{-20},
 \end{aligned} \tag{4.71}$$

de modo que o nosso resultado está preciso até a 19 casa decimal, e como o raio começa na 8 casa, temos 11 casas de precisão, conforme observamos anteriormente.

A tabela 4.1 sintetiza os resultados obtidos. Os valores sublinhado correspondem às correspondências com o resultado exato do problema.

4.3 APLICAÇÃO: LEI DE OHM

Ao aplicarmos uma mesma diferença de potencial V nos extremos de dois condutores geometricamente idênticos, obtemos diferentes correntes. Isto ocorre devido às diferentes *resistências* de cada condutor. Esta resistência pode ser computada através da expressão

$$R = \frac{V}{i}, \tag{4.72}$$

Tabela 4.1: Resultados para a força média segundo a Lei de Coulomb.

Método utilizado	Resultado obtido
Definição de esperança	$8,193822874944370 \times 10^{-8}$
Lei de Coulomb	$8,193093627625485 \times 10^{-8}$
Monte Carlo	$8,193844465022471 \times 10^{-8}$
<i>Unscented Transform</i>	$8,193822918216722 \times 10^{-8}$
<i>Unscented Transform Gaussiana</i>	$8,193822874943365 \times 10^{-8}$

onde V é a diferença de potencial aplicada (medida em *volts*) nos extremos e i e a corrente resultante (medida em *ampères*). A unidade de medida da resistência é o *ohm*, cujo símbolo é Ω .

Porém o termo resistência fica mais evidente se reescrevermos a equação 4.72 como

$$i = \frac{V}{R}, \quad (4.73)$$

que nos diz que, quanto maior for a resistência do condutor, menor será a corrente resultante.

É dito que um condutor obedece a **Lei de Ohm** se a resistência é independente do valor e da polaridade da diferença de potencial aplicada.

Retomando o problema apresentado na introdução deste trabalho, considere o seguinte problema: *qual é a corrente média resultante em um condutor, o qual obedece a Lei de Ohm, submetido a uma diferença de potencial média de 5 V com uma resistência média de 1 Ω ?*

Na introdução vimos que experimentos práticos contradizem a nossa intuição, que seria a aplicação direta da equação 4.73, que nos daria uma corrente de 5A. Novamente vamos resolver o problema apresentado usando as diferentes técnicas já abordada, estudando os resultados obtidos em cada uma delas.

4.3.1 Esperança da Lei de Ohm

Novamente estamos lidando com valores médios, e não valores exatos. Para fins de experimento, suponha que a resistência R é uma variável aleatória com distribuição uniforme no intervalo $I = [a, b]$, onde $a = 0,5\Omega$ e $b = 1,5\Omega$, e que V seja uma variável

aleatória com distribuição exponencial de parâmetro $\lambda = 0,2V$ (observe que a esperança de V é igual a $1/\lambda = 5V$). Como o condutor obedece a Lei de Ohm, então R e V são variáveis independentes.

Usando a expressão da proposição 2.40(b) e o fato de que R e V são independentes, temos

$$\begin{aligned}
 E\{i\} &= \int_a^b \int_0^\infty \left(\frac{1}{b-a}\right) \cdot (\lambda e^{\lambda V}) \cdot \left(\frac{V}{R}\right) dV dR \\
 &= \frac{1}{b-a} \int_a^b \frac{1}{R} \left(\int_0^\infty V \cdot \lambda e^{\lambda V} dV\right) dR \\
 &= \frac{1}{\lambda(b-a)} \int_a^b \frac{1}{R} dR \\
 &= \frac{1}{\lambda(b-a)} \cdot \ln R \Big|_a^b \\
 &= \frac{\ln(b/a)}{\lambda(b-a)} \\
 &= 5,49306144334055A.
 \end{aligned} \tag{4.74}$$

Observe que o valor 4.74 difere, em muito, dos $5A$ que poderíamos imaginar. Embora já citado na introdução, vamos repetir o método de Monte Carlo para comparar o resultado obtido com o valor obtido nesta seção.

4.3.2 Lei de Ohm com Monte Carlo

Realizaremos 1.000.000 repetições do nosso experimento, que consiste no cálculo da corrente i de um condutor que obedece a Lei de Ohm usando uma diferença de potencial $V \sim \text{Exp}(\lambda)$, com $\lambda = 0,2V$, e uma resistência escolhida uniformemente no intervalo $[0, 5\Omega; 1, 5\Omega]$. Ao final das repetições calculamos a média \bar{i}_{MC} das correntes obtidas.

O resultado da aplicação do método de Monte Carlo ao nosso problema foi

$$\bar{i}_{MC} = \mathbf{5,49508013614438}, \tag{4.75}$$

que concorda com o resultado 4.74 apenas nas 2 primeiras casas decimais.

Embora não seja muito preciso, o Método de Monte Carlo nos deu a evidência do perigo da aplicação direta da equação 4.73 à este problema. Vamos ver o quanto de precisão ganharemos ao utilizarmos a *Unscented Transform*.

4.3.3 Lei de Ohm com Unscented Transform

Novamente usando a formulação apresentada em [1], temos a seguinte escolha dos pontos-sigma e dos pesos: seja $R \sim U [0, 5\Omega; 1, 5\Omega]$ e $V \sim \text{Exp}(\lambda)$, com $\lambda = 0, 2V$. Logo temos

$$m_R = 1, \quad \sigma_R \approx 0, 28865 \quad m_V = \sigma_V = 0, 2, \quad (4.76)$$

onde m_R, σ_R, m_V e σ_V são as médias e os desvios-padrão de R e V , respectivamente.

Como $i = V/R$ é bidimensional, teremos $2 \cdot (2) + 1 = 5$ pontos-sigma $x_i = (V_i, R_i)$ dados pelas expressões

$$\begin{aligned} x_0 &= (m_R, m_V) \\ x_1 &= (m_R + \sigma_R\sqrt{5}, m_V), \quad x_2 = (m_R - \sigma_R\sqrt{5}, m_V), \\ x_3 &= (m_R, m_V + \sigma_V\sqrt{5}), \quad x_4 = (m_R, m_V - \sigma_V\sqrt{5}), \end{aligned} \quad (4.77)$$

e os pesos dados por

$$w_0 = 3/5, \quad w_1 = w_2 = w_3 = w_4 = 1/10. \quad (4.78)$$

Aplicando os valores do problema, obtemos

$$\begin{aligned} x_0 &= (1; 5) \\ x_1 &= (1, 64549722436790; 5) \\ x_2 &= (0, 35450277563210; 5) \\ x_3 &= (1; 16, 18033988749895) \\ x_4 &= (1; -6, 1803398874985) \end{aligned} \quad (4.79)$$

O resultado \bar{i}_{UT} obtido foi igual a

$$\bar{i}_{UT} = \sum_{i=0}^4 w_i \cdot \left(\frac{V_i}{R_i} \right) = 5,71428571428571, \quad (4.80)$$

o qual é ainda menos preciso do que Monte Carlo.

Esta divergência em relação ao resultado se deu, principalmente, devido ao fato de que a formulação original da *Unscented Transform*, como comentado anteriormente,

é baseada na extrapolação dos pontos-sigma e dos pesos de uma distribuição normal. Esta aproximação funciona bem em distribuições simétricas (veja a simetria presente nas tabelas A.1 e A.2 do apêndice), mas não tão bem em distribuições assimétricas (veja a tabela A.3 da distribuição exponencial). Este tipo de erro motivou o desenvolvimento da *Unscented Transform Gaussiana*, a qual iremos utilizar a seguir e estudar seu comportamento neste mesmo problema.

4.3.4 Lei de Ohm com Unscented Transform Gaussiana

Como as variáveis R e V são independentes (pois obedecem a Lei de Ohm), podemos utilizar a equação 4.49. Utilizando o algoritmo 4.11 com precisão $2m - 1 = 9$, intervalo $I = [0, 5; 1, 5]$ e função peso $w(x) = 1$ (ver exemplo 2.12), obtemos os pontos-sigma

$$\begin{aligned}
 R_0 &= 1,45308992290003 \\
 R_1 &= 1,26923465474404 \\
 R_2 &= 0,99999999948279 \\
 R_3 &= 0,73076534450042 \\
 R_4 &= 0,54691007690067
 \end{aligned} \tag{4.81}$$

e pesos

$$\begin{aligned}
 w_0^R &= 0,11846344269987 \\
 w_1^R &= 0,23931433551949 \\
 w_2^R &= 0,28444444454491 \\
 w_3^R &= 0,23931433500748 \\
 w_4^R &= 0,11846344222825.
 \end{aligned} \tag{4.82}$$

Aplicando novamente o algoritmo 4.11 com precisão $2m - 1 = 9$, intervalo $I = (0; \infty)$ e função peso $w(x) = \lambda e^{\lambda V}$, com $\lambda = 0,2$ (ver exemplo 2.14), obtemos os pontos-sigma

$$\begin{aligned}
 V_0 &= 63,20400422137566 \\
 V_1 &= 35,42905002929182 \\
 V_2 &= 17,98212885520247 \\
 V_3 &= 7,067015295532080 \\
 V_4 &= 1,317801598590610
 \end{aligned} \tag{4.83}$$

e os pesos

$$\begin{aligned}
 w_0^V &= 0,00002336997239 \\
 w_1^V &= 0,00361175867992 \\
 w_2^V &= 0,07594244968172 \\
 w_3^V &= 0,39866681108319 \\
 w_4^V &= 0,52175561058278
 \end{aligned}
 \tag{4.84}$$

O resultado final \bar{F}_{UTG} obtido é igual a

$$\bar{i}_{UT} = \sum_{i=0}^4 \sum_{j=0}^4 w_i^R w_j^V \cdot \left(\frac{V_j}{R_i} \right) = \mathbf{5,493004620906300},
 \tag{4.85}$$

que concorda com o resultado 4.74 nas 4 primeiras casas decimais.

Tabela 4.2: Resultados para a corrente média segundo a Lei de Ohm.

Método utilizado	Resultado obtido
Definição de esperança	<u>5,49306144334055</u>
Equação 4.73	<u>5,00000000000000</u>
Monte Carlo	<u>5,49508013614438</u>
<i>Unscented Transform</i>	<u>5,71428571428571</u>
<i>Unscented Transform Gaussiana</i>	<u>5,493004620906300</u>

Como podemos observar, novamente a *Unscented Transform Gaussiana* (UTG) nos forneceu, dentre todos os métodos, o resultado mais preciso, embora em escala menor quando comparada a precisão obtida no experimento da Lei de Coulomb. Esta diferença se deve ao comportamento da distribuição exponencial e a alta variância do denominador. Devido ao baixo esforço computacional exigido no cálculo dos pontos-sigma e dos pesos, o ideal seria ampliar a precisão do algoritmo para melhores resultados.

A tabela 4.2 sintetiza os resultados obtidos. Os valores sublinhados correspondem às correspondências com o resultado exato do problema.

Neste capítulo desenvolvemos técnicas para a mensuração da esperança e dos momentos de uma variável aleatória resultante da transformação de uma ou mais variáveis aleatórias. No próximo capítulo estudaremos o problema de se determinar a função de densidade de probabilidade destas variáveis.

5 ENCONTRANDO A FUNÇÃO DE DENSIDADE

A função de densidade de probabilidade de uma variável aleatória X (ou função de probabilidade, no caso de X ser discreta) determina a representação de X , explicitando sua distribuição. Através dela podemos computar, por exemplo, a probabilidade do valor de uma determinada variável aleatória se encontrar em um intervalo pré-definido.

Neste capítulo discutimos as possíveis formas de se determinar a função de densidade de uma variável Y obtida através da transformação de outras variáveis aleatórias, e apontamos as vantagens e desvantagens de cada modo. Também é apresentado um método que combina as principais características dos métodos apresentados até então, numa tentativa de viabilizar o cálculo de tal função num contexto de simulação, numa abordagem distinta da apresentada em [17].

5.1 MÉTODOS

5.1.1 Método Analítico

O método analítico consiste na obtenção da função de densidade de probabilidade através da definição da mesma. Considere X uma variável aleatória com densidade $f_X(x)$ e $Y = g(X)$. De acordo com a definição 2.10, o problema consiste em encontrar uma função $f(t)$ tal que

$$F_Y(y) = \int_{-\infty}^y f(t)dt. \quad (5.1)$$

Ilustraremos este método com um exemplo.

Exemplo 5.1. Considere $Y = aX + b$ com $a, b \in R$ e $a > 0$. Determine $f_Y(y)$.

Solução: Sabendo que

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t)dt \quad (5.2)$$

temos

$$\begin{aligned}F_Y(y) &= P(Y \leq y) \\&= P(aX + b \leq y) \\&= P(aX \leq y - b) \\&= P\left(X \leq \frac{y - b}{a}\right) \\&= \int_{-\infty}^{\frac{y-b}{a}} f_X(x) dx\end{aligned}\tag{5.3}$$

Aplicando na integral acima a seguinte mudança de variáveis

$$t = ax + b,\tag{5.4}$$

os novos limites de integração serão

caso 01: $x = (y - b)/a$

$$\begin{aligned}t &= a\left(\frac{y - b}{a}\right) + b \\&= (y - b) + b \\&= y\end{aligned}\tag{5.5}$$

caso 02: $x \rightarrow -\infty$

$$\begin{aligned}\lim_{x \rightarrow -\infty} t &= \lim_{x \rightarrow -\infty} (ax + b) \\&= -\infty\end{aligned}\tag{5.6}$$

Derivando em x , temos

$$dt = a \cdot dx.\tag{5.7}$$

Por fim,

$$x = \frac{t - b}{a}\tag{5.8}$$

Usando os resultados obtidos,

$$\begin{aligned}F_Y(y) &= \int_{-\infty}^{\frac{y-b}{a}} f_X(x) dx \\&= \int_{-\infty}^y f_X\left(\frac{t - b}{a}\right) \frac{dt}{a}\end{aligned}\tag{5.9}$$

Pela definição de densidade,

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right). \quad (5.10)$$

□

Exemplo 5.2. Considere $X \sim N(0, 1)$ e $Y = aX + b$ com $a, b \in R$ e $a > 0$. Determine $f_Y(y)$.

Solução: Sabemos que

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (5.11)$$

De acordo com o resultado obtido no exemplo 5.1, temos

$$\begin{aligned} f_Y(y) &= \frac{1}{a} f_X\left(\frac{y-b}{a}\right) \\ &= \frac{1}{a} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(\frac{y-b}{a})^2}{2}} \\ &= \frac{1}{\sqrt{2\pi}a} e^{-\frac{(y-b)^2}{2a^2}} \end{aligned} \quad (5.12)$$

Logo, dada uma variável $X \sim N(0, 1)$, a transformação afim $Y = aX + b$ produz uma variável aleatória de distribuição normal com média b e desvio-padrão a , isto é, $Y \sim N(b, a^2)$.

□

A solução do exemplo 5.1 traz consigo idéias interessantes, como a dedução da densidade de Y através da distribuição (e, por conseqüência, da densidade) de X , o cálculo de funções inversas e a mudança de variáveis na integral. A aplicação destas técnicas ao caso genérico $Y = g(X)$ resultará no Método do Jacobiano, que será apresentado na próxima seção.

5.1.2 Método do Jacobiano

Antes de apresentar o Método do Jacobiano, é importante ressaltar algumas características das funções inversíveis. A proposição abaixo determina a derivada da inversa de uma função.

Proposição 5.3. Seja f uma função inversível em um intervalo I e f^{-1} sua inversa. Então

$$\frac{\partial f^{-1}(x)}{\partial x} = \frac{1}{f'(f^{-1}(x))}, \quad \forall x \in f(I). \quad (5.13)$$

Demonstração: Como f é inversível com inversa f^{-1} temos, para todos $x \in f(I)$,

$$x = (f \circ f^{-1})(x). \quad (5.14)$$

Logo, derivando ambos lados da igualdade em x , obtemos

$$\begin{aligned} 1 &= \frac{\partial [f \circ f^{-1}(x)]}{\partial x} \\ &= \frac{\partial [f(f^{-1}(x))]}{\partial x} \\ &= f'(f^{-1}(x)) \cdot \frac{\partial f^{-1}(x)}{\partial x} \end{aligned} \quad (5.15)$$

Portanto,

$$\frac{\partial f^{-1}(x)}{\partial x} = \frac{1}{f'(f^{-1}(x))}. \quad (5.16)$$

□

A proposição 5.3, além de nos mostrar que a derivada da inversa pode ser calculada a partir da derivada da função, permite algumas conclusões, dentre elas:

1. Para que a derivada de f^{-1} em relação a x exista em I é necessário que $f'(x) \neq 0$, $\forall x \in I$.
2. A função f^{-1} tem a mesma orientação de f , pois suas derivadas têm o mesmo sinal. Em outras palavras, se f é crescente, f^{-1} é crescente; se f é decrescente, sua inversa também o será. Isto é importante porque as funções inversíveis em um dado intervalo são monotônicas no mesmo.

No teorema abaixo apresentaremos a versão unidimensional do **Método do Jacobiano** para o cálculo da função de densidade de probabilidade.

Teorema 5.4. Seja uma variável aleatória X com densidade $f_X(x)$ e $Y = g(X)$, onde $g(X)$ é uma função inversível. Então

$$F_Y(y) = \int_{-\infty}^y f_X(g^{-1}(u)) \left| \frac{\partial g^{-1}(u)}{\partial u} \right| du \quad (5.17)$$

Demonstração: Sabemos que

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t) dt \quad (5.18)$$

Considerando inicialmente g uma função crescente, temos

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(g(X) \leq y) \\ &= P(X \leq g^{-1}(y)) \\ &= \int_{-\infty}^{g^{-1}(y)} f_X(t) dt \end{aligned} \quad (5.19)$$

Aplicando a mudança de variáveis $u = g(t)$, teríamos

$$du = g'(t) dt \quad (5.20)$$

Os limites de integração seriam

caso 01: $t = g^{-1}(y)$

$$\begin{aligned} u &= g(t) \\ &= g(g^{-1}(y)) \\ &= y \end{aligned} \quad (5.21)$$

caso 02: $t \rightarrow -\infty$

$$\begin{aligned} \lim_{t \rightarrow -\infty} u &= \lim_{t \rightarrow -\infty} g(t) \\ &= -\infty \end{aligned} \quad (5.22)$$

Aplicando a mudança de variáveis, temos

$$F_Y(y) = \int_{-\infty}^y f_X(g^{-1}(u)) \frac{du}{|g'(t)|} \quad (5.23)$$

pois $g'(x) > 0$ (g é crescente por hipótese). Mas

$$\frac{1}{|g'(t)|} = \frac{1}{|g'(g^{-1}(u))|} = \left| \frac{\partial g^{-1}(u)}{\partial u} \right| \quad (5.24)$$

Portanto,

$$F_Y(y) = \int_{-\infty}^y f_X(g^{-1}(u)) \left| \frac{\partial g^{-1}(u)}{\partial u} \right| \quad (5.25)$$

Falta mostrar que o teorema é verdadeiro quando g é decrescente. Neste caso, temos

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(g(X) \leq y) \\ &= P(X > g^{-1}(y)) \\ &= 1 - P(X \leq g^{-1}(y)) \\ &= 1 - \int_{-\infty}^{g^{-1}(y)} f_X(t) dt \end{aligned} \quad (5.26)$$

Usando novamente a mudança de variáveis $u = g(t)$, teríamos $u = y$ quando $t = g^{-1}(y)$ e $u \rightarrow \infty$ quando $t \rightarrow -\infty$. Aplicando a mudança de variáveis, e lembrando que $g'(x) < 0$ (g é decrescente por hipótese), temos

$$\begin{aligned} F_Y(y) &= 1 - \int_{\infty}^y f_X(g^{-1}(u)) \left(\frac{-du}{|g'(t)|} \right) \\ &= 1 - \int_y^{\infty} f_X(g^{-1}(u)) \left| \frac{\partial g^{-1}(u)}{\partial u} \right| \\ &= 1 - \left(1 - \int_{-\infty}^y f_X(g^{-1}(u)) \left| \frac{\partial g^{-1}(u)}{\partial u} \right| \right) \\ &= \int_{-\infty}^y f_X(g^{-1}(u)) \left| \frac{\partial g^{-1}(u)}{\partial u} \right| \end{aligned} \quad (5.27)$$

□

Uma condição para a aplicação do Método do Jacobiano é que a função $g(X)$ seja inversível em seu domínio. Caso ela não seja inversível para todos os possíveis valores da variável X , basta subdividir o conjunto domínio D em subconjuntos D_i tais que

1. $D_i \cap D_j = \emptyset$ para $i \neq j$,
2. $D = \bigcup_i D_i$
3. as restrições g_i de g aos conjuntos D_i ($g_i : D_i \rightarrow I_i$, $g_i(x_i) = g(x_i) \forall x_i \in D_i$) sejam inversíveis.

Procedendo desta forma, a função de densidade f_Y seria dada pela soma das funções de densidade f_Y^i obtidas para cada restrição g_i de $g(X)$.

Exemplo 5.5. Considere $X \sim N(0, 1)$ e $Y = X^2$. Determine $f_Y(y)$.

Solução: A função $g(X) = X^2$ não é inversível em seu domínio $D = (-\infty, \infty)$. Tomando $D_1 = (-\infty, 0)$ e $D_2 = (0, \infty)$, as restrições g_1 e g_2 são inversíveis. Observe que $g(D) = I = (0, \infty)$.

Calculando as inversas, temos que $g_1^{-1}(y) = -\sqrt{y}$ e $g_2^{-1}(y) = \sqrt{y}$. Computando o módulo das derivadas, segue que

$$\left| \frac{\partial g_1^{-1}(y)}{\partial y} \right| = \left| \frac{\partial g_2^{-1}(y)}{\partial y} \right| = \frac{1}{2\sqrt{y}}. \quad (5.28)$$

Pelo Método do Jacobiano,

$$f_Y(y) = f_X(g_1^{-1}(y)) \cdot \left| \frac{\partial g_1^{-1}(y)}{\partial y} \right| + f_X(g_2^{-1}(y)) \cdot \left| \frac{\partial g_2^{-1}(y)}{\partial y} \right|. \quad (5.29)$$

Como X tem distribuição normal com média 0 e variância 1, sabemos pelo exemplo 2.13 que

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \quad (5.30)$$

Substituindo em 5.29, segue que

$$\begin{aligned} f_Y(y) &= f_X(g_1^{-1}(y)) \cdot \left| \frac{\partial g_1^{-1}(y)}{\partial y} \right| + f_X(g_2^{-1}(y)) \cdot \left| \frac{\partial g_2^{-1}(y)}{\partial y} \right| \\ &= \left| \frac{\partial g_1^{-1}(y)}{\partial y} \right| [f_X(g_1^{-1}(y)) + f_X(g_2^{-1}(y))] \\ &= \frac{1}{2\sqrt{y}} \left(\frac{1}{\sqrt{2\pi}} e^{\frac{(-\sqrt{y})^2}{2}} + \frac{1}{\sqrt{2\pi}} e^{\frac{(\sqrt{y})^2}{2}} \right) \\ &= \frac{1}{\sqrt{y}} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{y}{2}}. \end{aligned} \quad (5.31)$$

Logo,

$$f_Y(y) = \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}, \quad y \in I. \quad (5.32)$$

Neste caso, dizemos que Y tem distribuição *qui-quadrado* $\chi(1)^2$ com um grau de liberdade.

□

Abaixo temos a generalização do Método do Jacobiano para o caso multidimensional.

Teorema 5.6. Seja $\tilde{X} = (x_1, x_2, \dots, x_n)$ uma variável aleatória com densidade $f_X(x)$ e sejam $D, I \in \mathbb{R}^n$ regiões abertas, e $g : D \rightarrow I$ uma bijeção entre D e I . Considere a variável $\tilde{Y} = g(\tilde{X}) = (y_1, y_2, \dots, y_n)$. Se definirmos o *jacobiano de \tilde{X} em relação a \tilde{Y}* por

$$J(\tilde{X}, \tilde{Y}) = \frac{\partial(x_1, x_2, \dots, x_n)}{\partial(y_1, y_2, \dots, y_n)} = \det \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{bmatrix},$$

então, para todo $\tilde{y} \in I$,

$$f_Y(\tilde{y}) = f_X(g^{-1}(\tilde{y})) \cdot |J(\tilde{X}, \tilde{Y})(\tilde{y})|. \quad (5.33)$$

Observe que, para utilizarmos o método do Jacobiano, ambos vetores aleatórios devem ter a mesma dimensão. Para se computar a função de densidade de vetor aleatório \tilde{Y} cuja dimensão é inferior a do vetor \tilde{X} , é necessário completar a dimensão de \tilde{Y} acrescentando transformações extras do vetor \tilde{X} e, ao final do processo, calcular a densidade marginal da variável desejada. O próximo exemplo ilustrará este caso.

Exemplo 5.7. Sejam X e Y duas variáveis aleatórias independentes com distribuição uniforme no intervalo $I = [0, 1]$. Determine a função de densidade $f_Z(z)$ da variável aleatória $Z = X + Y$.

Solução: Observe que, para $g_1(X, Y) = X + Y$ e $D = [0, 2]$,

$$g_1 : I \times I \rightarrow D, \quad (5.34)$$

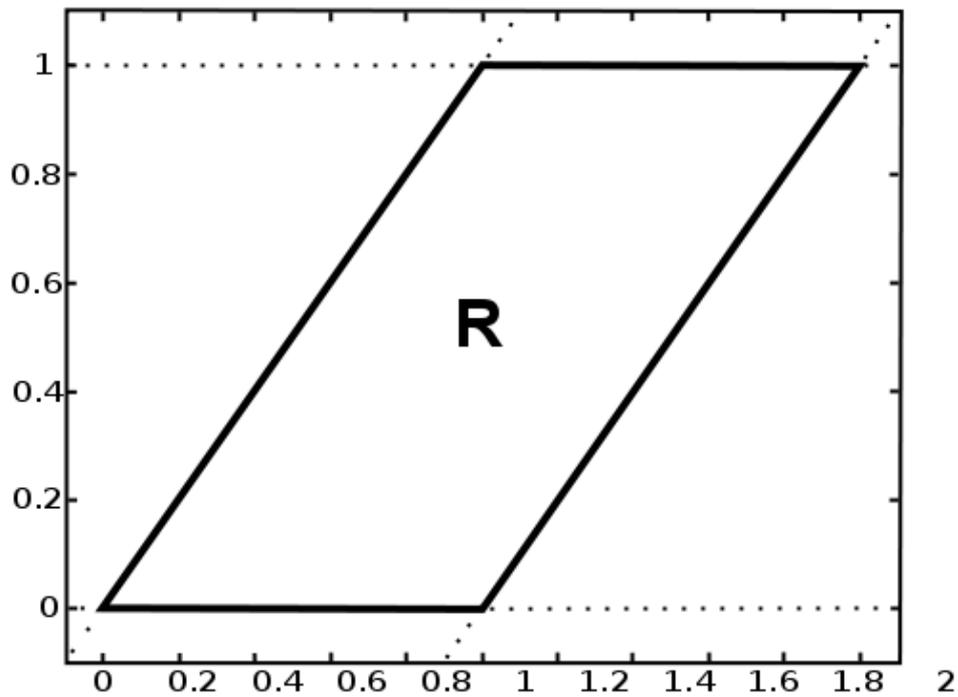


Figura 5.1: Região de integração de $f_{Z,W}(z, w)$

de modo que as dimensões não coincidem. Definiremos então uma segunda transformação $W = g_2(X, Y) = Y$ e a função

$$g(X, Y) = (g_1(X, Y), g_2(X, Y)) = (X + Y, Y) = (Z, W), \quad (5.35)$$

onde as dimensões coincidem.

Calculando as inversas, temos $Y = W$ e $X = Z - W$, de modo que o jacobiano será dado por

$$\begin{aligned} J(z, w) &= \frac{\partial(X, Y)}{\partial(Z, W)}(z, w) \\ &= \det \begin{bmatrix} \frac{\partial X}{\partial Z}(z, w) & \frac{\partial X}{\partial W}(z, w) \\ \frac{\partial Y}{\partial Z}(z, w) & \frac{\partial Y}{\partial W}(z, w) \end{bmatrix} \\ &= \det \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \\ &= 1. \end{aligned} \quad (5.36)$$

Segundo o Método do Jacobiano, a densidade conjunta $f_{Z,W}(z, w)$ (consulte a definição

2.34) será dada por

$$\begin{aligned}
 f_{Z,W}(z, w) &= f_{X,Y}(x = x(z, w), y = y(z, w))J(z, w) \\
 &= f_X(x = x(z, w))f_Y(y = y(z, w))J(z, w) \\
 &= 1.
 \end{aligned} \tag{5.37}$$

A função marginal $f_Z(z)$ (veja a definição 2.36) é dada por

$$f_Z(z) = \int_{-\infty}^{\infty} f_{Z,W}(z, w)dw. \tag{5.38}$$

Um erro comum que pode acontecer neste passo é o da determinação errônea dos limites de integração. No primeiro momento, pelo fato de que $W = Y$, pode parecer a integral 5.38 terá limite inferior igual a 0 e superior igual a 1, o que não é verdadeiro. Para se determinar os limites de integração devemos analisar a região de integração R , que é delimitada pelas quatro expressões

$$w = 0, \quad w = 1, \quad z - w = 0, \quad z - w = 1, \tag{5.39}$$

mostrada na figura 5.1.

Logo, para $0 \leq z \leq 1$, temos

$$f_Z(z) = \int_0^z f_{Z,W}(z, w)dw = z \tag{5.40}$$

e, para $1 < z \leq 2$,

$$f_Z(z) = \int_{z-1}^1 f_{Z,W}(z, w)dw = 2 - z. \tag{5.41}$$

Portanto,

$$f_Z(z) = \begin{cases} z, & \text{se } 0 \leq z \leq 1, \\ 2 - z, & \text{se } 1 < z \leq 2, \\ 0, & \text{caso contrário.} \end{cases} \tag{5.42}$$

Observe que, por outro lado,

$$f_W(w) = \int_w^{w+1} f_{Z,W}(z, w)dz = (w + 1) - w = 1, \tag{5.43}$$

conforme esperado, pois $W = Y \sim U(0, 1)$.

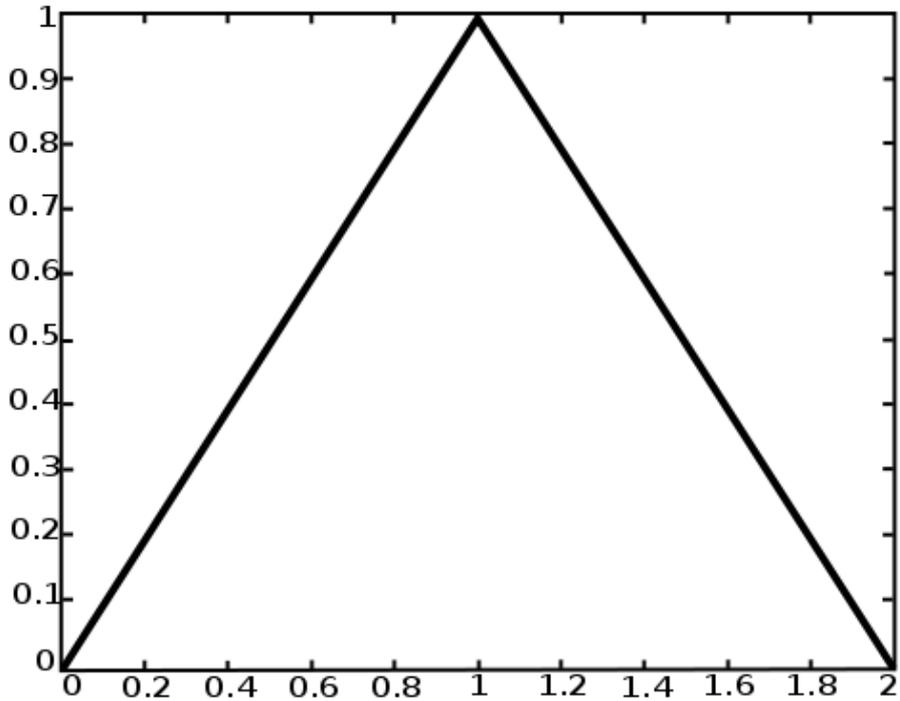


Figura 5.2: Função de densidade de $Z = X + Y$

O Método do Jacobiano é resultado, como dito anteriormente, da generalização do método analítico. Embora resolva o problema de se determinar a densidade de uma variável $Y = g(X)$, este método é de difícil implementação prática, principalmente pela dificuldade em se determinar as funções inversas e suas derivadas (ou o Jacobiano, no caso multidimensional). A próxima seção propõe uma adaptação do Método do Jacobiano, que visa a superação destas dificuldades.

5.1.3 Método do Jacobiano Numérico

Como visto nas seções anteriores, tanto o método analítico quanto o método do Jacobiano trazem consigo dificuldades em sua operacionalização, principalmente no que diz respeito ao cálculo de funções inversas, derivadas parciais e jacobianos e à determinação da região de integração no caso multidimensional. Nesta seção reformularemos o Método do Jacobiano, visando sua implementação computacional. Denominaremos esta adaptação de **Método do Jacobiano Numérico**.

O Método do Jacobiano Numérico é baseado nas seguintes estratégias e premissas:

1. substituição da função $\tilde{Y} = g(\tilde{X})$ por sua aproximação de Taylor $P_N(\tilde{x})$;
2. substituição do jacobiano $J(\tilde{X}, \tilde{Y})$ por $1/J(\tilde{Y}, \tilde{X})$;
3. particionamento do domínio das variáveis aleatórias e
4. definição da função inversa através do cálculo de zeros de polinômios.

A estratégia 1 se dá pelo fato de que o cálculo tanto do Jacobiano quanto das derivadas fica simplificado pelo fato de $P_N(\tilde{x})$ ser um polinômio (veja definição 3.3). Além disso, o cálculo da função inversa $g^{-1}(\tilde{y})$ se reduz ao problema de encontrar, pontualmente, os zeros do polinômio $p(\tilde{x}) = P_N(\tilde{x}) - \tilde{y}$, que é nossa premissa 4.

A estratégia 2 é a generalização da equação 5.13, e vai de encontro aos princípios da estratégia 1, pois as derivadas do jacobiano são, nesta formulação, tomadas em relação às variáveis do polinômio $P_N(\tilde{x})$, sem necessidade de inversões.

Por fim, o item 3 é uma necessidade pois, ao se trabalhar em ambientes computacionais, temos que lidar com quantidade finitas, daí a necessidade de se particionar os domínios. Ademais, um bom particionamento gera um algoritmo mais eficiente e com resultados mais precisos. A definição de partição é dada a seguir.

Definição 5.8. Seja $I \in \mathbb{R}$ um intervalo e M um número inteiro positivo. Dizemos que o conjunto de M intervalos $\{\Delta^i\}_M$ é uma *partição* Δ_I de M elementos de I se

1. $\Delta^i \subseteq I$ para $i = 0, 1, \dots, M - 1$,
2. $\Delta^i \cap \Delta^j = \emptyset$, para $i \neq j$,
3. $\bigcup \Delta^i = I$.

Em situações computacionais às vezes se torna conveniente escolher, para cada intervalo da partição, um número contido neste intervalo como seu representante. A definição abaixo formaliza este conceito.

Definição 5.9. Seja I um intervalo e $\Delta_I = \{\Delta^i\}_M$ uma partição de I . Dizemos que o conjunto de M números $\{\delta_i\}_M$ é um *conjunto de representantes da partição* Δ_I se, para $i = 0, 1, \dots, M - 1$, $\delta_i \in \Delta^i$.

Em geral, se escolhe partições Δ_I cujos intervalos Δ^i tem mesmo comprimento. Para os representantes, uma escolha comum é o ponto médio de cada intervalo. Porém estas são apenas sugestões, não constituindo regra nem requisito para os algoritmos que serão apresentados.

O algoritmo abaixo descreve o Método do Jacobiano Numérico para variáveis unidimensionais.

Algoritmo 5.10. *Método do Jacobiano Numérico unidimensional.*

ENTRADA: A variável aleatória X , sua densidade $f_X(x)$ e sua média m_X , a transformação g tal que $Y = g(X)$, a ordem N da aproximação $P_N(x)$ e a quantidade M de pontos para a aproximação da densidade.

SAÍDA: A função de densidade $f_Y(y)$ aproximada por M pontos.

1. Determine os $N + 1$ coeficientes c_i tais que

$$P_N(x) = \sum_{i=0}^N c_i x^i \quad (5.44)$$

seja a aproximação de Taylor de ordem N de $g(X)$ em m_X .

2. Calcule os N coeficientes s_i dados por

$$s_i = \sum_{i=0}^{N-1} (i + 1) \cdot c_{i+1}. \quad (5.45)$$

Deste modo,

$$S(x) = \sum_{i=0}^{N-1} s_i x^i \quad (5.46)$$

é a derivada primeira de $P_N(x)$.

3. Gere uma partição Δ_Y de M elementos do intervalo $g(I)$, onde I é o domínio da transformação $g(X)$.

4. Escolha um conjunto de representantes $\{\delta_i\}_M$ para a partição Δ_Y .
5. Para $i = 0, 1, \dots, M - 1$ faça
 - (a) Faça $f_Y(\delta_i) = 0$.
 - (b) Calcule os N zeros x_j do polinômio $p_i(x) = P_N(x) - \delta_i$.
 - (c) Para $j = 0, 1, \dots, N - 1$, se $x_j \in I$, faça

$$f_Y(\delta_i) = f_Y(\delta_i) + \frac{f_X(x_j)}{|S(x_j)|}. \quad (5.47)$$

Para computar $f_Y(y)$ para um valor $y \neq \delta_i$, podemos recorrer a interpolação. A definição abaixo mostra a interpolação mais simples de todas, que é a *interpolação linear* ([15], pág. 218). No entanto, a escolha do método de interpolação é livre, e deve atender aos requisitos de precisão do problema.

Definição 5.11. Dados dos pares de pontos $P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$, a *interpolação linear* de P_1 e P_2 é a reta que une estes dois pontos, isto é, dado $x_1 \leq x \leq x_2$,

$$y = y_1 + (x - x_1) \frac{y_2 - y_1}{x_2 - x_1}. \quad (5.48)$$

Para o caso multidimensional, devemos fazer alguns ajustes no algoritmo 5.10, que vão desde a substituição da derivada da função por seu Jacobiano até a escolha de uma variável aleatória auxiliar. O algoritmo abaixo mostra o caso bidimensional: os algoritmos para as demais dimensões são análogos e serão deixados a cargo do leitor.

Algoritmo 5.12. *Método do Jacobiano Numérico bidimensional.*

ENTRADA: As variáveis aleatórias X e Y , suas respectivas médias m_X e m_Y , a densidade conjunta $f_{X,Y}(x, y)$, a transformação g tal que $Z = g(X, Y)$, a ordem N da aproximação $P_N(x, y)$ e a quantidade M de pontos para a aproximação da densidade.

SAÍDA: A função de densidade $f_Z(z)$ aproximada por M pontos.

1. Determine os $(N + 1)^2$ coeficientes $c_{i,j}$ tais que

$$P_N(x, y) = \sum_{j=0}^N \sum_{i=0}^N c_{i,j} x^i y^j \quad (5.49)$$

seja a aproximação de Taylor de ordem N de $g(X, Y)$ em (m_X, m_Y) .

2. Calcule os N coeficientes $s_{i,j}$ dados por

$$s_{i,j} = \sum_{j=0}^N \sum_{i=0}^{N-1} (i+1) \cdot c_{i+1,j}. \quad (5.50)$$

Deste modo,

$$J(x, y) = \sum_{j=0}^N \sum_{i=0}^{N-1} s_{i,j} x^i y^j \quad (5.51)$$

é o Jacobiano de $P_N(x, y)$.

3. Gere uma partição Δ_Z de M elementos do intervalo $g(I, J)$, onde $I \times J$ é o domínio da transformação $g(X, Y)$, e uma partição Δ_Y de M elementos do intervalo J .

4. Escolha um conjunto de representantes $\{\delta_i\}_M$ para a partição Δ_Z e um conjunto de representantes $\{\omega_i\}_M$ para a partição Δ_Y .

5. Para cada $k = 0, 1, \dots, M - 1$ faça

(a) Faça $f_Z(\delta_k) = 0$.

(b) Para cada $j = 0, 1, \dots, M - 1$ faça

i. Calcule os N zeros x_i do polinômio $p_{k,j}(x) = P_N(x, \omega_j) - \delta_k$.

ii. Para $i = 0, 1, \dots, N - 1$, se $x_i \in I$, faça

$$f_Z(\delta_k) = f_Z(\delta_k) + \|\Delta_Y^j\| \frac{f_{X,Y}(x_i, y_j)}{|J(x_i, y_j)|}, \quad (5.52)$$

onde $\|\Delta_Y^j\|$ é o comprimento do intervalo Δ_Y^j .

Valem alguns esclarecimentos sobre o algoritmo 5.12:

- A expressão para o cálculo do Jacobiano dada no passo 2 decorre do seguinte fato: considere a função $h(X, Y) = (g(X, Y), Y) = (Z, W)$. Segundo o teorema 5.6, temos

$$\begin{aligned} J(x, y) &= \frac{\partial(Z, W)}{\partial(X, Y)}(x, y) \\ &= \det \begin{bmatrix} \frac{\partial Z}{\partial X} & \frac{\partial Z}{\partial Y} \\ \frac{\partial W}{\partial X} & \frac{\partial W}{\partial Y} \end{bmatrix} \\ &= \det \begin{bmatrix} \frac{\partial g(X, Y)}{\partial X} & \frac{\partial g(X, Y)}{\partial Y} \\ 0 & 1 \end{bmatrix} \\ &= \frac{\partial g(X, Y)}{\partial X}(x, y). \end{aligned} \quad (5.53)$$

- Como a densidade $f_Z(z)$ é a densidade marginal Z de $f_{Z,W}(z, w)$, então

$$f_Z(z) = \int_{-\infty}^{\infty} f_{Z,W}(z, w) dw. \quad (5.54)$$

Usando o teorema do Jacobiano e fazendo $Y = W$, uma aproximação para $f_Z(z)$ no ponto z_k seria

$$f_Z(z_k) \approx \sum_{P_k} \frac{f_{X,Y}(x_i, y_j)}{|J(x_i, y_j)|} \|\Delta_Y^j\|, \quad (5.55)$$

onde $P_k = \{(x_i, y_j) | z_k = g(x_i, y_j), x_i \in I, y_j \in J\}$, o que justifica a partição de J no passo 4 e o termo $\|\Delta_Y^j\|$ no passo 5(b).

Na próxima seção ilustraremos o uso de ambos algoritmos em problemas práticos do eletromagnetismo.

5.2 APLICAÇÃO: ENERGIA ARMAZENADA POR UM INDUTOR

Da mesma forma que é possível armazenar energia potencial elétrica no campo elétrico gerado por duas cargas de sinais opostos ao afastá-las, podemos armazenar energia num campo magnético, por exemplo, afastando dois fios longos, rígidos, paralelos que transportam uma corrente de mesmo sentido.

Segundo [16], pág. 241, a energia U_B total armazenada por um indutor L transportando um corrente i é dada pela expressão

$$U_B = \frac{1}{2} Li^2. \quad (5.56)$$

Considere o seguinte problema, baseado no problema apresentado em [16], pág. 241: *suponha uma bobina com indutância de 53 mH e uma resistência de 0,35 Ω . Qual é a probabilidade da energia U_B armazenada no campo magnético resultante da aplicação de uma fem $\mathcal{E} \sim N(12V; 0, 01)$ na referida bobina, após a corrente atingir seu valor de equilíbrio, ser maior do que 31,5J?*

Para a resolução deste problema, é necessário o conhecimento da função de densidade de $f_{U_B}(u)$. Também usaremos a expressão do crescimento da corrente ([16]), dada por

$$i = \frac{\mathcal{E}}{R} (1 - e^{-t/\tau_L}), \quad (5.57)$$

onde τ_L é a *constante de tempo indutiva*. De acordo com a equação 5.57, a corrente de equilíbrio é dada por

$$i_\infty = \lim_{t \rightarrow \infty} \frac{\mathcal{E}}{R} (1 - e^{-t/\tau_L}) = \frac{\mathcal{E}}{R}. \quad (5.58)$$

Resolveremos primeiramente o problema usando o método do jacobiano, e em seguida utilizaremos o método do jacobiano numérico, comparando os resultados obtidos.

5.2.1 Método do Jacobiano

Primeiramente devemos determinar a distribuição da variável i . De acordo com a equação 5.10, e sabendo que $\mathcal{E} \sim N(m_{\mathcal{E}}; \sigma_{\mathcal{E}})$ e $i = (1/R) \cdot \mathcal{E}$, temos

$$\begin{aligned} f_i(i) &= R \cdot f_{\mathcal{E}}\left(\frac{i - 0}{1/R}\right) \\ &= R \cdot \frac{1}{\sqrt{2\pi}\sigma_{\mathcal{E}}} \cdot e^{-(Ri - m_{\mathcal{E}})^2 / 2\sigma_{\mathcal{E}}^2} \\ &= \frac{1}{\sqrt{2\pi}(\sigma_{\mathcal{E}}/R)} \cdot e^{-(i - m_{\mathcal{E}}/R)^2 / 2(\sigma_{\mathcal{E}}/R)^2}, \end{aligned} \quad (5.59)$$

ou seja, $i = N(m_{\mathcal{E}}/R, \sigma_{\mathcal{E}}/R) = (2, 4; 0, 002)$.

A função inversa $i = g^{-1}(U_B)$ é dada por

$$i = \sqrt{\frac{2U_B}{L}} \quad (5.60)$$

e sua derivada primeira é

$$\frac{\partial i}{\partial U_B} = \sqrt{\frac{2}{L}} \cdot \frac{1}{2\sqrt{U_B}} = \frac{1}{\sqrt{2LU_B}}. \quad (5.61)$$

Deste modo, de acordo com o teorema 5.4,

$$f_{U_B}(u) = f_i(g^{-1}(u)) \cdot \left| \frac{\partial g^{-1}(u)}{\partial U_B} \right| = \frac{1}{0,002\sqrt{4\pi Lu}} \cdot e^{-(\sqrt{2u/L} - 2,4)^2 / 8 \times 10^{-6}}. \quad (5.62)$$

Por fim, a probabilidade de U_B ser maior do que 31,5 J é dada por

$$\begin{aligned} P(U_B > 31,5) &= 1 - P(U_B \leq 31,5) \\ &= \int_{-\infty}^{31,5} f_{U_B}(u) du \\ &= 0,25134636663983, \end{aligned} \quad (5.63)$$

de modo que em mais de 25% dos casos a energia acumulada U_B ultrapassaria o limite de 31,5 J.

A título de comparação, a aplicação direta da equação 5.56 para $\mathcal{E} = 12V$ nos daria

$$U_B = \frac{1}{2}Li^2 = U_B = \frac{1}{2}(53 \times 10^{-3}) \left(\frac{12}{0,35} \right)^2 = 31,15102040816326, \quad (5.64)$$

enquanto a aplicação do método da *Unscented Transform Gaussiana de ordem 9* nos daria

$$m_{U_B} = 31,15318367346939, \quad \sigma_{U_B} = 0,51919268699656, \quad (5.65)$$

mostrando a alta variância de U_B , informação que não está presente na equação 5.56 e que justifica a alta probabilidade de se superar o valor estabelecido.

5.2.2 Método do Jacobiano Numérico

Fazendo $i \sim N(2, 4; 0, 002)$ (veja a seção anterior para a dedução da distribuição de i), $U_B = g(i)$ e $N = 2$, teremos

$$\begin{aligned} P_2(i) &= g(m_i) + g'(m_i)(i - m_i) + \frac{g''(i)}{2}(i - m_i)^2 \\ &= \frac{1}{2}Lm_i^2 + Lm_i \cdot (i - m_i) + \frac{L}{2}(i - m_i)^2 \\ &= \frac{1}{2}Li^2, \end{aligned} \quad (5.66)$$

enquanto a derivada de $X(i)$ de $g(i)$ é igual a

$$S(i) = Li, \quad (5.67)$$

onde m_i é a média da variável aleatória i (neste exemplo, $m_i = 2,4A$).

As figuras 5.3, 5.4 e 5.5 mostram os comparativos entre as funções de densidade obtidas usando o Método do Jacobiano e o Método do Jacobiano Numérico usando o algoritmo 5.10 com $M = 15, 25$ e 50 , respectivamente. A interpolação utilizada nos três casos foi a interpolação linear (ver definição 5.11).

A probabilidade de U_B ser menor do que $31,5J$, computada usando a aproximação de 50 pontos, é igual a

$$P(U_B > 31,5) = 1 - P(U_B \leq 31,5) = 0.25222338302707, \quad (5.68)$$

que coincide com o resultado analítico nas duas primeiras casas. Para maiores precisões, é necessário aumentar o número de pontos M .

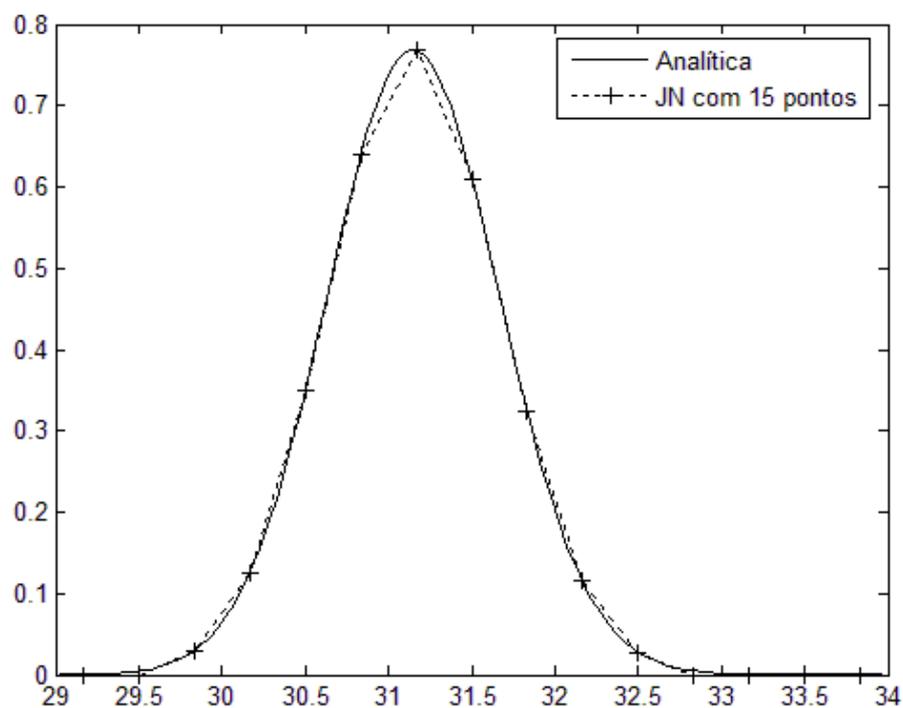


Figura 5.3: Comparativo entre a f_{U_B} obtida analiticamente e através do Método do Jacobiano Numérico com 15 pontos

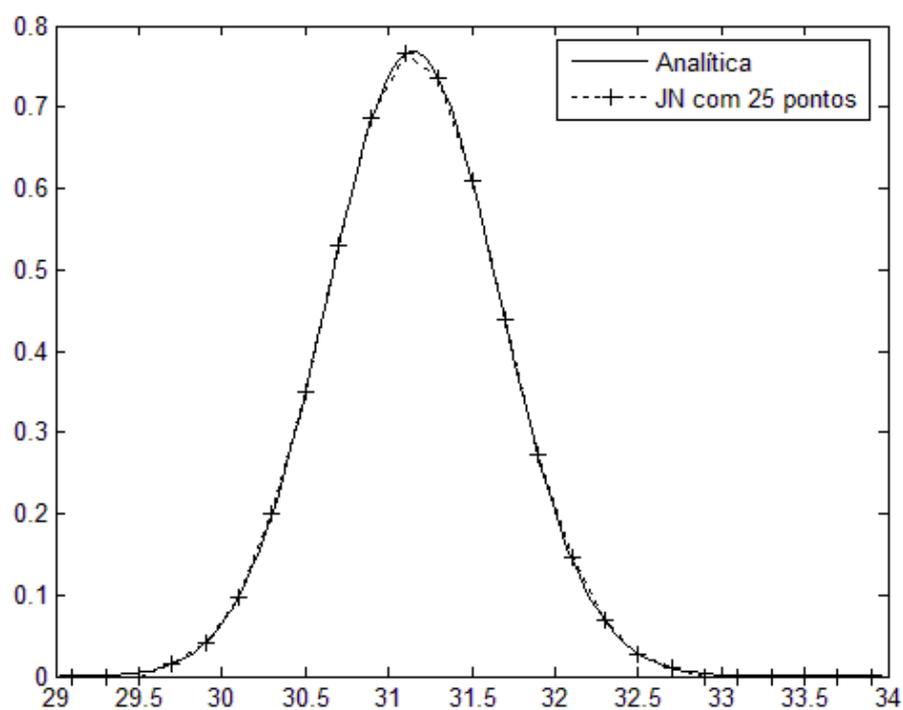


Figura 5.4: Comparativo entre a f_{U_B} obtida analiticamente e através do Método do Jacobiano Numérico com 25 pontos

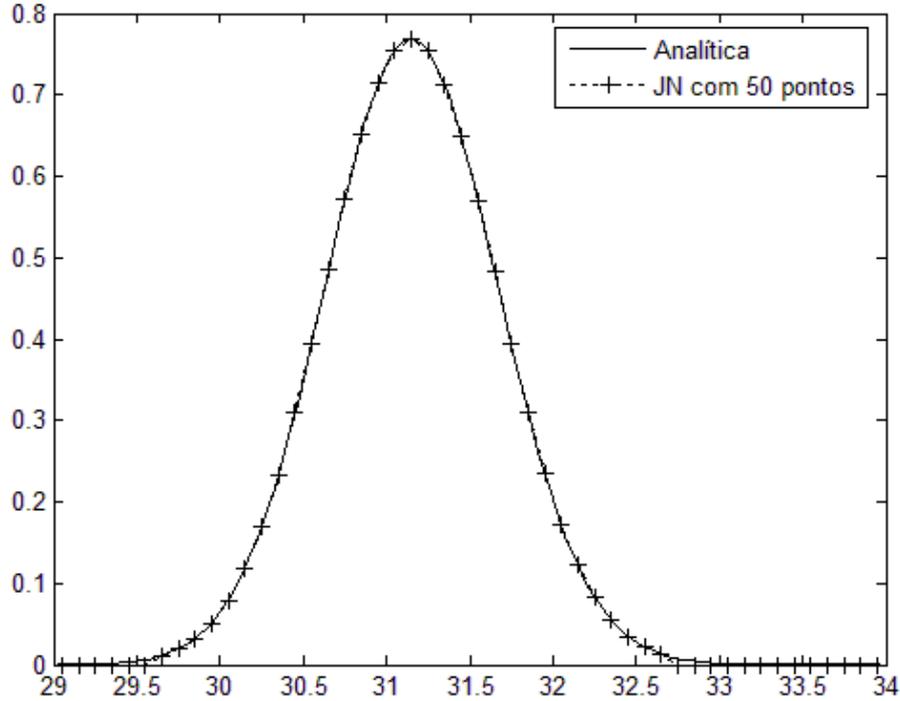


Figura 5.5: Comparativo entre a f_{U_B} obtida analiticamente e através do Método do Jacobiano Numérico com 50 pontos

5.3 APLICAÇÃO: CAPACITOR DE PLACAS PARALELAS

Segundo [16], pág. 94, a capacitância de um capacitor de placas paralelas separadas por uma distância de d metros e cuja carga positiva está envolta por uma superfície gaussiana de área A , é dada por

$$C = \epsilon_0 \frac{A}{d}, \quad (5.69)$$

onde ϵ_0 é a *constante de permissividade*, e vale

$$\epsilon_0 = 8,85 \times 10^{-12} F/m. \quad (5.70)$$

A figura 5.6 ilustra esta situação.

Usando os dados da questão encontrada em [16], pág. 104, formulemos o seguinte problema: *Sejam $d \sim U(0,01237; 0,01243)$ (em metros) e $A \sim U(0,01149; 0,01151)$ (em metros quadrados) duas variáveis aleatórias independentes onde d representa a distância entre duas placas paralelas de um capacitor cuja carga positiva está envolta por uma superfície gaussiana de área A . Determine a capacitância C_M tal que a probabilidade de C ser menor ou igual a C_M seja de 95%.*

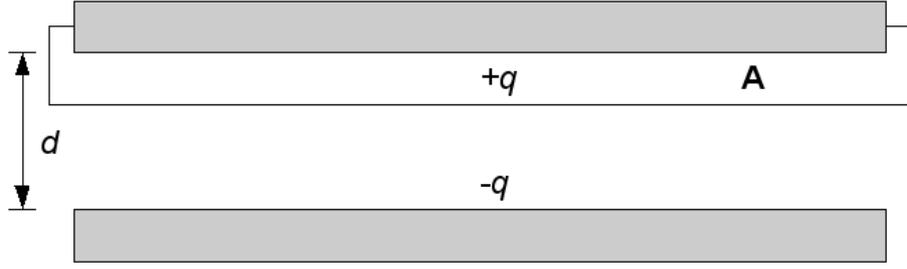


Figura 5.6: Capacitor de placas paralelas

Novamente resolveremos o problema tanto no Método do Jacobiano quanto no Método do Jacobiano Numérico.

5.3.1 Método do Jacobiano

Novamente a solução do problema exige a determinação da função de densidade de probabilidade da variável em questão. Seja $g(A, d) = (Z, W) = (\epsilon_0 A/d, d)$. Calculando a inversa $g^{-1}(Z, W)$, temos $(A, d) = (ZW/\epsilon_0, W)$. Deste modo

$$\begin{aligned}
 J(z, w) &= \frac{\partial(A, d)}{\partial(Z, W)}(z, w) \\
 &= \det \begin{bmatrix} \frac{\partial X}{\partial Z}(z, w) & \frac{\partial X}{\partial W}(z, w) \\ \frac{\partial Y}{\partial Z}(z, w) & \frac{\partial Y}{\partial W}(z, w) \end{bmatrix} \\
 &= \det \begin{bmatrix} \frac{w}{\epsilon_0} & \frac{z}{\epsilon_0} \\ 0 & 1 \end{bmatrix} \\
 &= \frac{w}{\epsilon_0}.
 \end{aligned} \tag{5.71}$$

Segundo o teorema 5.6,

$$\begin{aligned}
 f_{Z,W}(z, w) &= f_{A,d}(A = A(z, w), d = d(z, w))J(z, w) \\
 &= f_A(A = A(z, w))f_d(d = d(z, w))J(z, w) \\
 &= \frac{1}{(0, 01243 - 0, 01237)} \cdot \frac{1}{(0, 01151 - 0, 01149)} \cdot \frac{w}{\epsilon_0} \\
 &= \frac{w}{1, 2\epsilon_0 \times 10^{-9}}.
 \end{aligned} \tag{5.72}$$

Os limites de integração para o cálculo da densidade marginal $f_Z(z)$ são dados pela

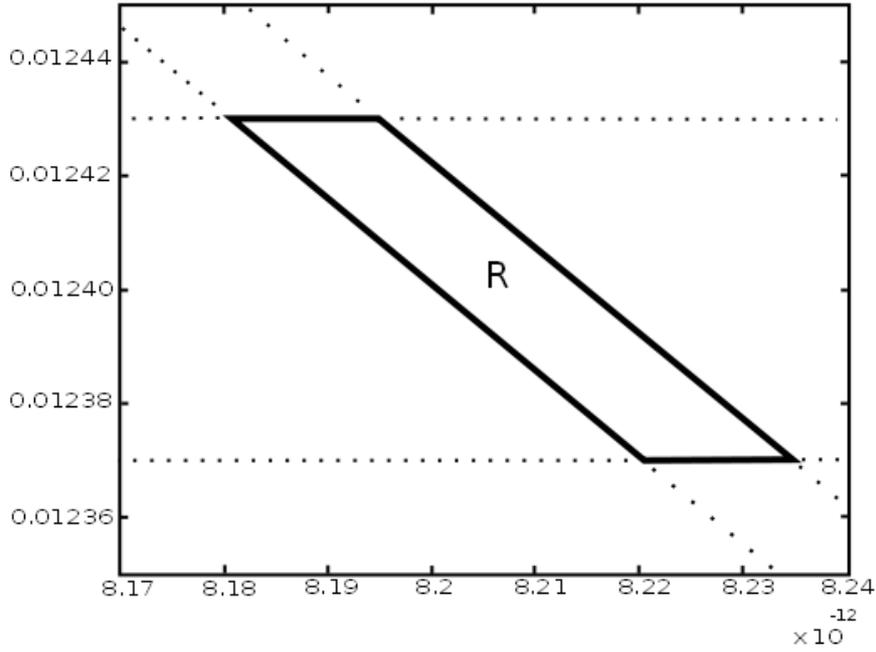


Figura 5.7: Região de integração R

região de integração R , mostrada na figura 5.7, que é delimitada pelas quatro expressões

$$w = 0,01237; \quad w = 0,01243; \quad w = \frac{0,01149 \cdot \epsilon_0}{z}; \quad w = \frac{0,01151 \cdot \epsilon_0}{z}. \quad (5.73)$$

A função marginal $f_Z(z)$ (veja a definição 2.36) é dada pela equação 5.38. As interseções das curvas das expressões 5.73 em z são os pontos

$$\begin{aligned} z_1 &= \frac{0,01237 \cdot \epsilon_0}{0,01151} \\ z_2 &= \frac{0,01243 \cdot \epsilon_0}{0,01151} \\ z_3 &= \frac{0,01237 \cdot \epsilon_0}{0,01149} \\ z_4 &= \frac{0,01243 \cdot \epsilon_0}{0,01149}. \end{aligned} \quad (5.74)$$

Logo, para $z_1 \leq z \leq z_2$,

$$f_Z(z) = \int_{0,01149 \cdot \epsilon_0/z}^{0,01151} f_{Z,W}(z, w) dw = (0,01151)^2 - \frac{(0,01237 \epsilon_0)^2}{z^2} \quad (5.75)$$

para $z_2 < z \leq z_3$,

$$f_Z(z) = \int_{0,01149 \cdot \epsilon_0/z}^{0,01151 \cdot \epsilon_0/z} f_{Z,W}(z, w) dw = \frac{(0,01243 \epsilon_0)^2}{z^2} - \frac{(0,01237 \epsilon_0)^2}{z^2} \quad (5.76)$$

e, para, $z_3 < z \leq z_4$,

$$f_Z(z) = \int_{0,01149}^{0,01151 \cdot \epsilon_0 / z} f_{Z,W}(z, w) dw = \frac{(0,01243\epsilon_0)^2}{z^2} - (0,01149)^2. \quad (5.77)$$

Portanto,

$$f_Z(z) = \begin{cases} (0,01151)^2 - \frac{(0,01237\epsilon_0)^2}{z^2}, & \text{se } z_1 \leq z \leq z_2, \\ \frac{(0,01243\epsilon_0)^2}{z^2} - \frac{(0,01237\epsilon_0)^2}{z^2}, & \text{se } z_2 < z \leq z_3, \\ \frac{(0,01243\epsilon_0)^2}{z^2} - (0,01149)^2, & \text{se } z_3 < z \leq z_4, \\ 0, & \text{caso contrário.} \end{cases} \quad (5.78)$$

Resolvendo a equação $P(C \leq C_M) = 95\%$, obtemos

$$C_M = 8,227188188188188 \times 10^{-12}. \quad (5.79)$$

Se computarmos o valor médio m_C de C através da definição de esperança, obtemos

$$\begin{aligned} m_C &= \int_{0,01237}^{0,01243} \int_{0,01149}^{0,01151} \left(\epsilon_0 \cdot \frac{A}{D} \right) dD dA \\ &= \epsilon_0 \cdot \left[\frac{(0,01237 + 0,01243)}{2 \cdot (0,01151 - 0,01149)} \right] \cdot \log \frac{0,01151}{0,01149} \\ &= 8,207677304286216 \times 10^{-12}, \end{aligned} \quad (5.80)$$

ou seja, se desejamos uma confiança de 95%, devemos tomar como capacitância o valor C_M , que é consideravelmente maior do que a média m_C .

5.3.2 Método do Jacobiano Numérico

Para solucionar o problema através do Método do Jacobiano Numérico, usaremos o algoritmo 5.12 com os seguintes parâmetros:

$$\begin{aligned} d &\sim U(0,01237; 0,01243), \quad A \sim U(0,01149; 0,01151), \\ f_{A,d}(a, d) &= \frac{1}{1,2\epsilon_0 \times 10^{-9}}, \quad N = 4. \end{aligned} \quad (5.81)$$

As figuras 5.8, 5.9 e 5.10 mostram os comparativos entre as funções de densidade obtidas usando o Método do Jacobiano e o Método do Jacobiano Numérico usando o

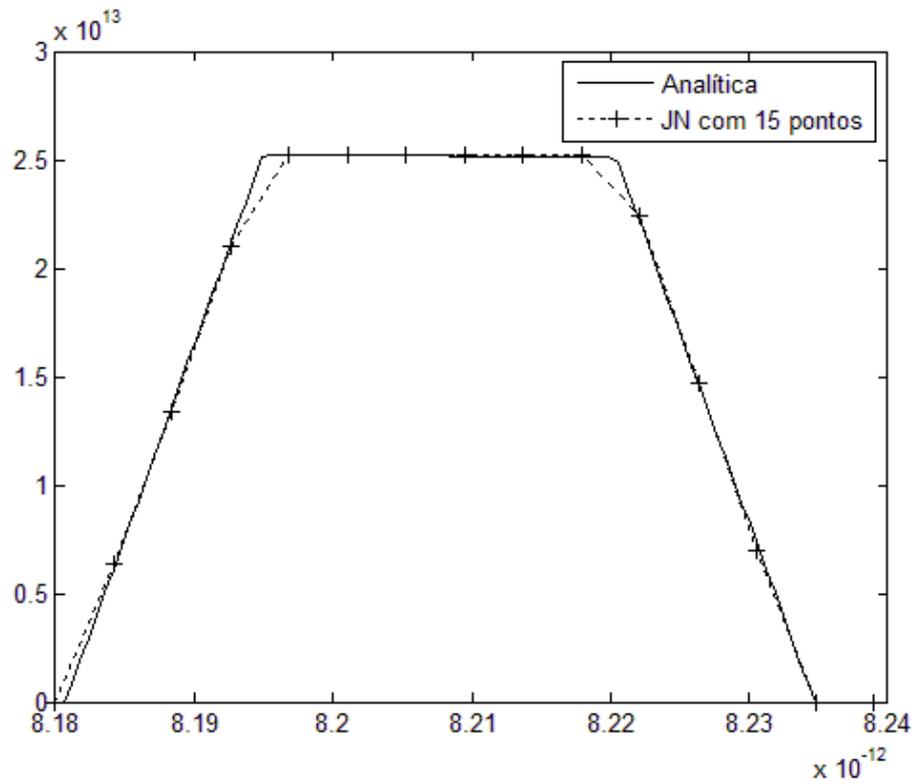


Figura 5.8: Comparativo entre a f_C obtida analiticamente e através do Método do Jacobiano Numérico com 15 pontos

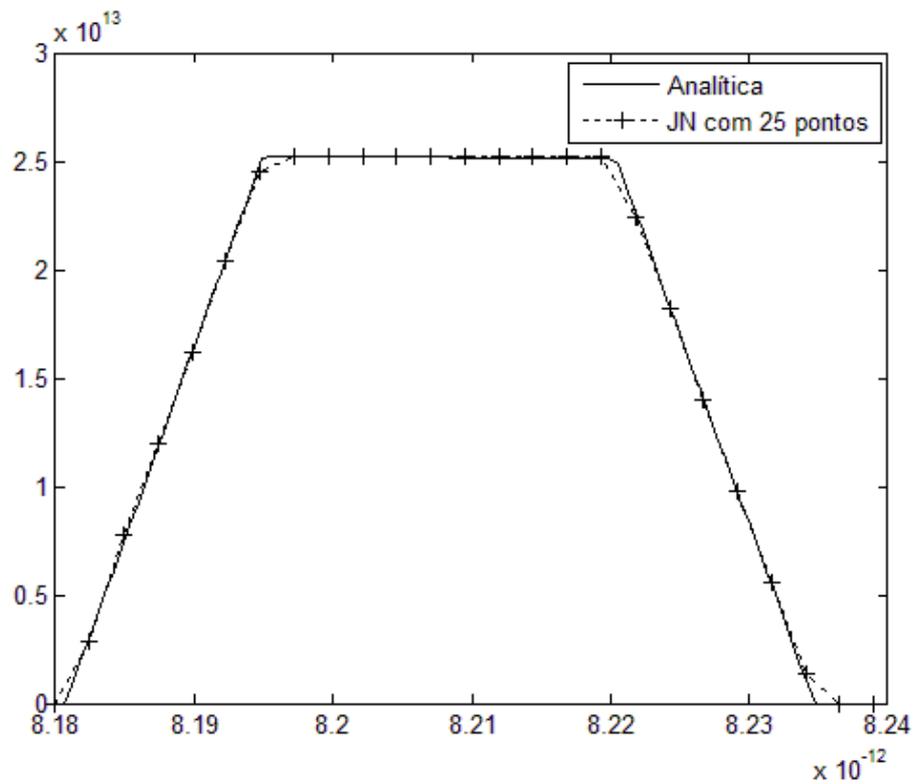


Figura 5.9: Comparativo entre a f_C obtida analiticamente e através do Método do Jacobiano Numérico com 25 pontos

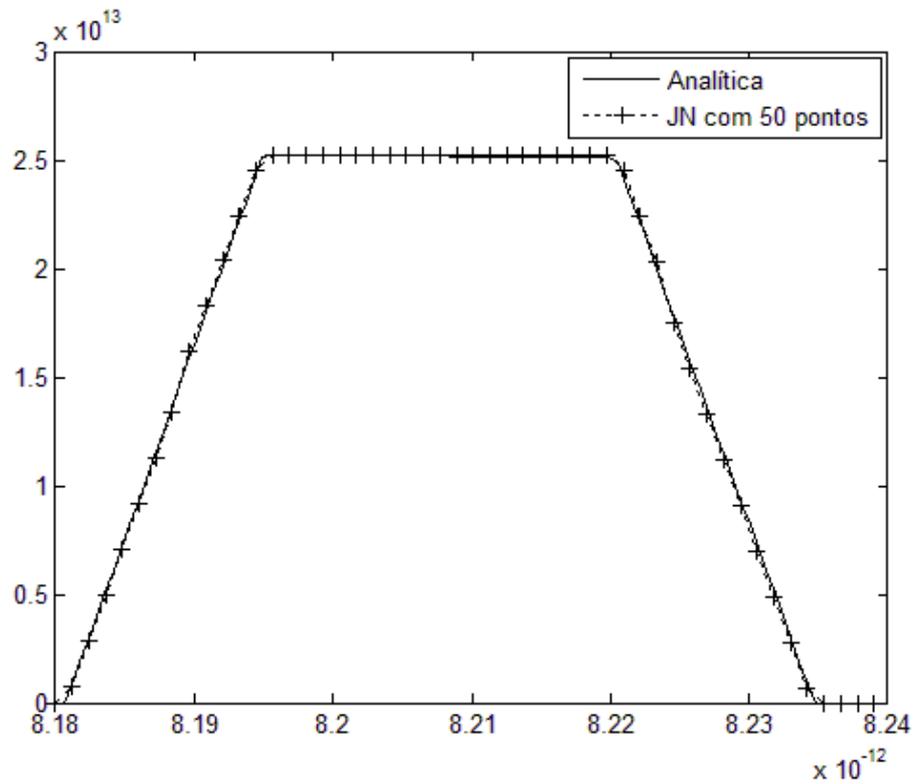


Figura 5.10: Comparativo entre a f_C obtida analiticamente e através do Método do Jacobiano Numérico com 50 pontos

algoritmo 5.12 com $M = 15, 25$ e 50 , respectivamente. A interpolação utilizada nos três casos foi a interpolação linear (ver definição 5.11).

Por fim, o ponto C_M computado usando a aproximação de 50 pontos é igual a

$$C_M = 8,226959183673469 \times 10^{-12}, \quad (5.82)$$

concordando com o resultado 5.79 nas duas primeiras casas decimais.

6 CONCLUSÃO

Ao longo deste trabalho foi evidenciada a importância de se caracterizar as variáveis de entrada de um sistema ou de um experimento através das incertezas associadas às mesmas, e foram observadas, analisadas e estudadas as maneiras como estas incertezas influenciam e alteram as variáveis resultantes.

A abordagem probabilística dos fenômenos naturais, em oposição à determinística, conduz a resultados mais fidedignos, embora seja mais sofisticada e complexa, o que justifica estudos como este.

A principal contribuição deste trabalho foi a formalização teórica da *Unscented Transform*. Esta formalização foi baseada nos rigores da matemática e da probabilidade, e pode ser utilizada para justificar os resultados das formulações anteriores da UT e também como alicerce da *Unscented Transform Gaussiana*, que consiste em uma evolução da UT apresentada neste trabalho, que amplia o escopo e a precisão da técnica original.

Outra importante adição foi a proposta de um algoritmo numérico para o cálculo da função de densidade de probabilidade. O problema de se determinar esta função tem várias dificuldades inerentes, tais como a determinação da região de integração, integrais sem primitivas, cálculo de funções inversas e jacobianos, e o método proposto visa a implementação prática da solução deste problema, contornando tais fatores adversos.

Dada a importância do cálculo dos momentos no algoritmo da *Unscented Transform Gaussiana*, uma contribuição pequena mas significativa foi a apresentação da fórmula geral para o cálculo dos momentos de n -ésima ordem da distribuição normal $N(\mu, \sigma^2)$ (ver exemplo 2.31). Também foi descrita as fórmulas gerais para as distribuições uniforme e exponencial, mas estas são mais recorrentes na literatura por serem mais simples e de fácil dedução.

Dentre os desenvolvimentos futuros a serem realizados a partir deste trabalho podemos destacar: um estudo mais aprofundado sobre a mensuração do erro cometido ao se utilizar a *Unscented Transform Gaussiana*, a tentativa de se deduzir a função de densidade

de probabilidade a partir dos pontos-sigma e dos pesos encontrados através da UTG e a avaliação das consequências da UTG ter desvinculado o número de pontos-sigma e de pesos do número de variáveis independentes do problema, como ocorre no caso da UT.

Por fim, conclui-se que este estudo, ao invés de ser o tratado definitivo sobre o assunto, constitui um ponto de partida para aqueles interessados em se aprofundar no estudo da propagação das incertezas, fornecendo embasamento teórico das técnicas já consagradas e propondo novas abordagens e caminhos a seguir.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] J. K. U. S. J. Julier and H. F. D. Durrant-Whyte, “A new approach for filtering nonlinear systems,” in *The Proceedings of the American Control Conference*, (Seattle, Washington), pp. 1628–1632, 1995.
- [2] B. R. James, *Probabilidade: um curso em nível intermediário*. Projeto Euclides, Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 1996.
- [3] R. Durrett, *Probability: Theory and Examples*. Pacific Coast, Califórnia: Wadsworth and Brooks/Cole, 1991.
- [4] J. L. Doob, *Measure Theory*. N.York: Springer-Verlag.
- [5] C. Christopoulos, *The Transmission-Line Modeling Method TLM*. Oxford: IEE Press.
- [6] E. A. C. Júnior, G. A. Borges, and L. R. Menezes, “Uncertainty propagation in one-dimensional transmission line modeling (tlm) method,” in *Proceedings of 2007 IEEE International Symposium on Antennas and Propagation*, vol. 1, (Honolulu), pp. 1–5, 2007.
- [7] G. A. Borges, E. A. C. Júnior, and L. R. A. X. Menezes, “Algorithms for uncertainty propagation in transmission-line matrix (tlm) method,” *International Journal of Numerical Modelling*, vol. 21, pp. 43–60, 2008.
- [8] M. A. Munem and D. J. Foulis, *Cálculo*. Rio de Janeiro: LTC Editora.
- [9] E. W. Swokowsky, *Cálculo com Geometria Analítica*. São Paulo: Makron Books.
- [10] S. Julier, “The scaled unscented transformation,” 1999.
- [11] J. R. V. Zandt, “A more robust unscented transform.”
- [12] F. B. Hildebrand, *Introduction to Numerical Analysis*. Mineola, Nova York: Dover, 1956.

- [13] T. S. Chihara, *An introduction to Orthogonal Polynomials*. New York: Gordon and Breach.
- [14] R. E. Attar, *Special Functions and Orthogonal Polynomials*. Morrisville NC 27560: Lulu Press.
- [15] M. A. G. Ruggiero, *Cálculo numérico: aspectos teóricos e computacionais*. São Paulo: Makron Books, 1996.
- [16] R. R. David Halliday and J. Walker, *Fundamentos da Física - Volume 3: Eletromagnetismo*. Rio de Janeiro: Editora LTC, 1996.
- [17] L. R. A. X. MENEZES, E. Júnior, M. N. d. SOUSA, and G. A. BORGES, “Estimation of the probability density function in electromagnetic propagation problems with the unscented transform and tlm,” in *The 24th Annual Review of Progress in Applied Computational Electromagnetics*, vol. 1 of *Proceedings of ACES 2008*, (Niagara Falls), pp. 615–620, 2008.

APÊNDICES

A TABELAS DE PONTOS-SIGMA E PESOS

A.1 Distribuição uniforme

Na tabela A.1, N é a precisão da aproximação e m é o número de pontos-sigma x_i e de pesos w_i para uma variável aleatória $X \sim U(0, 1)$.

Tabela A.1: Pontos-sigma e pesos para $X \sim U(0, 1)$

N	m	x_i	w_i
1	1	0,500000000000000	1,000000000000000
3	2	0,78867513459481	0,500000000000000
		0,21132486540519	0,500000000000000
5	3	0,88729833462074	0,277777777777778
		0,499999999999999	0,444444444444445
		0,11270166537926	0,277777777777777
7	4	0,93056815579702	0,17392742256875
		0,66999052179239	0,32607257743131
		0,33000947820751	0,32607257743126
		0,06943184420295	0,17392742256868
9	5	0,95308992296944	0,11846344252783
		0,76923465505333	0,23931433524923
		0,500000000000083	0,284444444444432
		0,23076534494779	0,23931433525014
		0,04691007703082	0,11846344252847
11	6	0,96623475710877	0,08566224617170
		0,83060469326609	0,18038078649336
		0,61930959310306	0,23395696726399
		0,38069040702890	0,23395696729304
		0,16939530681500	0,18038078655911
		0,03376524291062	0,08566224621880

A.2 Distribuição normal

Na tabela A.2, N é a precisão da aproximação e m é o número de pontos-sigma x_i e de pesos w_i para uma variável aleatória $X \sim N(0, 1)$.

Tabela A.2: Pontos-sigma e pesos para $X \sim N(0, 1)$

N	m	x_i	w_i
1	1	0,000000000000000	1,000000000000000
3	2	1,000000000000000	0,500000000000000
		-1,000000000000000	0,500000000000000
5	3	0,000000000000000	0,666666666666667
		1,73205080756888	0,166666666666667
		-1,73205080756888	0,166666666666667
7	4	2,33441421833898	0,04587585476807
		-2,33441421833898	0,04587585476807
		0,74196378430273	0,45412414523193
		-0,74196378430273	0,45412414523193
9	5	0,000000000000000	0,533333333333333
		2,85697001387281	0,01125741132772
		-2,85697001387280	0,01125741132772
		1,35562617997427	0,22207592200561
		-1,35562617997427	0,22207592200561
11	6	3,32425743355212	0,00255578440206
		-3,32425743355212	0,00255578440206
		1,88917587775371	0,08861574604191
		-1,88917587775371	0,08861574604191
		0,61670659019259	0,40882846955603
		-0,61670659019259	0,40882846955603

A.3 Distribuição exponencial

Na tabela A.3, N é a precisão da aproximação e m é o número de pontos-sigma x_i e de pesos w_i para uma variável aleatória $X \sim \text{Exp}(1)$.

Tabela A.3: Pontos-sigma e pesos para $X \sim \text{Exp}(1)$

N	m	x_i	w_i
1	1	1,000000000000000	1,000000000000000
3	2	3,41421356237309	0,14644660940673
		0,58578643762690	0,85355339059327
5	3	6,28994508293748	0,01038925650159
		2,29428036027904	0,27851773356924
		0,41577455678348	0,71109300992917
7	4	9,39507091230116	0,00053929470556
		4,53662029692115	0,03888790851500
		1,74576110115837	0,35741869243780
		0,32254768961940	0,60315410434164
9	5	12,64080084427579	0,00002336997239
		7,08581000585886	0,00361175867992
		3,59642577104074	0,07594244968171
		1,41340305910652	0,39866681108317
		0,26356031971814	0,52175561058281
11	6	15,98287398060219	0,00000089854791
		9,83746741838317	0,00026101720281
		5,77514356910473	0,01039919745315
		2,99273632605951	0,11337338207403
		1,18893210167270	0,41700083077211
		0,22284660417928	0,45896467394999