UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PÓS-GRADUAÇÃO EM ESTATÍSTICA

**RAPHAELA LIMA BELCHIOR DE ARAÚJO**

# FAMÍLIA COMPOSTA POISSON-TRUNCADA:
## Propriedades e Aplicações

DISSERTAÇÃO DE MESTRADO


Recife
2015

**RAPHAELA LIMA BELCHIOR DE ARAÚJO**

# FAMÍLIA COMPOSTA POISSON-TRUNCADA:
## Propriedades e Aplicações

Dissertação apresentada ao programa de Pós-graduação em Estatística da Universidade Federal de Pernambuco como requisito parcial para obtenção do título de **Mestre em Estatística**.

**Área de Concentração:** Probabilidade

**ORIENTADOR:** Prof. Ph.D. Leandro Chaves Rêgo

**CO-ORIENTADOR:** Prof. Dr. Abraão David Costa Do Nascimento

Recife
2015

Catalogação na fonte
Bibliotecária Joana D'Arc Leão Salvador CRB4-572

A663f Araújo, Raphaela Lima Belchior de.
Família composta Poisson-Truncada: propriedades e aplicações / Raphaela Lima Belchior de Araújo. – Recife: O Autor, 2015.
74 f.: fig., tab.

Orientador: Leandro Chaves Rêgo.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Estatística, 2015.
Inclui referências e apêndice.

1. Probabilidades. 2. Distribuição (Teoria da probabilidade). 3. Métodos de estimação. I. Rêgo, Leandro Chaves (Orientador). II. Titulo.

519.2 CDD (22. ed.) UFPE-MEI 2015-116

**RAPHAELA LIMA BELCHIOR DE ARAÚJO**

**FAMÍLIA COMPOSTA POISSON-TRUNCADA: PROPRIEDADES E APLICAÇÕES**

Dissertação apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Estatística.

Aprovada em: 31 de julho de 2015.

**BANCA EXAMINADORA**

---

Prof. Ph.D. Leandro Chaves Rêgo
UFPE

---

Prof. Dr. Francisco José de Azevêdo Cysneiros (Examinador Interno)
UFPE

---

Prof. Dr. Juvêncio Santos Nobre (Examinador Externo)
UFC

# Agradecimentos

Inicio meus agradecimentos por Deus. Quem melhor para se espelhar do que no Mestre dos mestres?! Agradeço por sentir Tua presença sempre em minha vida.

Agradeço a minha mãe, pelo amor, esforço e dedicação. Sem a perseverança de Dona Irene, em oferecer uma vida melhor do que a que teve, não estaria realizando mais essa conquista. Mãe, obrigada mesmo.

À minha família, pelo apoio em todos esses anos. E, em especial, à minha irmã Juliana. Não consigo computar quantas discussões tivemos, nem é preciso; o importante é que não consigo imaginar minha vida sem você. Agradeço por permitir, nem sei ao certo se você teve escolha, ser minha primeira aluna. Foi um chute inicial.

Ao meu esposo Daniel. Até hoje tentamos saber como nosso casamento aconteceu em meio ao mestrado... mas, tenho certeza de que sua paciência, fé e amor fizeram toda a diferença. Obrigada por ser meu companheiro, por acreditar e nunca desistir de mim.

Agradeço a todos os amigos. Em particular, ao Fábio. Pela persistência e tão boa vontade. Obrigada pela confiança e por aguentar a sessão de dúvidas desde a graduação.

Agradeço a todos os professores que contribuíram para a minha formação. Ressalto a admiração que tenho pelos meus orientadores, os professores Leandro e Abraão. A esses, meus sinceros agradecimentos pela orientação, apoio e tempo dedicado a esse trabalho, sem contar a paciência com tantas dúvidas "descabidas". Aqui, gostaria de lembrar dos professores desse Programa de Pós-graduação que também tive o prazer de conhecer e aprender: Francisco Cribari, Francisco Cysneiros, Gauss Cordeiro e Patrícia Ospina.

À banca examinadora, pelas observações e contribuições dadas ao trabalho.

Aos funcionários Valéria Bittencourt e Lódino Serbim. Agradeço por toda a atenção.

Agradeço aos colegas e amigos, conquistados nesse período, pelos momentos de descontração e de estudos. Parece impossível, mas experimentamos que é possível. Espero que a distância não quebre o laço que formamos. Da minha entrada, restaram 5 (comigo), guardarei lembranças: do Sébastien, pelas trocas de conhecimento em "portunhol"; da Wanessa, que não aguenta ficar sem comentar; da Evelyne, por nos lembrar a importância de "o porquê das coisas"; da Laura, que precisou ser "adotada" para se sentir em casa. Obrigada pelas inúmeras ajudas. Meninas, brincadeiras a parte, nosso quarteto passou por poucas e boas... Que nossa amizade dure por muito tempo... Já sinto saudades!

Finalmente, gostaria de agradecer à Universidade Federal de Pernambuco pela oportunidade de ingressar no seu Programa de Pós-graduação em Estatística, e também à CAPES pelo apoio financeiro. Sem esse elo, esse meu sonho não seria concretizado.

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

# Resumo

Este trabalho analisa propriedades da família de distribuições de probabilidade Composta $N$ e propõe a sub-família Composta Poisson-Truncada como um meio de compor distribuições de probabilidade. Suas propriedades foram estudadas e uma nova distribuição foi investigada: a distribuição Composta Poisson-Truncada Normal. Esta distribuição possui três parâmetros e tem uma flexibilidade para modelar dados multimodais. Demonstramos que sua densidade é dada por uma mistura infinita de densidades normais em que os pesos são dados pela função de massa de probabilidade da Poisson-Truncada. Dentre as propriedades exploradas desta distribuição estão a função característica e expressões para o cálculo dos momentos. Foram analisados três métodos de estimação para os parâmetros da distribuição Composta Poisson-Truncada Normal, sendo eles, o método dos momentos, o da função característica empírica (FCE) e o método de máxima verossimilhança (MV) via algoritmo EM. Simulações comparando estes três métodos foram realizadas e, por fim, para ilustrar o potencial da distribuição proposta, resultados numéricos com modelagem de dados reais são apresentados.

**Palavras-chave:** Família Composta $N$. Família Composta Poisson-Truncada. Distribuição Composta Poisson-Truncada Normal. Estimação pelo Método dos Momentos. Estimação por Função Característica Empírica. Algoritmo EM.

# Abstract

This work analyzes properties of the Compound $N$ family of probability distributions and proposes the sub-family Compound Poisson-Truncated as a means of composing probability distributions. Its properties were studied and a new distribution was investigated: the Compound Poisson-Truncated Normal distribution. This distribution has three parameters and has the flexibility to model multimodal data. We demonstrated that its density is given by an infinite mixture of normal densities where in the weights are given by the Poisson-Truncated probability mass function. Among the explored properties of this distribution are the characteristic function end expressions for the calculation of moments. Three estimation methods were analyzed for the parameters of the Compound Poisson-Truncated Normal distribution, namely, the method of moments, the empirical characteristic function (ECF) and the method of maximum likelihood (ML) by EM algorithm. Simulations comparing these three methods were performed and, finally, to illustrate the potential of the proposed distribution numerical results with real data modeling are presented.

**Keywords:** Compound $N$ Family. Compound Poisson-Truncated Family. Compound Poisson-Truncated Normal Distribution. Estimation by Method of Moments. Estimation by Empirical Characteristic Function. EM Algorithm.

# Lista de Figuras

2.1 Funções de distribuição acumulada da Composta Poisson-Truncada Normal 18
2.2 Funções densidade da distribuição Composta Poisson-Truncada Normal . . 20
2.3 Função taxa de falha da Composta Poisson-Truncada Normal . . . . . 22
2.4 Entropia de Shannon da Composta Poisson-Truncada Normal . . . . . 29

3.1 Densidades teórica e ajustadas pelos métodos FCE e MV via EM. . . . . 43
3.2 Densidades ajustadas aos dados. . . . . 60

# Lista de Tabelas


3.1 Total de amostras com estimativas pelos métodos FCE e MV via EM na simulação. . . . . 41
3.2 Comparação das estimações para o método FCE e MV via EM. . . . . 41
3.3 Ajustes dos métodos de estimação na amostra $N = 100$. . . . . 42
3.4 Ajustes dos métodos de estimação na amostra $N = 150$. . . . . 42
3.5 Total de amostras sem estimativas pelo MM na simulação Monte Carlo. . . 45
3.6 Simulações para os métodos dos momentos e MV via EM variando o parâmetro $\lambda$. . . . . 48
3.7 Simulações para os métodos dos momentos e MV via EM variando o parâmetro $\mu$. . . . . 49
3.8 Simulações para os métodos dos momentos e MV via EM variando o parâmetro $\sigma$. . . . . 50
3.9 Simulações para os métodos dos momentos e MV via EM variando o parâmetro $\lambda$. . . . . 53
3.10 Simulações para os métodos dos momentos e MV via EM variando o parâmetro $\mu$. . . . . 54
3.11 Simulações para os métodos dos momentos e MV via EM variando o parâmetro $\sigma$. . . . . 55
3.12 Conjunto de dados de Foulum. . . . . 57
3.13 Descrição do conjunto de dados de Foulum. . . . . 57
3.14 Parâmetros estimados ao conjunto de dados de Foulum. . . . . 59
3.15 Bondade de ajuste. . . . . 60

# Sumário

**1 Introdução** **13**

1.1 Referencial teórico . . . . . 13

1.2 Contribuições . . . . . 15

1.3 Organização da dissertação . . . . . 16

**2 Família Composta N** **17**

2.1 Apresentação da Família . . . . . 17

2.1.1 Sub-família Composta Poisson-Truncada . . . . . 17

2.2 Propriedades . . . . . 19

2.2.1 Densidade . . . . . 19

2.2.2 Função de risco . . . . . 20

2.2.3 Função característica e função geradora de momentos . . . . . 22

2.2.4 Momentos . . . . . 25

2.2.5 Entropia de Shannon . . . . . 28

**3 Métodos de estimação e resultados numéricos** **30**

3.1 Método dos momentos . . . . . 30

3.2 Método por função característica empírica . . . . . 32

3.3 Método de máxima verossimilhança . . . . . 33

3.3.1 Função escore . . . . . 34

3.3.2 Via algoritmo EM . . . . . 37

3.4 Estudo de casos: FCE versus MV via algoritmo EM . . . . . 40

3.5 Simulações . . . . . 44

3.5.1 Análise da viabilidade do método dos momentos . . . . . 45

3.5.2 Pequenas amostras . . . . . 45

3.5.3 Grandes amostras . . . . . 51
3.6 Aplicação . . . . . 56

**4 Conclusões e sugestões para trabalhos futuros** **61**
4.1 Conclusões . . . . . 61
4.2 Sugestões para trabalhos futuros . . . . . 62

**Referências** **63**

**A Apêndice** **66**
A.1 Cálculo dos momentos através da função característica . . . . . 66
A.2 Cálculo para a estimação dos parâmetros por FCE . . . . . 71

# CAPÍTULO 1

# Introdução

## 1.1 Referencial teórico

Segundo Cobb (1978), distribuições empíricas multimodais consistem em uma realidade em muitas aplicações. Estes dados requerem grande flexibilidade do modelo probabilístico adotado, o que destoa da proposta de grande parte das distribuições clássicas. Na prática, o problema costuma ser resolvido através do uso de misturas de distribuições de probabilidade, entretanto este suposto se confronta com a questão de parcimônia (BARNETT, 1999). Nesta dissertação, objetiva-se apresentar uma solução parcimoniosa à descrição de distribuições empíricas multimodais.

Aplicações em reconhecimento de voz e em processamento de imagens médicas são cenários cujos dados resultantes são multimodais. Adicionalmente, aplicações em imagens de radar também carecem de modelos para descrever dados multimodais (EL-ZAART; ZIOU, 2007).

Antes de definir elementos da nossa proposta, faz-se necessário apresentar uma discussão rápida sobre mistura de distribuições. Uma mistura de distribuições é representada por uma variável aleatória em que tanto sua função densidade de probabilidade (fdp) como sua função de distribuição acumulada (fda) são definidas como combinações lineares convexas de fdps ou de fdas, respectivamente. Seja $X_i$ uma variável aleatória com fda $F_i(x) = F(x;\theta_i)$ e fdp ou função de massa de probabilidade (fmp) $f_i(x) = f(x,\theta_i)$, em que $\theta_i$ representa um ponto do espaço paramétrico $\Theta$. Uma mistura finita de $M$ variáveis aleatórias $X_1,\ldots,X_M$ tem fda e fdp (ou fmp) dadas, respectivamente, por:

$$F(x;\delta) = \sum_{i=1}^{M} \pi_i F_i(x) \quad \text{e} \quad f(x;\delta) = \sum_{i=1}^{M} \pi_i f_i(x), \tag{1.1}$$

com $\delta = (\theta_1^\top,\ldots,\theta_M^\top)^\top$, em que $\theta_i$ e $\theta_j$ são elementos distintos de $\Theta$, $\sum_{i=1}^{M} \pi_i = 1$ e $\pi_i > 0$.

As distribuições combinadas para obter a mistura são chamadas componentes, e as probabilidades, $\pi_i$ para $i = 1, \ldots, M$, associadas a elas são denominadas de pesos da mistura. Muitas vezes, a quantidade dessas componentes é finita; não excluindo a possibilidade de um número infinito de componentes. Wirjanto e Xu (2009) discutiram sobre aplicações de misturas da família normal. Everitt *et al.* (1981) mostraram evidências de que a mistura de normais também pode ser aplicada para investigar a robustez de certas técnicas estatísticas, quando os dados não estão normalmente distribuídos.

Como um gerador para novos modelos, uma distribuição de probabilidade composta é o resultado de assumir que uma variável aleatória (v.a.) leve em conta outros componentes aleatórios na sua definição: como exemplos, (a) uma variável definida como o mínimo, o máximo ou a soma de outras $N$ variáveis tal que (como outro componente aleatório) $N$ assume uma distribuição discreta e estritamente positiva e (b) uma variável tal que, no mínimo, um de seus parâmetros seja também aleatório. Vários trabalhos têm contribuído nesta frente. Grushka (1972) e Golubev (2010) apresentaram estudos sobre a distribuição Normal exponencialmente modificada. Esta distribuição é o resultado de compor a distribuição Normal com a média distribuída de acordo com uma distribuição Exponencial modificada. Compondo uma distribuição Poisson com o parâmetro distribuído conforme uma distribuição Gama, produz-se uma distribuição Binomial Negativa (TEICH; DIAMENT, 1989).

Karlis e Xekalaki (2005) abordaram um tipo importante dessas composições, a denominada distribuição Composta $N$, em que uma variável aleatória $S = X_1 + X_2 + \cdots + X_N$ é constituída a partir de um número $N$ de variáveis aleatórias subjacentes, sendo que $N$ também é uma variável aleatória. Um caso mencionado, nesse artigo, é a distribuição Composta Poisson. Uma variável aleatória que segue essa distribuição representa a soma de um conjunto de variáveis aleatórias independentes identicamente distribuídas segundo uma dada distribuição, e que o número de elementos desse conjunto é uma v.a. $N$ que segue uma distribuição Poisson e é independente das variáveis a serem somadas.

Distribuições geradas da Composta Poisson são usadas em aplicações cotidianas. A Composta Poisson Exponencial foi aplicada por Revfeim (1984) para modelar a distribuição da precipitação total em um dia, em que cada dia contém um número com uma distribuição Poisson de eventos, cada um dos quais fornece uma quantidade de precipitação que tem uma distribuição Exponencial. Thompson (1984) aplicou o mesmo modelo para dados de chuvas totais mensais. Panjer (1981) apontou que distribuições compostas, tais como a Composta Poisson e a Composta Binomial Negativa são usadas extensivamente na teoria do risco.

## 1.2 Contribuições

Um dos objetivos desta dissertação é estudar propriedades estatísticas da família de distribuições denominada *Composta N*. Supondo uma variável aleatória $S = \sum_{j=1}^{N} X_j$ que segue essa distribuição, e particularizando para o caso em que $N$ é uma v.a. com distribuição Poisson truncada no zero, propomos uma forma de gerar outras distribuições de probabilidade, sugerindo deste modo, a sub-família *Composta Poisson-Truncada*. Que ao contrário da distribuição Composta Poisson, que sempre possui probabilidade positiva no zero, a distribuição Composta Poisson-Truncada é contínua, se as variáveis aleatórias $X$'s tiverem distribuição contínua. Motivando o seu uso, demonstramos algumas propriedades matemáticas, aproveitando para introduzir a distribuição *Composta Poisson-Truncada Normal* (CPTN), em que os $X_j$'s são variáveis aleatórias independentes com distribuição Normal.

Patil (1964), com uma nomenclatura similar, apresentou uma variável aleatória que segue a distribuição Composta Poisson Normal-Truncada. Entretanto, resulta de uma variável aleatória com distribuição Composta Poisson que possui como média, uma v.a. $\theta$, distribuída por uma Normal truncada à esquerda do zero. A nova distribuição CPTN resulta numa soma de distribuições normais, nos remetendo a um modelo de "Misturas de Normais", o que permite uma grande flexibilidade na captura de várias formas de densidade; no entanto, esta mesma flexibilidade também leva a alguns problemas de estimação.

No Capítulo de estimação dos parâmetros, apresentamos três métodos: O método dos momentos, o método da função característica empírica e o método de máxima verossimilhança. Introduzido na literatura por Karl Pearson, o primeiro método é considerado um dos mais antigos para se encontrar um estimador, de acordo com Bolfarine e Sandoval (2010). Em muitos casos, trata-se de um bom ponto de partida, pois esse método produz estimadores consistentes.

Quandt e Ramsey (1978) utilizaram o método da máxima verossimilhança (MV) para estimar os parâmetros do modelo Misturas de Normais (MN). Porém, em muitas circunstâncias, a maximização da função de verossimilhança não tem forma fechada. Observando esta dificuldade, recorremos ao método da FCE (Função Característica Empírica). Segundo Carrasco *et al.* (2000), no caso de estimação de misturas, a estimação via função característica é uma boa alternativa para o MV, pois a FCE produz um estimador eficiente quando utilizado com uma função de ponderação específica, isso foi demonstrado por Feuerverger e McDunnough (1981). Dessa forma, aplicamos o método na distribuição CPTN, mas fazendo uso da função peso proposta para a distribuição Normal em Yu (2004).

Como outra alternativa, propomos o método de máxima verossimilhança via algoritmo EM. O método permite estimar os parâmetros na presença de dados latentes; a ideia consiste em ampliar os dados observados com dados não observados, com a finalidade de facilitar os cálculos. No caso da distribuição CPTN, os valores não observados são os valores de $N$.

## 1.3 Organização da dissertação

Esta dissertação está estruturada da seguinte forma. No Capítulo 2, desenvolvemos propriedades estatísticas da família de distribuições Composta $N$, a saber, a função densidade de probabilidade, a função de risco, a função característica e a função geradora de momentos, expressões para o cálculo dos momentos e a entropia de Shannon. No Capítulo 3, apresentamos três métodos de estimação para a distribuição Composta Poisson-Truncada Normal: Método dos momentos (CASELLA; BERGER, 2002), estimação por função característica empírica (FCE) (ver: Press (1972), Paulson *et al.* (1975), Heathcote (1977), Yu (2004)) e estimação por máxima verossimilhança via algoritmo EM (DEMPSTER *et al.*, 1977). Em seguida, fizemos alguns estudos de casos comparando a eficiência dos métodos FCE e máxima verossimilhança via algoritmo EM. Como os resultados do método FCE foram computacionalmente ineficientes, não o utilizamos em um estudo de simulação Monte Carlo subsequente, em que comparamos as estimativas obtidas pelo método dos momentos com as obtidas pelo método de máxima verossimilhança via algoritmo EM. Finalmente, concluímos o Capítulo 3, apresentando uma aplicação da distribuição CPTN na modelagem de um banco de dados de imagem SAR (Synthetic ApertureRadar - SAR) extraída da imagem de Foulum (Dinamarca). Por fim, no Capítulo 4, as conclusões e propostas de trabalhos futuros são apresentadas.

# CAPÍTULO 2

## Família Composta N: Propriedades e aplicação ao caso da Distribuição Composta Poisson-Truncada Normal

## 2.1 Apresentação da Família

Sejam $X_1, \ldots, X_n$ uma amostra aleatória (a.a.) (entenda-se a.a. como uma amostra independente e identicamente distribuída (i.i.d.)) de $X \sim G$, em que $G$ é uma função de distribuição acumulada qualquer, e $n$ uma possível realização de uma variável aleatória $N \in \mathbb{Z}_+$. A variável $S = \sum_{j=1}^{N} X_j$ é denotada como "Composta N" e tem fda dada por

$$\begin{aligned} F_S(x) = P(S \leqslant x) &= P\left(\sum_{j=1}^{N} X_j \leqslant x\right) \\ &= \sum_{k=1}^{\infty} P(N=k) P\left(\sum_{j=1}^{N} X_j \leqslant x | N=k\right) \\ &= \sum_{k=1}^{\infty} P(N=k) P\left(\sum_{j=1}^{k} X_j \leqslant x\right) \\ &= \sum_{k=1}^{\infty} P(N=k)\, F_{S_k}(x), \end{aligned} \tag{2.1}$$

em que $F_{S_k}(x)$ é a fda da soma de $k$ variáveis aleatórias independentes com distribuição $G$, ou seja, $S_k = X_1 + \ldots + X_k$.

### 2.1.1 Sub-família Composta Poisson-Truncada

Seja $N$ uma variável aleatória (v.a.) com distribuição Poisson truncada no zero, com parâmetro $\lambda > 0$, $N \sim PT(\lambda)$. Consideremos as propriedades da família decorrente de

assumir $N \sim PT(\lambda)$, denotada por *Composta Poisson-Truncada* com a fda dada por:

$$F_S(x) = \sum_{k=1}^{\infty} P(N=k)\, F_{S_k}(x) = \sum_{k=1}^{\infty} \left(\frac{1}{e^{\lambda}-1}\right) \frac{\lambda^k}{k!}\, F_{S_k}(x)$$
$$= \left(\frac{1}{e^{\lambda}-1}\right) \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} F_{S_k}(x). \tag{2.2}$$

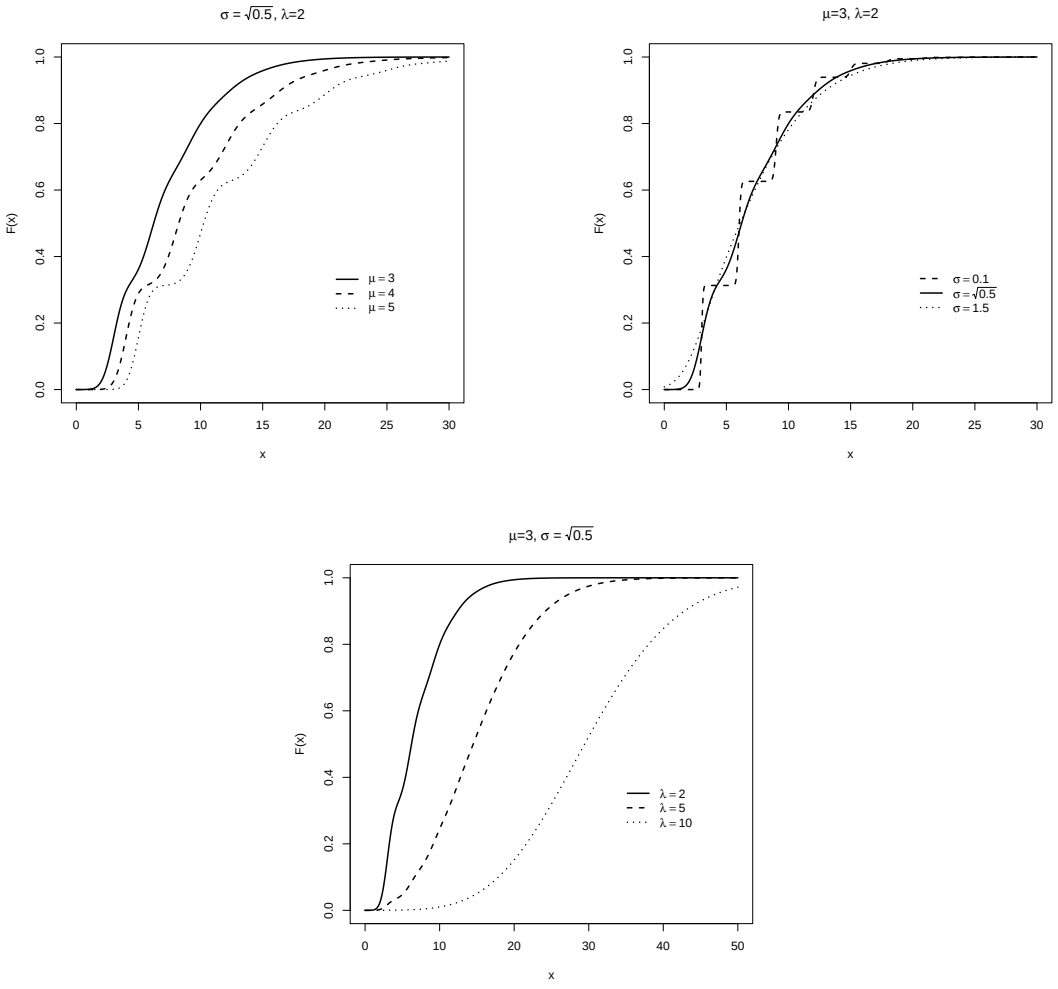
**Exemplo 1.** Um caso especial de (2.2) decorrente de $G \sim N(\mu, \sigma^2)$ é a distribuição Composta Poisson-Truncada Normal (CPTN). Este modelo será utilizado nos estudos de simulação e aplicação desta dissertação. A fda da CPTN é dada por:

$$F_S(x) = \left(\frac{1}{e^{\lambda}-1}\right) \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \cdot \Phi\left(\frac{x-k\mu}{\sigma\sqrt{k}}\right),$$

em que $\Phi(\cdot)$ é a fda da $N(0,1)$, $\lambda > 0$, $\mu \in \mathbb{R}$ e $\sigma > 0$. Nesse caso, usa-se a notação $S \sim CPTN(\lambda, \mu, \sigma^2)$.

A Figura 2.1 apresenta o comportamento da fda da distribuição Composta Poisson-Truncada Normal para diferentes valores de parâmetros.

**Figura 2.1:** Funções de distribuição acumulada da Composta Poisson-Truncada Normal



## 2.2 Propriedades

Nesta seção, apresentamos e demonstramos algumas propriedades da família e sub-família em análise. A cada propriedade da sub-família Composta Poisson-Truncada, considere $N$ uma v.a. discreta com distribuição Poisson truncada no zero.

### 2.2.1 Densidade

Se a variável aleatória $X$ for absolutamente contínua, a função densidade de probabilidade (fdp) da família Composta $N$ é dada por:

$$f_S(x) = \sum_{k=1}^{\infty} P(N = k)\, f_{S_k}(x), \tag{2.3}$$

em que $f_{S_k}(x)$ é a densidade da soma de $k$ variáveis aleatórias independentes com distribuição igual a de $X$.

#### Sub-família Composta Poisson-Truncada

Dessa forma, a função densidade da sub-família Composta Poisson-Truncada pode ser expressa por

$$\begin{aligned} f_S(x) &= \sum_{k=1}^{\infty} P(N = k)\, f_{S_k}(x) = \sum_{k=1}^{\infty} \left(\frac{1}{e^{\lambda}-1}\right) \frac{\lambda^k}{k!}\, f_{S_k}(x) \\ &= \left(\frac{1}{e^{\lambda}-1}\right) \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} f_{S_k}(x), \quad \text{com } \lambda > 0. \end{aligned}$$

**Exemplo 2.** Sendo $X$ uma variável aleatória com distribuição Normal $(\mu, \sigma^2)$, podemos representar a fdp da distribuição $CPTN(\lambda, \mu, \sigma^2)$ por:

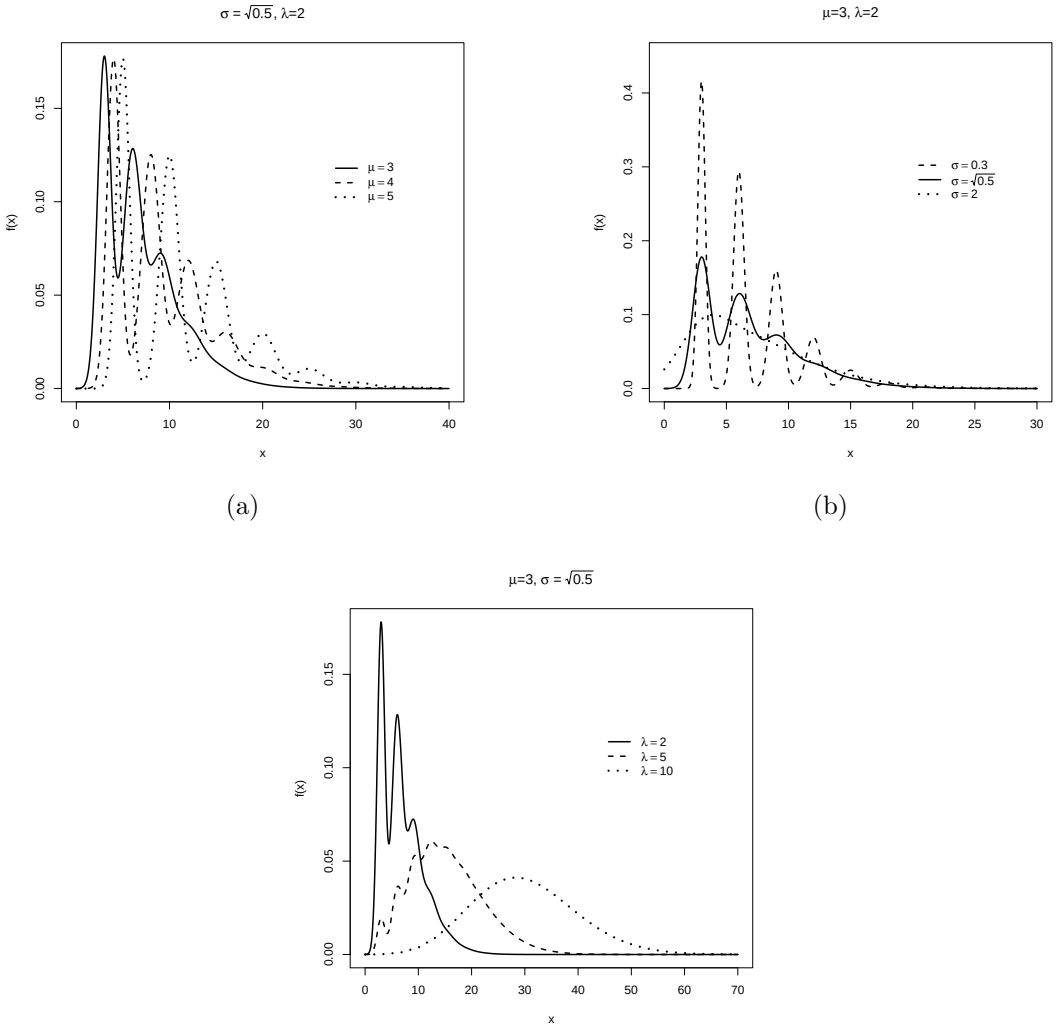
$$\begin{aligned} f_S(x) &= \sum_{k=1}^{\infty} \frac{\lambda^k}{k!(e^{\lambda}-1)} f_{N(k\mu, k\sigma^2)}(x) \\ &= \left(\frac{1}{e^{\lambda}-1}\right) \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} (2\pi k\sigma^2)^{-1/2} \exp\left[-\frac{1}{2}\frac{(x-k\mu)^2}{k\sigma^2}\right], \end{aligned} \tag{2.4}$$

em que $x \in \mathbb{R}$, $\lambda > 0$, $\mu \in \mathbb{R}$ e $\sigma > 0$.

A Figura 2.2 ilustra diferentes formas da fdp da distribuição CPTN. No gráfico 2.2(a), fixando os parâmetros $\lambda = 2$ e $\sigma = \sqrt{0.5}$ e adotando alguns valores para $\mu$, encontramos densidades multimodais. Quando fixamos $\lambda = 2$ e $\mu = 3$, no gráfico 2.2(b), notamos uma mudança de forma e no número de modas da CPTN. Admitindo diferentes valores para o parâmetro $\lambda$, note que à medida que ele cresce as caudas se tornam mais pesadas. Desse modo, um caso em que a densidade é unimodal é obtido quando $\lambda = 10$ em 2.2(c). Observe que o número de modas, presentes em cada gráfico, aumenta quando o valor de $\mu$ cresce em

módulo, ou o de $\sigma$ decresce ou quando o valor de $\lambda$ decresce. Outro fato interessante, é que analisando o gráfico 2.2(b), por exemplo, pode ser conferido que a distância entre modas é igual ao valor do parâmetro $\mu$. Veja que para o caso $\sigma = 0.3$, a densidade apresenta 5 modas no intervalo $x \in [0, 15]$, em que $\mu = 3$.

**Figura 2.2:** Funções densidade da distribuição Composta Poisson-Truncada Normal




(a) (b)

(c)

### 2.2.2 Função de risco

Em análise de sobrevivência, a função de risco ou função taxa de falha (ftf), como também é conhecida, descreve como a taxa instantânea de falhas varia com o tempo. Mais informações podem ser encontradas em Lawless (2003). A função taxa de falha da família

Composta $N$ é definida por

$$h_S(x) = \frac{f_S(x)}{1-F_S(x)} = \frac{\sum_{k=1}^{\infty} P(N=k)\, f_{S_k}(x)}{1-\sum_{k=1}^{\infty} P(N=k)\, F_{S_k}(x)} = \frac{\sum_{k=1}^{\infty} P(N=k)\, h_{S_k}(x)(1-F_{S_k}(x))}{\sum_{k=1}^{\infty} P(N=k)\,(1-F_{S_k}(x))}.$$

Essencialmente, sabe-se que a função taxa de falha costuma ser associada a variáveis aleatórias estritamente positivas. No nosso caso, embora a distribuição CPTN tenha suporte no conjunto dos números reais, a Figura 2.3 apresenta casos para os quais a probabilidade do evento $[X < 0]$ é negligível e, portanto, parece razoável analisar as taxas de risco associadas.

**Sub-família Composta Poisson-Truncada**

Substituindo no resultado anterior a fmp da distribuição Poisson truncada no zero, temos a seguinte função de risco:

$$\begin{aligned} h_S(x) &= \frac{f_S(x)}{1-F_S(x)} = \frac{\sum_{k=1}^{\infty} P(N=k)\, h_{S_k}(x)(1-F_{S_k}(x))}{\sum_{k=1}^{\infty} P(N=k)\,(1-F_{S_k}(x))} \\ &= \frac{\left(\frac{1}{e^{\lambda}-1}\right)\sum_{k=1}^{\infty} \frac{\lambda^k}{k!}\cdot h_{S_k}(x)(1-F_{S_k}(x))}{\left(\frac{1}{e^{\lambda}-1}\right)\sum_{k=1}^{\infty} \frac{\lambda^k}{k!}\cdot (1-F_{S_k}(x))} \\ &= \frac{\sum_{k=1}^{\infty} \frac{\lambda^k}{k!}\cdot h_{S_k}(x)(1-F_{S_k}(x))}{\sum_{k=1}^{\infty} \frac{\lambda^k}{k!}\cdot (1-F_{S_k}(x))}. \end{aligned}$$

**Exemplo 3.** Suponha que $S$ é uma variável aleatória com distribuição $CPTN(\lambda, \mu, \sigma^2)$ e fdp dada em (2.4). A função de risco de $S$ é dada por:

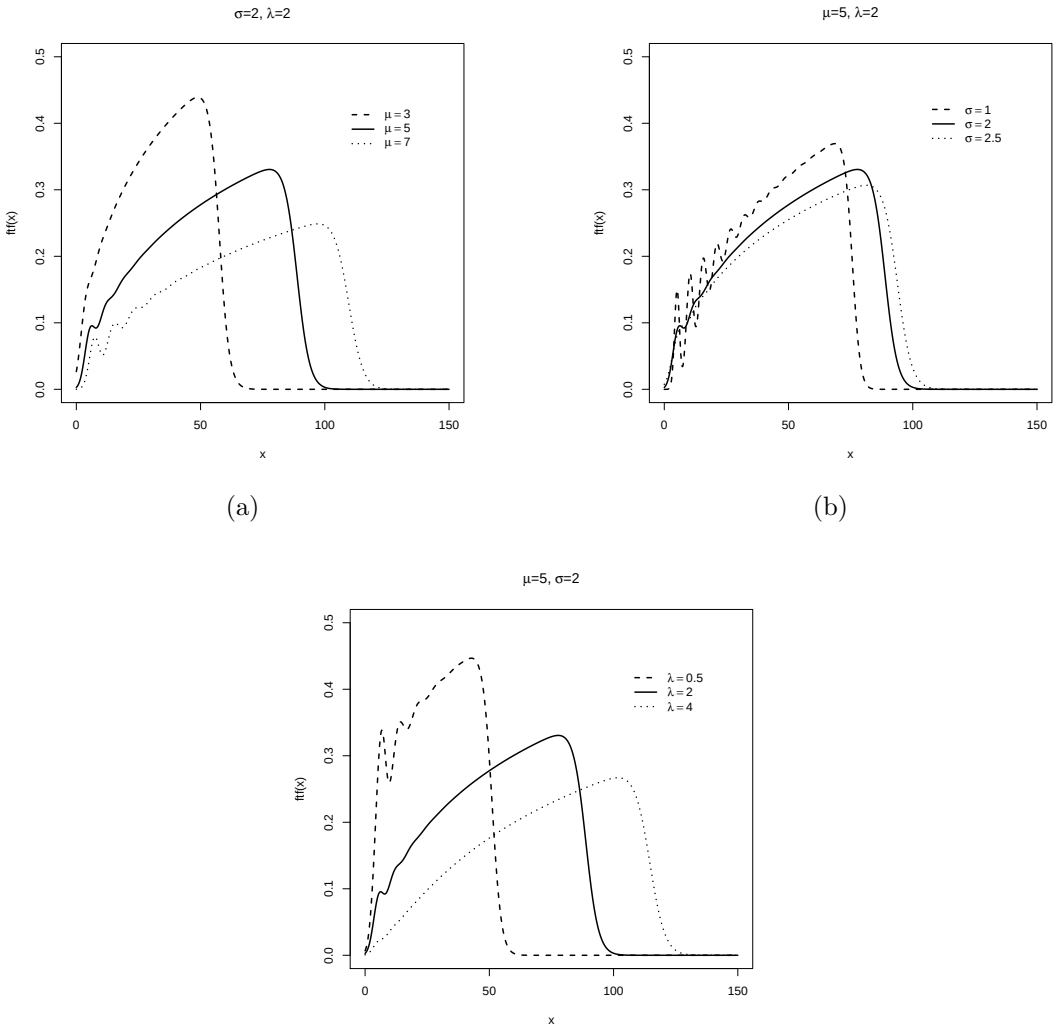
$$h_S(x) = \frac{\frac{1}{e^{\lambda}-1}\sum_{k=1}^{\infty} \frac{\lambda^k}{k!}\cdot f_{N(k\mu,k\sigma^2)}(x)}{1-\left(\frac{1}{e^{\lambda}-1}\right)\sum_{k=1}^{\infty} \frac{\lambda^k}{k!}\cdot \Phi\left(\frac{x-k\mu}{\sigma\sqrt{k}}\right)},$$

ou ainda, utilizando o resultado anterior,

$$h_S(x) = \frac{\sum_{k=1}^{\infty} \frac{\lambda^k}{k!}\cdot h_{S_k}(x)\left(1-\Phi\left(\frac{x-k\mu}{\sigma\sqrt{k}}\right)\right)}{\sum_{k=1}^{\infty} \frac{\lambda^k}{k!}\cdot \left(1-\Phi\left(\frac{x-k\mu}{\sigma\sqrt{k}}\right)\right)}.$$

A seguir, o comportamento da função de risco da variável aleatória $S$. Note que o número de modas presente na função aumenta conforme o valor do parâmetro $\mu$ cresce (Figura 2.3(a)) ou quando o valor de $\sigma$ ou de $\lambda$ decresce (Figuras 2.3(b) e 2.3(c), respectivamente).

**Figura 2.3:** Função taxa de falha da Composta Poisson-Truncada Normal




(a) (b)

(c)

### 2.2.3 Função característica e função geradora de momentos

a. **Função Característica**

A função característica $\varphi_X$ de uma variável aleatória $X$ é dada por

$$\varphi_X(t) = Ee^{itX} = E\cos(tX) + iE\,\mathrm{sen}(tX),$$

em que $i = \sqrt{-1}$ e $t \in \mathbb{R}$.

Numa variável aleatória discreta, tem-se:

$$\varphi_X(t) = \sum_k e^{itx_k}\, p(x_k),$$

em que $p(x_k)$ é a função de probabilidade de $X$.

Sendo $X$ uma variável aleatória absolutamente contínua, sua função característica é definida por

$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} f_X(x)\, dx,$$

em que $f_X(x)$ é a densidade de probabilidade de $X$.

Baseado no teorema (17.10.1) de soma de variáveis aletórias apresentado em Fine (2006), poderemos apresentar a função característica da família Composta $N$, utilizando o seguinte teorema.

**Teorema:** *Se $N$ é uma variável aleatória inteira e positiva, $S = \sum_{j=1}^{N} X_j$, em que $X_j$, $j \geqslant 1$ são i.i.d. com função característica comum $\varphi_X$, e elas são independentes de $N$ que é descrita pela função característica $\varphi_N$, então*

$$\varphi_S(t) = \varphi_N(-i \log\, \varphi_X(t)), \text{ sendo } i^2 = -1.$$

**Sub-família Composta Poisson-Truncada**

Sendo $N$ uma variável aleatória com distribuição Poisson truncada no zero, com função característica $\varphi_N(t) = \dfrac{e^{\lambda e^{it}} - 1}{e^{\lambda} - 1}$, podemos encontrar uma função característica para a sub-família Composta Poisson-Truncada que dependerá da distribuição da v.a. $X$. Então,

$$\varphi_S(t) \;=\; \frac{e^{\lambda e^{i(-i \log \varphi_X(t))}} - 1}{e^{\lambda} - 1} = \frac{e^{\lambda \varphi_X(t)} - 1}{e^{\lambda} - 1}.$$

É interessante observar que em Navarrete (2013), a função de distribuição acumulada da classe *G-Poisson*, tem a mesma estrutura do resultado encontrado anteriormente. No entanto, a função característica $\varphi_X(t)$ é substituída por uma função $G(x)$, que representa a função de distribuição de qualquer distribuição *baseline* contínua.

**Exemplo 4.** Supondo a $\varphi_X(t)$ a função característica da distribuição Normal $(\mu, \sigma^2)$, temos que

$$\varphi_S(t) \;=\; \frac{e^{\lambda \exp[i\mu t - \sigma^2 t^2/2]} - 1}{e^{\lambda} - 1},$$

em que $\varphi_S(t)$ é a função característica da distribuição Composta Poisson-Truncada Normal.

b. **Função geradora de momentos**

A função geradora de momentos (fgm) de uma variável $X$ é definida por

$$M_X(t) = E(e^{tX}),$$

desde que a esperança seja finita para todo $t$ em um intervalo contendo o ponto zero no seu interior. Dado a fgm $M_X(t)$, a função característica pode ser expressa por $\varphi(x) = M_X(it)$. Como a fgm nem sempre existe, a fc é utilizada porque sempre existe.

Deste modo, podemos representar a função geradora de momentos da família Composta $N$ da seguinte forma:

$$\begin{aligned} M_S(t) &= \int_{-\infty}^{\infty} e^{tx} \cdot f_S(x)\, dx \\ &= \int_{-\infty}^{\infty} e^{tx} \cdot \sum_{k=1}^{\infty} P(N=k)\, f_{S_k}(x)\, dx \\ &= \sum_{k=1}^{\infty} P(N=k) \int_{-\infty}^{\infty} e^{tx} \cdot f_{S_k}(x)\, dx, \end{aligned}$$

em que $f_{S_k}(x)$ é a densidade da soma de $k$ variáveis aleatórias independentes com distribuição igual a de $X$. Logo, a função geradora de momentos também pode ser expressa por

$$M_S(t) = \sum_{k=1}^{\infty} P(N=k) \cdot [M_X(t)]^k = M_N(\log\, M_X(t)),$$

em que $M_N(\cdot)$ e $M_X(\cdot)$ são as fgms das variáveis aleatórias $N$ e $X$, respectivamente.

A função geradora de cumulantes (fgc) é definida como o logaritmo da função geradora de momentos, então

$$K_S(t) = \log\{M_S(t)\} = \log\{M_N(\log\, M_X(t))\} = K_N(K_X(t)).$$

**Sub-família Composta Poisson-Truncada**

Com os resultados obtidos anteriormente, podemos escrever a função geradora de momentos para a sub-família como

$$\begin{aligned} M_S(t) &= M_N(\log\, M_X(t)) \\ &= \left(\frac{1}{e^{\lambda}-1}\right) \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \cdot \exp\{(\log M_X(t))k\} \\ &= \left(\frac{1}{e^{\lambda}-1}\right) \sum_{k=1}^{\infty} \frac{(\lambda \cdot M_X(t))^k}{k!} \\ &= \frac{e^{\lambda \cdot M_X(t)}-1}{e^{\lambda}-1}. \end{aligned}$$

Dessa forma, a função geradora de cumulantes é dada por

$$K_S(t) = \log\left\{\frac{e^{\lambda \cdot M_X(t)}-1}{e^{\lambda}-1}\right\}$$

**Exemplo 5.** Considere $X \sim N(\mu, \sigma^2)$ uma v.a. com distribuição Normal, temos que a fgm da distribuição $CPTN(\lambda, \mu, \sigma^2)$ é dada por

$$M_S(t) = \left(\frac{1}{e^\lambda - 1}\right) \left\{\exp\left[\lambda \exp(t\mu + \sigma^2 t^2/2)\right] - 1\right\}.$$

Dessa maneira, a sua função geradora de cumulantes é dada por

$$K_S(t) = \log\left\{\frac{\exp\left[\lambda \exp(t\mu + \sigma^2 t^2/2)\right] - 1}{e^\lambda - 1}\right\}.$$

### 2.2.4 Momentos

Utilizando a expressão dada para $f_S(x)$ é possível obter os momentos ordinários de $S$, em termos de uma série dos momentos de $X$:

$$\begin{aligned} E(S^j) &= \int_{-\infty}^{\infty} x^j \, f_S(x)\, dx \\ &= \int_{-\infty}^{\infty} x^j \sum_{k=1}^{\infty} P(N = k)\, f_{S_k}(x)\, dx \\ &= \sum_{k=1}^{\infty} P(N = k) \int_{-\infty}^{\infty} x^j \, f_{S_k}(x)\, dx \end{aligned}$$

Note que

$$\int_{-\infty}^{\infty} x^j \, f_{S_k}(x)\, dx = E(X_1 + \ldots + X_k)^j,$$

em que $X_1, \ldots, X_k$ têm distribuição igual a de $X$ e são independentes. Logo, utilizando o teorema multinomial e a independência dos $X_i$'s, temos:

$$E[(X_1 + \ldots + X_k)^j] = \sum_{j_1=0}^{j} \sum_{j_2=0}^{j-j_1} \cdots \sum_{j_{k-1}=0}^{j-j_1-j_2-\ldots j_{k-2}} \binom{j}{j_1 \ldots j_k} \times E(X_1^{j_1}) E(X_2^{j_2}) \ldots E(X_k^{j_k}),$$

em que $j_k = j - j_1 - j_2 - \ldots - j_{k-1}$. Como os $X_i$'s têm a mesma distribuição de $X$, segue que:

$$E[(X_1 + \ldots + X_k)^j] = \sum_{j_1=0}^{j} \sum_{j_2=0}^{j-j_1} \cdots \sum_{j_{k-1}=0}^{j-j_1-j_2-\ldots j_{k-2}} \binom{j}{j_1 \ldots j_k} \times E(X^{j_1}) E(X^{j_2}) \ldots E(X^{j_k}),$$

em que $j_k = j - j_1 - j_2 - \ldots - j_{k-1}$.

Portanto,

$$E(S^j) = \sum_{k=0}^{\infty} P(N = k) \sum_{j_1=0}^{j} \sum_{j_2=0}^{j-j_1} \cdots \sum_{j_{k-1}=0}^{j-j_1-j_2-\ldots j_{k-2}} \binom{j}{j_1 \ldots j_k} \prod_{i=1}^{k} E(X^{j_i}),$$

em que $j_k = j - j_1 - j_2 - \ldots - j_{k-1}$.

**Sub-família Composta Poisson-Truncada**

Nesse caso, os momentos ordinários da variável aleatória $S$ podem ser encontrados pela expressão que segue.

$$E(S^j) = \left(\frac{1}{e^\lambda - 1}\right) \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \sum_{j_1=0}^{j} \sum_{j_2=0}^{j-j_1} \cdots \sum_{j_{k-1}=0}^{j-j_1-j_2-\ldots j_{k-2}} \binom{j}{j_1 \ldots j_k} \prod_{i=1}^{k} E(X^{j_i}),$$

em que $j_k = j - j_1 - j_2 - \ldots - j_{k-1}$.

Outra forma de calcular os momentos ordinários da v.a. $S$ é através da sua função característica $\varphi_S(t)$ apresentada na seção anterior. Dessa forma, o $j$-ésimo momento de $S$ é a razão entre a $j$-ésima derivada da sua função característica, calculada no ponto zero, e o número imaginário $i$ elevado a $j$, isto é,

$$E(S^j) = \frac{\varphi_S^{(j)}(0)}{i^j}.$$

A seguir, serão apresentadas as equações para o cálculo dos três primeiros momentos. A expressão para o primeiro momento é conhecida como *equação de Wald* para o cálculo de valores esperados por condicionamento. Mais informações podem ser encontradas em Resnick (1992). O desenvolvimento dos cálculos encontra-se no Apêndice A.1.

1. **Primeiro Momento**

$$E(S) = \frac{\varphi_S^{(1)}(0)}{i} = E(N) \cdot E(X).$$

2. **Segundo Momento**

$$E(S^2) = \frac{\varphi_S^{(2)}(0)}{i^2} = E(N^2) \cdot [E(X)]^2 + E(N) \cdot \mathrm{Var(X)}.$$

3. **Terceiro Momento**

$$\begin{aligned} E(S^3) &= \frac{\varphi_S^{(3)}(0)}{i^3} \\ &= E(N^3) \cdot [E(X)]^3 - 3 \cdot E(N^2) \cdot [E(X)]^3 + 3 \cdot E(N^2) \cdot E(X) \cdot E(X^2) \\ &\quad -3 \cdot E(N) \cdot E(X) \cdot E(X^2) + E(N) \cdot E(X^3) + 2 \cdot E(N)[E(X)]^3 \end{aligned}$$

**Sub-família Composta Poisson-Truncada**

Calculando os três primeiros momentos da distribuição Poisson-Truncada, obtemos os seguintes momentos da sub-família proposta.

1. **Primeiro Momento**

$$E(S) = \left(\frac{\lambda}{1 - e^{-\lambda}}\right) \cdot E(X).$$

2. **Segundo Momento**

$$E(S^2) = \left(\frac{1}{1-e^{-\lambda}}\right)\cdot\{(\lambda+\lambda^2)\cdot[E(X)]^2+\lambda\cdot \mathrm{Var(X)}\}.$$

3. **Terceiro Momento**

$$\begin{aligned} E(S^3) &= \left(\frac{1}{1-e^{-\lambda}}\right)\cdot\{(\lambda+3\lambda^2+\lambda^3)\cdot[E(X)]^3-3\cdot(\lambda+\lambda^2)\cdot[E(X)]^3\} \\ &\quad +3\cdot(\lambda+\lambda^2)\cdot E(X)\cdot E(X^2)-3\cdot\lambda\cdot E(X)\cdot E(X^2) \\ &\quad +\lambda\cdot E(X^3)+2\cdot\lambda\cdot[E(X)]^3\}. \end{aligned}$$

**Exemplo 6.** Considere $X \sim N(\mu,\sigma^2)$. Os três primeiros momentos da distribuição $CPTN(\lambda,\mu,\sigma^2)$, são dados por:

**Primeiro Momento.**

$$\begin{aligned} E(S) &= E(X)\cdot E(N) \\ E(S) &= \mu\cdot\left(\frac{\lambda}{1-e^{-\lambda}}\right). \end{aligned}$$

**Segundo Momento.**

$$\begin{aligned} E(S^2) &= E(N^2)\cdot[E(X)]^2+E(N)\cdot \mathrm{Var(X)} \\ E(S^2) &= \left(\frac{1}{1-e^{-\lambda}}\right)\cdot\{(\lambda+\lambda^2)\cdot\mu^2+\lambda\cdot\sigma^2\}. \end{aligned}$$

**Terceiro Momento.**

$$\begin{aligned} E(S^3) &= E(N^3)\cdot[E(X)]^3-3\cdot E(N^2)\cdot[E(X)]^3 \\ &\quad +3\cdot E(N^2)\cdot E(X)\cdot E(X^2)-3\cdot E(N)\cdot E(X)\cdot E(X^2) \\ &\quad +E(N)\cdot E(X^3)+2\cdot E(N)[E(X)]^3 \\ E(S^3) &= \left(\frac{\lambda+3\lambda^2+\lambda^3}{1-e^{-\lambda}}\right)\cdot\mu^3-3\left(\frac{\lambda+\lambda^2}{1-e^{-\lambda}}\right)\cdot\mu^3 \\ &\quad +3\left(\frac{\lambda+\lambda^2}{1-e^{-\lambda}}\right)\cdot\mu\cdot(\sigma^2+\mu^2) \\ &\quad -3\left(\frac{\lambda}{1-e^{-\lambda}}\right)\cdot\mu\cdot(\sigma^2+\mu^2) \\ &\quad +\left(\frac{\lambda}{1-e^{-\lambda}}\right)\cdot(\mu^3+3\sigma^2\mu)+2\left(\frac{\lambda}{1-e^{-\lambda}}\right)\cdot\mu^3 \\ (1-e^{-\lambda})\cdot E(S^3) &= \lambda\mu^3+3\lambda^2\mu^3+\lambda^3\mu^3-3\lambda\mu^3-3\lambda^2\mu^3 \\ &\quad +3\lambda\mu\sigma^2+3\lambda\mu^3+3\lambda^2\mu\sigma^2+3\lambda^2\mu^3-3\lambda\mu\sigma^2 \\ &\quad -3\lambda\mu^3+\lambda\mu^3+3\lambda\sigma^2\mu+2\lambda\mu^3 \\ E(S^3) &= \left(\frac{1}{1-e^{-\lambda}}\right)\cdot\left[\lambda\mu^3+\lambda^3\mu^3+3\lambda\mu\sigma^2+3\lambda^2\mu\sigma^2+3\lambda^2\mu^3\right]. \end{aligned}$$

### 2.2.5 Entropia de Shannon

A entropia de Shannon mede a quantidade de incerteza associada a uma variável aleatória. Essa teoria tem sido utilizada com sucesso em diversas aplicações, especialmente na área de Teoria da Informação (SHANNON, 1948). Essa medida calculada para uma v.a. $S$ com distribuição Composta $N$ é representada por:

$$\begin{aligned}\mathbb{H}_{Sh}[f(x)] &= E(-\log[f(X)]) = -\int_{-\infty}^{\infty} (\log[f_S(x)]) \cdot f_S(x)\, dx \\ &= -\int_{-\infty}^{\infty} \left\{ \log\left[ \sum_{k=1}^{\infty} P(N=k)\, f_{S_k}(x) \right] \times \sum_{k=1}^{\infty} P(N=k)\, f_{S_k}(x) \right\} dx\end{aligned}$$

**Sub-família Composta Poisson-Truncada**

Utilizando o resultado obtido anteriormente, a Entropia de Shannon para a sub-família é dada da seguinte forma:

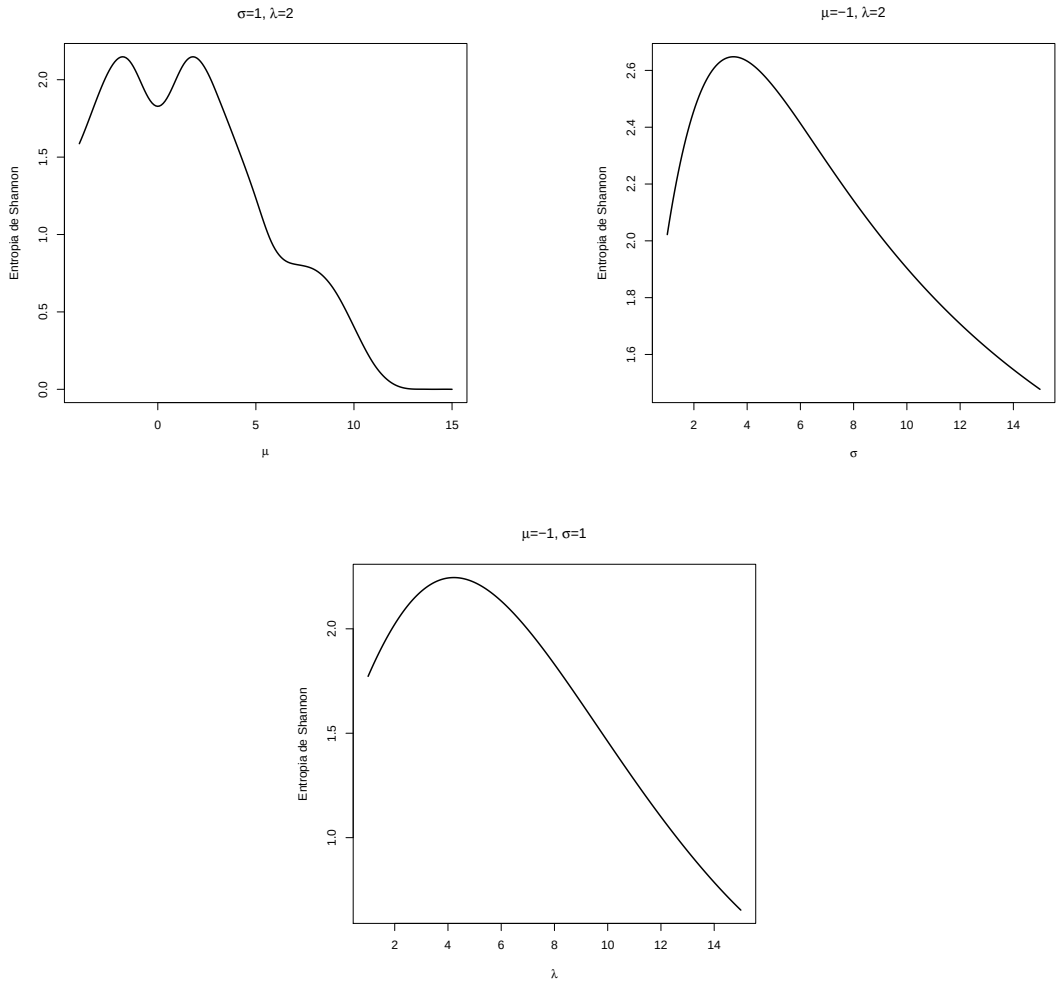
$$\begin{aligned}\mathbb{H}_{Sh}[f(x)] &= -\int_{-\infty}^{\infty} \left\{ \log\left[ \sum_{k=1}^{\infty} P(N=k)\, f_{S_k}(x) \right] \times \sum_{k=1}^{\infty} P(N=k)\, f_{S_k}(x) \right\} dx \\ &= -\int_{-\infty}^{\infty} \left\{ \log\left[ \left(\frac{1}{e^\lambda - 1}\right) \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \cdot f_{S_k}(x) \right] \times \left(\frac{1}{e^\lambda - 1}\right) \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \cdot f_{S_k}(x) \right\} dx \\ &= [\log(e^\lambda - 1)] \int_{-\infty}^{\infty} \left(\frac{1}{e^\lambda - 1}\right) \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \cdot f_{S_k}(x)\, dx \\ &\quad -\int_{-\infty}^{\infty} \left(\frac{1}{e^\lambda - 1}\right) \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \cdot f_{S_k}(x) \cdot \log\left( \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \cdot f_{S_k}(x) \right) dx \\ &= \log(e^\lambda - 1) - E\left[ \log\left( \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \cdot f_{S_k}(X) \right) \right].\end{aligned}$$

**Exemplo 7.** Considere $S \sim CPTN(\lambda, \mu, \sigma^2)$ com densidade dada em (2.4). Assim, a Entropia de Shannon de $S$ é definida como

$$\begin{aligned}\mathbb{H}_{Sh}[f(x)] &= \log(e^\lambda - 1) - E\left[ \log\left( \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \cdot f_{S_k}(x) \right) \right] \\ &= \log(e^\lambda - 1) - E\left[ \log\left( \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \cdot f_{N(k\mu, k\sigma^2)}(x) \right) \right] \\ &= \log(e^\lambda - 1) - E\left[ \log\left( \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \cdot (2\pi k\sigma^2)^{-1/2} \exp\left( -\frac{1}{2} \frac{(x-k\mu)^2}{k\sigma^2} \right) \right) \right].\end{aligned}$$

Algumas formas dessa entropia são apresentadas na Figura 2.4.

**Figura 2.4:** Entropia de Shannon da Composta Poisson-Truncada Normal



# CAPÍTULO 3

## Métodos de estimação e resultados numéricos

Neste capítulo, apresentamos três métodos para a estimação dos parâmetros da distribuição CPTN: Método dos momentos (MM), método da função característica empírica (FCE) e método de máxima verossimilhança (MV). Em uma primeira análise, fazemos alguns estudos de casos comparando a eficiência dos métodos de estimação FCE e MV via algoritmo EM. Como nestes estudos de casos, o método de estimação FCE foi computacionalmente ineficiente, na análise de simulação Monte Carlo, não utilizamos este método, mas comparamos as estimativas obtidas pelo método dos momentos, com as obtidas pelo método de MV via algoritmo EM, utilizando como chute inicial o resultado do método dos momentos. Ainda, nesta análise, obtivemos as estimativas pelo método de máxima verossimilhança via algoritmo EM, utilizando como chute inicial o próprio parâmetro utilizado para gerar a amostra, com o intuito de analisar como os parâmetros reais influenciam na performance do método. Finalmente, concluímos este capítulo apresentando uma aplicação desta nova distribuição proposta na modelagem de uma imagem SAR (Synthetic Aperture Radar - SAR) extraída da imagem de Foulum (Dinamarca).

## 3.1 Método dos momentos

Seja $S_1, S_2, \ldots, S_n$ uma amostra aleatória simples de uma população $S$ que possua distribuição $CPTN(\lambda, \mu, \sigma^2)$. No método dos momentos, encontramos os estimadores dos parâmetros igualando os momentos populacionais aos seus respectivos momentos amostrais. Sendo assim,

$$E(S^j) = \frac{1}{n}\sum_{i=1}^{n} S_i^j.$$

Deste modo, obtemos o seguinte sistema:

$$\begin{cases} E(S) &= \frac{1}{n}\sum_{i=1}^{n} S_i &\equiv \overline{S} \\ E(S^2) &= \frac{1}{n}\sum_{i=1}^{n} S_i^2 &\equiv \overline{S^2} \\ E(S^3) &= \frac{1}{n}\sum_{i=1}^{n} S_i^3 &\equiv \overline{S^3} \end{cases}$$

Pela Seção 2.2.4, seguem os seguintes resultados:

1. $\overline{S} = E(S)$

$$\begin{aligned} \overline{S} &= \mu \cdot \left(\frac{\lambda}{1-e^{-\lambda}}\right) \\ \mu &= \overline{S} \cdot \left(\frac{1-e^{-\lambda}}{\lambda}\right) \end{aligned} \tag{3.1}$$

2. $\overline{S^2} = E(S^2)$

$$\begin{aligned} \overline{S^2} &= \left(\frac{\lambda+\lambda^2}{1-e^{-\lambda}}\right)\cdot\mu^2 + \left(\frac{\lambda}{1-e^{-\lambda}}\right)\cdot\sigma^2 \\ \sigma^2 &= \overline{S^2}\cdot\left(\frac{1-e^{-\lambda}}{\lambda}\right) - \left(\frac{\lambda+\lambda^2}{\lambda}\right)\cdot\mu^2 \\ &= \overline{S^2}\cdot\left(\frac{1-e^{-\lambda}}{\lambda}\right) - \mu^2 - \lambda\mu^2. \end{aligned} \tag{3.2}$$

3. $\overline{S^3} = E(S^3)$

$$(1-e^{-\lambda})\cdot\overline{S^3} = \lambda\mu^3 + \lambda^3\mu^3 + 3\lambda\mu\sigma^2 + 3\lambda^2\mu\sigma^2 + 3\lambda^2\mu^3.$$

Substituindo $\sigma^2$, temos:

$$\begin{aligned} (1-e^{-\lambda})\cdot\overline{S^3} &= \lambda\mu^3 + \lambda^3\mu^3 + 3\lambda^2\mu^3 \\ &\quad +3\lambda\mu\cdot\left[\overline{S^2}\cdot\left(\frac{1-e^{-\lambda}}{\lambda}\right) - \mu^2 - \lambda\mu^2\right] \\ &\quad +3\lambda^2\mu\cdot\left[\overline{S^2}\cdot\left(\frac{1-e^{-\lambda}}{\lambda}\right) - \mu^2 - \lambda\mu^2\right] \\ (1-e^{-\lambda})\cdot\overline{S^3} &= \lambda\mu^3 + \lambda^3\mu^3 + 3\lambda^2\mu^3 + 3\mu\overline{S^2}\cdot(1-e^{-\lambda}) - 3\lambda\mu^3 \\ &\quad -3\lambda^2\mu^3 + 3\lambda\mu\overline{S^2}\cdot(1-e^{-\lambda}) - 3\lambda^2\mu^3 - 3\lambda^3\mu^3 \\ &= 3\mu\overline{S^2}\cdot(1-e^{-\lambda}) + 3\lambda\mu\overline{S^2}\cdot(1-e^{-\lambda}) \\ &\quad -3\lambda^2\mu^3 - 2\lambda\mu^3 - 2\lambda^3\mu^3. \end{aligned}$$

Substituindo $\mu$, segue:

$$\begin{aligned}
(1-e^{-\lambda})\cdot\overline{S^3} &= 3\left[\overline{S}\cdot\left(\frac{1-e^{-\lambda}}{\lambda}\right)\right]\cdot\overline{S^2}\cdot(1-e^{-\lambda}) \\
&\quad +3\lambda\left[\overline{S}\cdot\left(\frac{1-e^{-\lambda}}{\lambda}\right)\right]\cdot\overline{S^2}\cdot(1-e^{-\lambda}) \\
&\quad -3\lambda^2\left[\overline{S}\cdot\left(\frac{1-e^{-\lambda}}{\lambda}\right)\right]^3-2\lambda\left[\overline{S}\cdot\left(\frac{1-e^{-\lambda}}{\lambda}\right)\right]^3 \\
&\quad -2\lambda^3\left[\overline{S}\cdot\left(\frac{1-e^{-\lambda}}{\lambda}\right)\right]^3 \\
\overline{S^3} &= 3\overline{S}\cdot\overline{S^2}\cdot\left(\frac{1-e^{-\lambda}}{\lambda}\right)+3\overline{S}\cdot\overline{S^2}\cdot(1-e^{-\lambda}) \\
&\quad -\frac{3(\overline{S})^3\cdot(1-e^{-\lambda})^2}{\lambda}-2(\overline{S})^3\cdot\left(\frac{1-e^{-\lambda}}{\lambda}\right)^2 \\
&\quad -2\cdot(\overline{S})^3\cdot(1-e^{-\lambda})^2. \qquad (3.3)
\end{aligned}$$

Para a estimação dos parâmetros, utilizamos a equação não-linear em (3.3) para encontrar a estimativa do parâmetro $\lambda$ e utilizando o resultado obtido, estimamos os outros parâmetros. No Algoritmo 1 encontram-se os passos dessa estimação.

**Algoritmo 1:** Algoritmo para o método dos momentos.

**Passo** 1. Encontre as soluções de $\lambda$ que satisfazem (3.3). Diga-se $\hat{\lambda}_k$, para $k = 1, \ldots, m$;
**Passo** 2. Se m for infinito ou for nulo, não existe solução;
**Passo** 3. Para $k = 1, \ldots, m$, encontre $\hat{\mu}_k$ resolvendo (3.1) e $\hat{\sigma}_k^2$ resolvendo (3.2);
**Passo** 4. Se para todo $k = 1, \ldots, m$, $\hat{\sigma}_k^2 \leqslant 0$, não existe solução;
**Passo** 5. Para cada $\hat{\theta}_k = (\hat{\lambda}_k, \hat{\mu}_k, \hat{\sigma}_k^2)$, tal que $\hat{\sigma}_k^2 > 0$, determine a função de log-verossimilhança;
**Passo** 6. Forneça como solução o terno do Passo 5 com maior função de log-verossimilhança.

## 3.2 Método por função característica empírica

Seja $F(x;\theta)$ a fda de $X$ que depende de $\theta$ (um vetor $K$-dimensional de parâmetros). A Função Característica é definida por

$$\varphi(r;\theta) = E[\exp(irX)] = \int \exp(irx)\, dF(x;\theta),$$

e a função característica empírica é a transformada de Fourier-Stieltjes da fda empírica $F_n(x)$. Segue, então

$$\varphi_n(r;\theta) = \frac{1}{n}\sum_{j=1}^{n}\exp(irX_j)] = \int \exp(irx)\, dF_n(x;\theta),$$

em que $i^2 = -1$, $r$ é a variável transformada, e $X_1, \cdots, X_n$ são variáveis aleatórias iid.

Quando o parâmetro $r$ é contínuo, Press (1972) e Paulson *et al.* (1975) propuseram um estimador que minimize a seguinte distância entre a função característica empírica e a função característica teórica,

$$D(\theta; x) = \int_{-\infty}^{\infty} |\varphi_n(r) - \varphi(r)|^{\nu} \, g(r) \, dr,$$

sendo $g(r)$ uma função de peso contínua.

Quando $\nu = 2$, Knight e Yu (2002) definiram um estimador que minimiza a distância

$$D(\theta; x) = \int_{-\infty}^{\infty} |\varphi_n(r) - \varphi(r)|^{2} \, g(r) \, dr,$$

em que $g(r)$ assegura a convergência da integral.

Utilizando $g(r) = \exp(-r^2)$ como sugerido em Yu (2004), aplicamos o método para a função característica da distribuição $CPTN(\lambda, \mu, \sigma^2)$. Então,

$$\int_{-\infty}^{\infty} \left| \frac{1}{n} \sum_{j=1}^{n} e^{irX_j} - \left( \frac{e^{\lambda \exp[i\mu r - \sigma^2 r^2/2]} - 1}{e^{\lambda} - 1} \right) \right|^2 \exp(-r^2) \, dr$$

$$= \frac{\pi^{1/2}}{n^2} \sum_{j=1}^{n} \sum_{t=1}^{n} e^{-\frac{1}{4}(X_j - X_t)^2} - \left( \frac{2}{n(e^{\lambda} - 1)} \right) \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \left( \frac{2\pi}{2 + k\sigma^2} \right)^{1/2} \sum_{j=1}^{n} \left( \exp\left\{ \frac{-(X_j - k\mu)^2}{4 + 2k\sigma^2} \right\} \right)$$
$$+ \left( \frac{2\pi^{1/2}}{n(e^{\lambda} - 1)} \right) \sum_{j=1}^{n} \exp\left\{ -\frac{1}{4} X_j^2 \right\} - \left( \frac{2}{(e^{\lambda} - 1)} \right) \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \left( \frac{2\pi}{2 + k\sigma^2} \right)^{1/2} \exp\left\{ \frac{-k^2\mu^2}{4 + 2k\sigma^2} \right\}$$
$$+ \left( \frac{\pi}{(e^{\lambda} - 1)^4} \right)^{1/2} + \left( \frac{1}{(e^{\lambda} - 1)^2} \right) \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \sum_{p=0}^{k} \binom{k}{p} \left( \frac{2\pi}{2 + k\sigma^2} \right)^{1/2} \exp\left\{ \frac{-(k - 2p)^2 \mu^2}{4 + 2k\sigma^2} \right\}.$$

Para maiores esclarecimentos o desenvolvimento dos cálculos encontra-se no Apêndice A.2.

Para estimar os parâmetros da distribuição CPTN, minimizamos a distância $D(\theta, x)$ obtida, utilizando a função `Optim()` da plataforma `R` com os seguintes argumentos: um vetor de parâmetros iniciais, a função a ser minimizada ($D(\theta, x)$) e o método utilizado para a otimização, nesse caso, o método BFGS.

## 3.3 Método de máxima verossimilhança

Seja $S_N = (S_1, \ldots, S_n)^{\top}$ uma a.a. de tamanho $n$ de $S \sim CPTN(\lambda, \mu, \sigma^2)$ com densidade dada em (2.4). Dada uma amostra observada $s_N = (x_1, \ldots, x_n)^{\top}$, a função de verossimilhança e o logaritmo da função de verossimilhança para o vetor de parâmetros $\theta = (\lambda, \mu, \sigma)^{\top}$ são dadas, respectivamente, por

$$L(\theta; x) = \prod_{i=1}^{n} f_S(x_i) = \prod_{i=1}^{n} \left\{ \left[ \frac{1}{e^{\lambda} - 1} \right] \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \, f_{N(k\mu, k\sigma^2)}(x_i) \right\}$$

e

$$\log L(\theta; x) = l(\theta) = \sum_{i=1}^{n} \log f_S(x_i) = -n \log(e^{\lambda} - 1) + \sum_{i=1}^{n} \log \left\{ \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} f_{N(k\mu, k\sigma^2)}(x_i) \right\} \tag{3.4}$$

### 3.3.1 Função escore

A seguinte discussão define um algoritmo a fim de encontrar os estimadores de máxima verossimilhança (EMVs) para o vetor $\theta$ da distribuição CPTN. Como a função de verossimilhança não apresenta solução analítica explícita, os estimadores podem ser obtidos através de métodos numéricos. Seja $U(\theta) = \frac{\partial l(\theta)}{\partial \theta}$ a função escore, para encontrar o estimador de máxima verossimilhança $\hat{\theta}$, é necessário resolver $U(\hat{\theta}) = 0$. Com base em (3.4), as funções escore são dadas por:

1. Função escore de $\lambda$

$$U_\lambda = \frac{\partial l(\theta)}{\partial \lambda} = -n \left[ \frac{e^{\lambda}}{e^{\lambda} - 1} \right] + \sum_{i=1}^{n} \left\{ \frac{\sum_{k=1}^{\infty} \frac{k\lambda^{k-1}}{k!} f_{N(k\mu, k\sigma^2)}(x_i)}{\sum_{k=1}^{\infty} \frac{\lambda^k}{k!} f_{N(k\mu, k\sigma^2)}(x_i)} \right\}$$

$$U_\lambda = -n \left[ \frac{e^{\lambda}}{e^{\lambda} - 1} \right] + \left[ \frac{1}{e^{\lambda} - 1} \right] \sum_{i=1}^{n} \frac{1}{f_S(x_i; \theta)} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} f_{N(k\mu, k\sigma^2)}(x_i).$$

De $U_\lambda \Big|_{(\lambda, \mu, \sigma^2) = (\hat{\lambda}, \hat{\mu}, \hat{\sigma}^2)} = 0$, temos:

$$n \left[ \frac{e^{\hat{\lambda}}}{e^{\hat{\lambda}} - 1} \right] = \left[ \frac{1}{e^{\hat{\lambda}} - 1} \right] \sum_{i=1}^{n} \left[ \frac{1}{f_S(x_i; \hat{\theta})} \right] \sum_{k=1}^{\infty} \frac{\hat{\lambda}^{k-1}}{(k-1)!} f_{N(k\hat{\mu}, k\hat{\sigma}^2)}(x_i). \tag{3.5}$$

Por simplicidade, denotemos $\hat{f}_S(x_i) \equiv f_S(x_i; \hat{\theta})$ e
$\hat{f}_N(x_i) \equiv f_{N(k\hat{\mu}, k\hat{\sigma}^2)}(x_i)$. Assim, a Equação (3.5) resulta em

$$n \left[ \frac{e^{\hat{\lambda}}}{e^{\hat{\lambda}} - 1} \right] = \sum_{i=1}^{n} \left[ \frac{1}{\hat{f}_S(x_i)} \right] \left[ \frac{1}{e^{\hat{\lambda}} - 1} \right] \sum_{k=1}^{\infty} \frac{\hat{\lambda}^{k-1}}{(k-1)!} \hat{f}_N(x_i).$$

No que segue, adotaremos

$$\tilde{f}_S(x_i) \equiv \left[ \frac{1}{e^{\hat{\lambda}} - 1} \right] \sum_{k=1}^{\infty} \frac{\hat{\lambda}^{k-1}}{(k-1)!} \hat{f}_N(x_i). \tag{3.6}$$

Logo,

$$\frac{e^{\hat{\lambda}}}{e^{\hat{\lambda}} - 1} = \frac{1}{n} \sum_{i=1}^{n} \left[ \frac{\tilde{f}_S(x_i)}{\hat{f}_S(x_i)} \right]. \tag{3.7}$$

2. Função escore de $\mu$

$$U_\mu = \frac{\partial l(\theta)}{\partial \mu} = \sum_{i=1}^{n} \left\{ \frac{\sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \left[ \frac{\partial f_{N(k\mu,k\sigma^2)}(x_i)}{\partial \mu} \right]}{\sum_{k=1}^{\infty} \frac{\lambda^k}{k!} f_{N(k\mu,k\sigma^2)}(x_i)} \right\}.$$

Dessa forma, temos:

$$U_\mu \Big|_{(\lambda,\mu,\sigma^2)=(\hat{\lambda},\hat{\mu},\hat{\sigma}^2)} = \left[ \frac{1}{e^{\hat{\lambda}} - 1} \right] \sum_{i=1}^{n} \left\{ \left[ \frac{1}{\hat{f}_S(x_i)} \right] \times \sum_{k=1}^{\infty} \left[ \frac{\hat{\lambda}^k}{k!} \hat{f}_N(x_i) \times \frac{\partial \log[f_N(x_i; k\mu, k\sigma^2)]}{\partial \mu} \Big|_{(\lambda,\mu,\sigma^2)=(\hat{\lambda},\hat{\mu},\hat{\sigma}^2)} \right] \right\}$$

Fazendo $U_\mu \Big|_{(\lambda,\mu,\sigma^2)=(\hat{\lambda},\hat{\mu},\hat{\sigma}^2)} = 0$, segue:

$$\left[ \frac{1}{e^{\hat{\lambda}} - 1} \right] \sum_{i=1}^{n} \left[ \frac{1}{\hat{f}_S(x_i)} \right] \sum_{k=1}^{\infty} \frac{\hat{\lambda}^k}{k!} \hat{f}_N(x_i) \times \left( \frac{x_i - k\hat{\mu}}{\hat{\sigma}^2} \right) = 0$$

$$\left[ \frac{1}{e^{\hat{\lambda}} - 1} \right] \sum_{i=1}^{n} \left[ \frac{(x_i/\hat{\sigma}^2)}{\hat{f}_S(x_i)} \right] \sum_{k=1}^{\infty} \frac{\hat{\lambda}^k}{k!} \hat{f}_N(x_i) =$$
$$\left[ \frac{1}{e^{\hat{\lambda}} - 1} \right] \times \sum_{i=1}^{n} \left[ \frac{(\hat{\mu}/\hat{\sigma}^2)}{\hat{f}_S(x_i)} \right] \sum_{k=1}^{\infty} \frac{\hat{\lambda}^k}{(k-1)!} \hat{f}_N(x_i)$$

$$\sum_{i=1}^{n} \left[ \frac{x_i}{\hat{f}_S(x_i)} \times \hat{f}_S(x_i) \right] =$$
$$\left[ \frac{\hat{\mu}}{e^{\hat{\lambda}} - 1} \right] \times \sum_{i=1}^{n} \left[ \frac{1}{\hat{f}_S(x_i)} \right] \sum_{k=1}^{\infty} \frac{\hat{\lambda}^k}{(k-1)!} \hat{f}_N(x_i)$$

$$\sum_{i=1}^{n} x_i = \hat{\mu} \sum_{i=1}^{n} \left[ \frac{1}{\hat{f}_S(x_i)} \right] \left[ \frac{1}{e^{\hat{\lambda}} - 1} \right] \sum_{k=1}^{\infty} \frac{\hat{\lambda}^k}{(k-1)!} \hat{f}_N(x_i)$$

$$\overline{x} = \hat{\mu} \frac{1}{n} \sum_{i=1}^{n} \left[ \frac{1}{\hat{f}_S(x_i)} \right] \left[ \frac{1}{e^{\hat{\lambda}} - 1} \right] \sum_{k=1}^{\infty} \frac{\hat{\lambda}^k}{(k-1)!} \hat{f}_N(x_i)$$

$$\overline{x} = \hat{\mu}\, \hat{\lambda} \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[ \frac{1}{\hat{f}_S(x_i)} \right] \left[ \frac{1}{e^{\hat{\lambda}} - 1} \right] \sum_{k=1}^{\infty} \frac{\hat{\lambda}^{k-1}}{(k-1)!} \hat{f}_N(x_i) \right\}$$

Utilizando a equivalência da Equação (3.6), temos

$$\overline{x} = \hat{\mu}\, \hat{\lambda} \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[ \frac{\tilde{f}_S(x_i)}{\hat{f}_S(x_i)} \right] \right\}, \tag{3.8}$$

em que $\hat{\mu}$ é o EMV de $\mu$.

Adicionalmente, substituindo o resultado da Equação (3.7) na (3.8), a média amostral é dada por

$$\overline{x} = \hat{\mu}\,\hat{\lambda}\left[\frac{e^{\hat{\lambda}}}{e^{\hat{\lambda}}-1}\right].$$

3. Função escore de $\sigma^2$

$$U_{\sigma^2} = \frac{\partial l(\theta)}{\partial \sigma^2} = \sum_{i=1}^{n}\left\{\frac{\sum_{k=1}^{\infty}\frac{\lambda^k}{k!}\left[\frac{\partial f_{N(k\mu,k\sigma^2)}(x_i)}{\partial\sigma^2}\right]}{\sum_{k=1}^{\infty}\frac{\lambda^k}{k!}\,f_{N(k\mu,k\sigma^2)}(x_i)}\right\}$$

Segue, então

$$U_{\sigma}\Big|_{(\lambda,\mu,\sigma^2)=(\hat{\lambda},\hat{\mu},\hat{\sigma}^2)} = \left[\frac{1}{e^{\hat{\lambda}}-1}\right]\sum_{i=1}^{n}\left\{\left[\frac{1}{\hat{f}_S(x_i)}\right]\times \sum_{k=1}^{\infty}\left[\frac{\hat{\lambda}^k}{k!}\,\hat{f}_N(x_i)\times\frac{\partial\log[f_N(x_i;k\mu,k\sigma^2)]}{\partial\sigma^2}\Big|_{(\lambda,\mu,\sigma^2)=(\hat{\lambda},\hat{\mu},\hat{\sigma}^2)}\right]\right\}$$

De $U_{\sigma^2}\Big|_{(\lambda,\mu,\sigma^2)=(\hat{\lambda},\hat{\mu},\hat{\sigma}^2)} = 0$, tem-se:

$$\left[\frac{1}{e^{\hat{\lambda}}-1}\right]\sum_{i=1}^{n}\left[\frac{1}{\hat{f}_S(x_i)}\right]\sum_{k=1}^{\infty}\frac{\hat{\lambda}^k}{k!}\,\hat{f}_N(x_i)\times\left(-\frac{1}{2\hat{\sigma}^2}+\frac{(x_i-k\hat{\mu})^2}{2k(\hat{\sigma}^2)^2}\right)=0$$

$$\left(\frac{1}{2\hat{\sigma}^2}\right)\sum_{i=1}^{n}\left[\frac{1}{\hat{f}_S(x_i)}\right]\left[\frac{1}{e^{\hat{\lambda}}-1}\right]\sum_{k=1}^{\infty}\frac{\hat{\lambda}^k}{k!}\,\hat{f}_N(x_i) = \left(\frac{1}{2(\hat{\sigma}^2)^2}\right)\times\sum_{i=1}^{n}\left[\frac{1}{\hat{f}_S(x_i)}\right]\left[\frac{1}{e^{\hat{\lambda}}-1}\right]\sum_{k=1}^{\infty}\frac{\hat{\lambda}^k}{k!}\,\frac{(x_i-k\hat{\mu})^2}{k}\,\hat{f}_N(x_i)$$

$$\sum_{i=1}^{n}\left[\frac{1}{\hat{f}_S(x_i)}\times\hat{f}_S(x_i)\right] = \left(\frac{1}{\hat{\sigma}^2}\right)\sum_{i=1}^{n}\left[\frac{1}{\hat{f}_S(x_i)}\right]\times\left[\frac{1}{e^{\hat{\lambda}}-1}\right]\sum_{k=1}^{\infty}\left(\frac{x_i-k\hat{\mu}}{\sqrt{k}}\right)^2\frac{\hat{\lambda}^k}{k!}\,\hat{f}_N(x_i)$$

Temos, portanto,

$$\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{\hat{f}_S(x_i)}\right]\tilde{\tilde{f}}_S(x_i), \tag{3.9}$$

em que

$$\tilde{\tilde{f}}_S(x_i) \equiv \left[\frac{1}{e^{\hat{\lambda}}-1}\right]\sum_{k=1}^{\infty}\left(\frac{x_i-k\hat{\mu}}{\sqrt{k}}\right)^2\frac{\hat{\lambda}^k}{k!}\,\hat{f}_N(x_i).$$

**Algoritmo 2:** Estimadores de máxima verossimilhança para a distribuição CPTN.

**Passo 1.** Calcule $\overline{x}$;
**Passo 2.** Escolha um chute inicial $\theta_0 = (\lambda_0, \mu_0, \sigma_0^2)^\top$;
**Passo 3.** Faça $h = 0$;
**Passo 4.** Encontre a solução do sistema formado pelas equações não lineares (3.7)-(3.9), diga-se $(\lambda_{h+1}, \mu_{h+1}, \sigma_{h+1}^2)$ dado que a obtida da iteração anterior foi $(\lambda_h, \mu_h, \sigma_h^2)$;
**Passo 5.** Calcule $E = |\lambda_{h+1} - \lambda_h| + |\mu_{h+1} - \mu_h| + |\sigma_{h+1}^2 - \sigma_h^2|$. Se $E < \epsilon$, em que $\epsilon$ é um nível de precisão pré-estabelecido, pare. Caso contrário, faça $h = h + 1$ e volte ao **Passo** 4.

### 3.3.2 Via algoritmo EM

Dempster *et al.* (1977) apresentaram o algoritmo EM (*Expectation Maximization*) como um processo iterativo para estimar parâmetros através da função de máxima verossimilhança. Este método é utilizado principalmente em problemas envolvendo dados incompletos.

Sejam $S = \sum_{i=1}^{N} X_i$, $N \sim$ Poisson Truncada $(\lambda)$, $X_i \sim N(\mu, \sigma^2)$ em que são iid e independentes de $N$. Suponha que observa-se uma amostra de tamanho $n$ da variável aleatória $S$. Logo, $S$ é observável e $N$ é uma variável aleatória não-observável, ou seja, observa-se $\mathbf{S} = \mathbf{x} = (x_1, x_2, \ldots, x_n)^\top$, mas não se observa $\mathbf{N} = \mathbf{k} = (k_1, k_2, \ldots, k_n)^\top$. Sendo o conjunto de dados completo $c = (x, k)$ o conjunto $x$ ampliado por $k$, sua função de verossimilhança é $f_{S,N}(x, k; \theta) = L^c(\theta|\mathbf{N}, \mathbf{S})$. Cada iteração do algoritmo EM é composta pelos seguintes passos:

- Passo E (Esperança): Calcule $Q(\theta|\theta_0, \underset{\sim}{S}) := E_{\theta_0}\left[\log L^c\left(\theta|\mathbf{N}, \mathbf{S}\right)\right]$, em que $E_{\theta_0}$ é o valor esperado com respeito a $\mathbf{N}|\theta_0, \mathbf{S}$ com fdp $f(\mathbf{N}|\theta, \mathbf{S})$.

- Passo M (Maximização): Encontre um $\hat{\theta}$ que maximiza $Q(\theta|\theta_0, \mathbf{S})$, ou seja,

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}\ Q(\theta|\theta_0, \mathbf{S})$$

Estes passos devem ser repetidos até se atingir uma convergência, adotando como um critério de parada $||\theta_{i+1} - \theta_i|| < \epsilon$, em que $||\cdot||$ é uma função denominada Norma, que a cada vetor de um espaço vetorial ela associa um número real não-negativo, e $\epsilon$ é um valor especificado maior que zero.

Seja $\theta = (\lambda, \mu, \sigma^2)^\top$. Deste modo, apliquemos o método. Para o Passo E, segue que:

$$Q(\theta|\theta_0, \mathbf{S}) = Q = E_{\theta_0}\left[\log L^c\left(\theta|\mathbf{N}, \mathbf{S}\right)\right] = E_{\theta_0}\left[\log f_{\mathbf{N},\mathbf{S}}\left(\mathbf{k}, \mathbf{x}|\theta\right)\right],$$

em que $E_{\theta_0}$ é o valor esperado com respeito a $\mathbf{N}|\theta_0, \mathbf{S}$. Então,

$$\begin{aligned}
Q &= \sum_{k_n=1}^{\infty}\cdots\sum_{k_1=1}^{\infty}\left\{\log\left[\prod_{i=1}^{n} f_{S|N}(x_i|k_i,\theta)\cdot P(N=k_i|\theta)\right]\right\}\cdot P(\mathbf{N}=K|\theta_0,\mathbf{S}) \\
&= \sum_{k_n=1}^{\infty}\cdots\sum_{k_1=1}^{\infty}\left\{\log\left[\prod_{i=1}^{n} f_{S_{k_i}}(x_i|\theta)\cdot P(N=k_i|\theta)\right]\right\}\prod_{j=1}^{n} P(N=k_j|\theta_0,S=x_j) \\
&= \sum_{k_n=1}^{\infty}\cdots\sum_{k_1=1}^{\infty}\left\{\sum_{i=1}^{n}\Big[\log f_{S_{k_i}}(x_i|\theta)+\log P(N=k_i|\theta)\Big]\right\}\prod_{j=1}^{n} P(N=k_j|\theta_0,S=x_j),
\end{aligned}$$

em que

$$\begin{aligned}
P(N=k_i|\theta_0,S=x_i) &= \frac{f_{S_{k_i}}(x_i|\theta_0)\cdot P(N=k_i|\theta_0)}{f_S(x_i|\theta_0)} \\
&= \frac{(2\pi k_i\sigma_0^2)^{-1/2}\exp\left[-\frac{1}{2}\frac{(x_i-k_i\mu_0)^2}{k_i{\sigma^2}_0}\right]\cdot\left(\frac{\lambda_0^{k_i}}{(e^{\lambda_0}-1)k_i!}\right)}{\sum_{j=1}^{\infty}\left(\frac{\lambda_0^{j}}{(e^{\lambda_0}-1)j!}\right)\cdot(2\pi j\sigma_0^2)^{-1/2}\exp\left[-\frac{1}{2}\frac{(x_i-j\mu_0)^2}{j\sigma_0^2}\right]}.
\end{aligned}$$

Sabendo que $f_{S_{k_i}}(x_i|\theta)=(2\pi k_i\sigma^2)^{-1/2}\exp\left[-\frac{1}{2}\frac{(x_i-k_i\mu)^2}{k_i\sigma^2}\right]$ e $P(N=k_i|\theta)=\dfrac{\lambda^{k_i}}{(e^{\lambda}-1)k_i!}$,

$$\begin{aligned}
Q &= \sum_{k_n=1}^{\infty}\cdots\sum_{k_1=1}^{\infty}\left\{\sum_{i=1}^{n}\Bigg[-\frac{1}{2}\log 2\pi-\frac{1}{2}\log\sigma^2-\frac{1}{2}\log k_i-\frac{(x_i-k_i\mu)^2}{2k_i\sigma^2}\right. \\
&\quad \left.-\log(e^{\lambda}-1)+k_i\log\lambda-\log k_i!\Bigg]\right\}\prod_{j=1}^{n} P(N=k_j|\theta_0,S=x_j).
\end{aligned}$$

No 2º passo do algoritmo, precisamos estimar o parâmetro $\theta$ que maximiza $Q(\theta|\theta_0,\mathbf{S})$, ou seja, obter

$$\hat{\theta}=\underset{\theta\in\Theta}{\operatorname{argmax}}\, Q(\theta|\theta_0,\mathbf{S}),$$

em que $\Theta$ é o espaço paramétrico.

1. Seja $Q_\lambda=\dfrac{\partial Q}{\partial\lambda}$, então

$$\begin{aligned}
\frac{\partial Q}{\partial\lambda} &= \sum_{k_n=1}^{\infty}\cdots\sum_{k_1=1}^{\infty}\left[\sum_{i=1}^{n}\left(-\frac{e^{\lambda}}{e^{\lambda}-1}+\frac{k_i}{\lambda}\right)\right]\prod_{j=1}^{n} P(N=k_j|\theta_0,S=x_j) \\
&= -\frac{n}{1-e^{-\lambda}}+\frac{1}{\lambda}\sum_{i=1}^{n}\sum_{k_n=1}^{\infty}\cdots\sum_{k_1=1}^{\infty} k_i\prod_{j=1}^{n} P(N=k_j|\theta_0,S=x_j) \\
&= -\frac{n}{1-e^{-\lambda}}+\frac{1}{\lambda}\sum_{i=1}^{n}\sum_{k_i=1}^{\infty} k_i\cdot P(N=k_i|\theta_0,S=x_i) \\
&= -\frac{n}{1-e^{-\lambda}}+\frac{1}{\lambda}\sum_{i=1}^{n} E(N|\theta_0,S=x_i). \quad \text{Assim,}
\end{aligned}$$

$$Q_\lambda\Big|_{\lambda=\hat{\lambda}} = 0 \quad \therefore \quad \frac{\hat{\lambda}}{1-e^{-\hat{\lambda}}} = \frac{1}{n}\sum_{i=1}^{n} E(N|\theta_0, S = x_i)$$

2. Seja $Q_\mu = \dfrac{\partial Q}{\partial \mu}$, então

$$\begin{aligned}
\frac{\partial Q}{\partial \mu} &= \sum_{k_n=1}^{\infty}\cdots\sum_{k_1=1}^{\infty}\left[\sum_{i=1}^{n}\frac{(x_i-k_i\mu)}{\sigma^2}\right]\prod_{j=1}^{n}P(N=k_j|\theta_0,S=x_j)\\
&= \frac{n\overline{x}}{\sigma^2} - \frac{\mu}{\sigma^2}\sum_{i=1}^{n}\sum_{k_n=1}^{\infty}\cdots\sum_{k_1=1}^{\infty}k_i\prod_{j=1}^{n}P(N=k_j|\theta_0,S=x_j)\\
&= \frac{n\overline{x}}{\sigma^2} - \frac{\mu}{\sigma^2}\sum_{i=1}^{n}\sum_{k_i=1}^{\infty}k_i\cdot P(N=k_i|\theta_0,S=x_i)\\
&= \frac{n\overline{x}}{\sigma^2} - \frac{\mu}{\sigma^2}\sum_{i=1}^{n}E(N|\theta_0,S=x_i). \quad \text{Assim,}
\end{aligned}$$

$$Q_\mu\Big|_{\mu=\hat{\mu}} = 0 \quad \therefore \quad \hat{\mu} = \frac{n\overline{x}}{\sum_{i=1}^{n}E(N|\theta_0,S=x_i)}.$$

3. Seja $Q_{\sigma^2} = \dfrac{\partial Q}{\partial \sigma^2}$, então

$$\begin{aligned}
\frac{\partial Q}{\partial \sigma^2} &= \sum_{k_n=1}^{\infty}\cdots\sum_{k_1=1}^{\infty}\left[\sum_{i=1}^{n}\left(-\frac{1}{2\sigma^2}+\frac{(x_i-k_i\mu)^2}{2k_i(\sigma^2)^2}\right)\right]\prod_{j=1}^{n}P(N=k_j|\theta_0,S=x_j)\\
&= -\frac{n}{2\sigma^2}+\frac{1}{2(\sigma^2)^2}\left\{\sum_{k_n=1}^{\infty}\cdots\sum_{k_1=1}^{\infty}\left[\sum_{i=1}^{n}\left(\frac{{x_i}^2}{k_i}-2\mu x_i+\mu^2k_i\right)\right]\right\}\\
&\quad\times\prod_{j=1}^{n}P(N=k_j|\theta_0,S=x_j)\\
&= -\frac{n}{2\sigma^2}+\frac{1}{2(\sigma^2)^2}\Bigg[-2\mu n\overline{x}+\mu^2\sum_{i=1}^{n}E(N|\theta_0,S=x_i)\\
&\quad+\sum_{i=1}^{n}\sum_{k_n=1}^{\infty}\cdots\sum_{k_1=1}^{\infty}\frac{{x_i}^2}{k_i}\Bigg]\prod_{j=1}^{n}P(N=k_j|\theta_0,S=x_j)\\
&= -\frac{n}{2\sigma^2}+\frac{1}{2(\sigma^2)^2}\Bigg[-2\mu n\overline{x}+\mu^2\sum_{i=1}^{n}E(N|\theta_0,S=x_i)\\
&\quad+\sum_{i=1}^{n}{x_i}^2\sum_{k_n=1}^{\infty}\cdots\sum_{k_1=1}^{\infty}\frac{1}{k_i}\Bigg]\prod_{j=1}^{n}P(N=k_j|\theta_0,S=x_j).
\end{aligned}$$

Assim, de $Q_{\sigma^2}\Big|_{\sigma^2=\hat{\sigma^2}} = 0$, tem-se

$$\hat{\sigma}^2 = \frac{1}{n}\left[-2\hat{\mu}n\overline{x} + \hat{\mu}^2\sum_{i=1}^{n} E(N|\theta_0, S = x_i) + \sum_{i=1}^{\infty} {x_i}^2 \sum_{k_i=1}^{\infty}\frac{1}{k_i} P(N = k_i|\theta_0, S = x_i)\right].$$

Logo,

$$\hat{\sigma}^2 = \frac{1}{n}\left[-2\hat{\mu}n\overline{x} + \hat{\mu}^2\sum_{i=1}^{n} E(N|\theta_0, S = x_i) + \sum_{i=1}^{n} {x_i}^2 E\left(\frac{1}{N}|\theta_0, S = x_i\right)\right].$$

A fim de encontrar os estimadores, utilizou-se o seguinte algoritmo.

**Algoritmo 3:** Algoritmo para o método MV via EM.

**Passo** 1. Escolha um chute inicial $\theta_0 = (\lambda_0, \mu_0, \sigma_0^2)^\top$;
**Passo** 2. Faça $k = 0$;
**Passo** 3. Encontre $\hat{\lambda}_{k+1}$ tal que $\dfrac{\lambda_{k+1}}{1 - e^{-\lambda_{k+1}}} = \frac{1}{n}\sum_{i=1}^{n} E(N|\theta_k, S = x_i)$;
**Passo** 4. Encontre $\hat{\mu}_{k+1}$ tal que $\mu_{k+1} = \dfrac{n\overline{x}}{\sum_{i=1}^{n} E(N|\theta_k, S = x_i)}$;
**Passo** 5. Encontre $\hat{\sigma}^2_{k+1}$ tal que

$$\sigma^2_{k+1} = \frac{1}{n}\left[-2\mu_{k+1}n\overline{x} + \mu^2_{k+1}\sum_{i=1}^{n} E(N|\theta_k, S = x_i) + \sum_{i=1}^{n} {x_i}^2 E\left(\frac{1}{N}|\theta_k, S = x_i\right)\right];$$

**Passo** 6. Se $|\lambda_{k+1} - \lambda_k| + |\mu_{k+1} - \mu_k| + |\sigma^2_{k+1} - \sigma^2_k| < \epsilon$, tal que $\epsilon$ é um nível de precisão pré-especificado, retorne $(\hat{\lambda}_{k+1}, \hat{\mu}_{k+1}, \hat{\sigma}^2_{k+1},)$. Caso contrário, faça $k = k + 1$ e retorne ao **Passo** 2.

## 3.4 Estudo de casos: FCE versus MV via algoritmo EM

Nessa seção é apresentado um estudo sobre a estimação dos parâmetros da distribuição CPTN utilizando os métodos FCE e máxima verossimilhança via EM. A partir das amostras 1 e 2 de tamanho $N = 100$ e $N = 150$, respectivamente, uma simulação com 100 réplicas Monte Carlo foi feita, a fim de compararmos as estimativas, o viés de cada parâmetro, o erro de ajuste e o tempo de execução de cada programa para obter estes resultados.

Foi utilizado a função `Optim()` para o Método FCE, como mencionado, e o algoritmo 3 para a estimação pelo Método MV com o próprio valor do parâmetro como chute inicial e o $\epsilon = 10^{-4}$, o tempo dado em horas (h), minutos (min) ou em segundos (s) e como casos para a análise, foram admitidos vetores de parâmetros na forma $(\lambda, \mu, \sigma)$, apresentados na Tabela 3.1, na qual é computado o número de amostras obtidas por cada método na simulação.

**Tabela 3.1:** Total de amostras com estimativas pelos métodos FCE e MV via EM na simulação.

| **CENÁRIOS** | | **Réplicas (Total** 100) | |
|---|---|---|---|
| | | **FCE** | **MV via EM** |
| **CASO 1** | $N = 100$ | 13 | 100 |
| $(5, 2, 1)$ | $N = 150$ | 80 | 100 |
| **CASO 2** | $N = 100$ | 100 | 100 |
| $(0.5, 4, 0.5)$ | $N = 150$ | 100 | 100 |
| **CASO 3** | $N = 100$ | 51 | 100 |
| $(2, 3, \sqrt{0.5})$ | $N = 150$ | 41 | 100 |

A seguir, a Tabela 3.2 traz as médias das estimações obtidas na simulação e os vieses de cada parâmetro para os três casos, em que $C = 10^{-3}$ é uma constante para auxiliar na visualização dos resultados.

**Tabela 3.2:** Comparação das estimações para o método FCE e MV via EM.

| **Estimações** | **Método FCE** | | | **Método MV via EM** | | |
|---|---|---|---|---|---|---|
| | CASO 1 | CASO 2 | CASO 3 | CASO 1 | CASO 2 | CASO 3 |
| **Amostra** 1 **(N=100)** | | | | | | |
| $\hat{\lambda}$ | 4.4832 | 0.5130 | 2.1237 | 5.0246 | 0.5093 | 2.0441 |
| $\hat{\mu}$ | 2.2565 | 4.0066 | 2.9935 | 2.0025 | 4.0033 | 3.0020 |
| $\hat{\sigma}$ | 0.9018 | 0.2546 | 0.5843 | 1.0000 | 0.5015 | 0.7093 |
| B $(\hat{\lambda})$ | -0.5168 | 0.0130 | 0.1237 | 0.0246 | 0.0093 | 0.0441 |
| B $(\hat{\mu})$ | 0.2565 | 0.0066 | -0.0065 | 0.0025 | 0.0033 | 0.0020 |
| B $(\hat{\sigma})$ | -0.0982 | -0.2454 | -0.1229 | 0.0000 | 0.0015 | 0.0022 |
| ERRO FIT | 22.14×C | 88.78×C | 4.24×C | 0.01×C | 1.30×C | 0.08×C |
| TEMPO | 27.37min | 57.36min | 45.47min | 8.10s | 5.51s | 9.00s |
| **Amostra** 2 **(N=150)** | | | | | | |
| $\hat{\lambda}$ | 4.8506 | 0.5067 | 2.0602 | 5.0217 | 0.5077 | 2.0318 |
| $\hat{\mu}$ | 2.1336 | 4.0069 | 2.9989 | 2.0020 | 4.0015 | 3.0017 |
| $\hat{\sigma}$ | 0.9333 | 0.2590 | 0.5199 | 0.9999 | 0.5042 | 0.7090 |
| B $(\hat{\lambda})$ | -0.1494 | 0.0067 | 0.0602 | 0.0217 | 0.0077 | 0.0318 |
| B $(\hat{\mu})$ | 0.1336 | 0.0069 | -0.0011 | 0.0020 | 0.0015 | 0.0017 |
| B $(\hat{\sigma})$ | -0.0667 | -0.2410 | -0.1872 | -0.0001 | 0.0042 | 0.0019 |
| ERRO FIT | 0.96×C | 80.62×C | 4.62×C | 0.009×C | 0.95×C | 0.05×C |
| TEMPO | 4.25h | 39.95min | 48.93min | 10.78s | 6.66s | 11.76s |

Como podemos observar, o método FCE, no CASO 1, segue como esperado: quando aumentamos o tamanho da amostra as estimativas dos parâmetros, seus vieses e o ajuste melhoraram. Vale ressaltar o tempo neste caso. Note que para o tamanho amostral $N = 100$, só obtivemos 13 estimativas em um pouco mais de 27 minutos, já no tamanho $N = 150$, o programa durou mais de 4 horas para gerar as 80 estimativas. No entanto, para obter as 100 réplicas para o CASO 2, o tempo diminuiu quando o tamanho da amostra aumentou. No CASO 3, há uma incoerência na estimativa do parâmetro $\sigma$, pois

o resultado obtido se distancia do verdadeiro valor com o aumento da amostra. Nessa situação, também pode ser verificado que o total de réplicas com estimativas diminuiu e o erro de ajuste aumentou.

Considerando o método MV, obtivemos as 100 réplicas para todos os cenários. Verifique que nos dois primeiros casos, a estimativa de $\sigma$ apresenta o melhor resultado no tamanho de amostra $N = 100$, apesar dos resultados serem muito próximos para ambos os tamanhos amostrais. O CASO 3, segue como previsto, pois suas estimações são mais satisfatórias quando o tamanho da amostra aumenta.

As Tabelas 3.3 e 3.4 apresentam o Critério de Informação de Akaike (AIC), o Critério de Informação de Akaike corrigido (AICc) e o Critério de Informação Bayesiano (BIC). Estes critérios são definidos por

$$AIC = -2 \cdot \log(L) + 2 \cdot k$$

$$AICc = AIC + \frac{2 \cdot k \cdot (k-1)}{(N-k-1)}$$

$$BIC = -2 \cdot \log(L) + k \cdot \log(N),$$

em que L é a função de verossimilhança, k é o número de parâmetros do modelo e N é o tamanho da amostra. De um modo geral, podemos verificar que o método EM teve a melhor performance em ambos os tamanhos de amostra e em todos os critérios, porque obteve os menores valores para estas medidas, exceto no primeiro caso para $N = 100$.

**Tabela 3.3:** Ajustes dos métodos de estimação na amostra $N = 100$.

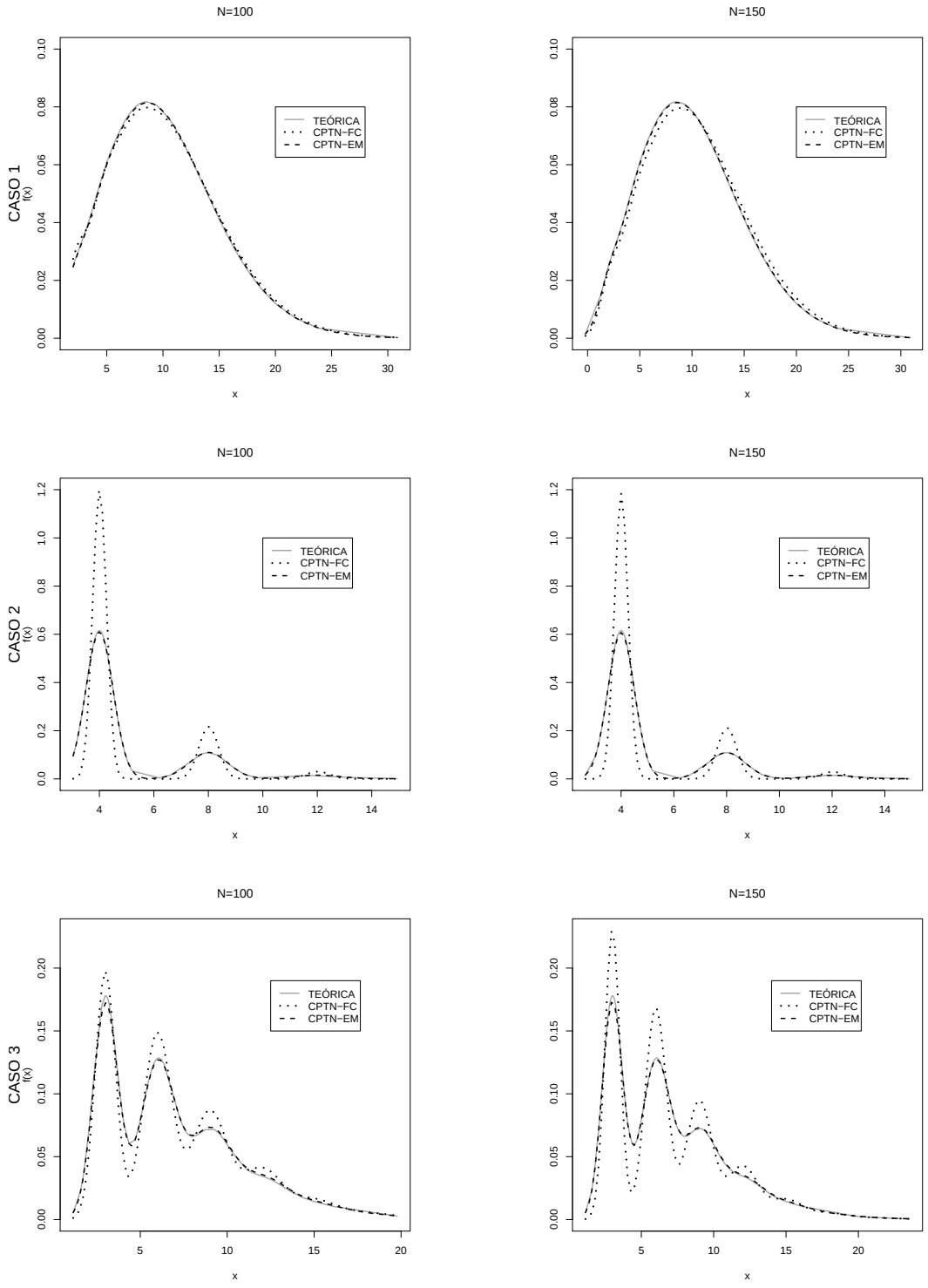
| | | **AIC** | **AICc** | **BIC** |
|---|---|---|---|---|
| CASO 1 | **FCE** | 603.57 | 611.38 | 603.82 |
| | **EM** | 603.77 | 611.58 | 604.02 |
| CASO 2 | **FCE** | 551.21 | 551.46 | 559.03 |
| | **EM** | 325.63 | 325.88 | 333.45 |
| CASO 3 | **FCE** | 573.77 | 574.02 | 581.59 |
| | **EM** | 567.24 | 567.49 | 575.05 |

**Tabela 3.4:** Ajustes dos métodos de estimação na amostra $N = 150$.

| | | **AIC** | **AICc** | **BIC** |
|---|---|---|---|---|
| CASO 1 | **FCE** | 911.26 | 911.42 | 920.29 |
| | **EM** | 911.19 | 911.35 | 920.22 |
| CASO 2 | **FCE** | 748.99 | 749.16 | 758.02 |
| | **EM** | 473.82 | 473.99 | 333.45 |
| CASO 3 | **FCE** | 843.83 | 843.99 | 852.86 |
| | **EM** | 829.51 | 829.68 | 838.54 |

A Figura 3.1 ilustra o ajuste das densidades em cada terno (método, ponto paramétrico, tamanho de amostra(N)) tal que $N = 100$ e $N = 150$, respectivamente. Dos resultados, vê-se que o método FCE não se ajusta bem na presença de multimodalidade, justificando os problemas apresentados nos casos 2 e 3. Entretanto, o método de máxima verossimilhança via algoritmo EM tem melhores ajustes exatamente nesses casos.

**Figura 3.1:** Densidades teórica e ajustadas pelos métodos FCE e MV via EM.



Do exposto, concluímos que além das estimativas do método MV via algoritmo EM e os ajustes serem mais satisfatórios, o seu custo computacional é mais vantajoso do que o do método FCE. O tempo para se executar uma estimação foi descrito na Tabela 3.2. Por esta razão, o método FCE foi excluído do estudo de simulação Monte Carlo a seguir.

## 3.5 Simulações

Nesta seção, apresentam-se resultados de simulações Monte Carlo para estudar o desempenho do método dos momentos (MM) e do método MV via algoritmo EM na estimação dos parâmetros da CPTN. Para este último, consideramos dois chutes iniciais: como primeiro chute, as estimativas do método dos momentos, a fim de comparar os resultados obtidos com os do método dos momentos; como segundo chute, o verdadeiro valor dos parâmetros, com o intuito de analisar a variabilidade das estimativas dos parâmetros. O algoritmo a seguir, foi utilizado para as simulações com pequenas e grandes amostras. Modificando, no segundo caso, o primeiro tópico do Passo 1, para os tamanhos $N_1 = 500$, $N_2 = 1000$ e $N_3 = 2000$.

**Algoritmo 4:** Algoritmo para a simulação.

**Passo 1.** Para cada $i = 1, \ldots, 1000$:
1.1. Gere $N_3 = 150$ realizações de $X \sim CPTN(\lambda, \mu, \sigma^2)$, $\theta = (\lambda, \mu, \sigma^2)^\top$, $y_{3i} = (x_1, \ldots, x_{150})^\top$.
Crie outros dois vetores de tamanhos $N_1 = 50$ e $N_2 = 100$, dados por $y_{1i} = (x_1, \ldots, x_{50})^\top$ e $y_{2i} = (x_1, \ldots, x_{100})^\top$;
1.2. Faça CONT$j = 0$, $j = 1, 2, 3$;
1.3. Com base em $y_{1i}, y_{2i}$ e $y_{3i}$, obtenha as estimativas $\hat{\theta}_{1i}, \hat{\theta}_{2i}$ e $\hat{\theta}_{3i}$;
1.4. Caso não se obtenha estimativa válida para $\theta_{ji}$, descarte a amostra e faça CONT$j$ = CONT$j$ + 1;
1.5. Registre o seguinte vetor de informações:

$$
\begin{aligned}
m_i^\top = \Bigg[ & \hat{\lambda}_{1i}, \hat{\lambda}_{2i}, \hat{\lambda}_{3i}, (\hat{\lambda}_{1i} - \lambda)^2, (\hat{\lambda}_{2i} - \lambda)^2, (\hat{\lambda}_{3i} - \lambda)^2, \hat{\mu}_{1i}, \hat{\mu}_{2i}, \hat{\mu}_{3i}, (\hat{\mu}_{1i} - \mu)^2, \\
& (\hat{\mu}_{2i} - \mu)^2, (\hat{\mu}_{3i} - \mu)^2, \hat{\sigma}^2_{1i}, \hat{\sigma}^2_{2i}, \hat{\sigma}^2_{3i}, (\hat{\sigma}^2_{1i} - \sigma^2)^2, (\hat{\sigma}^2_{2i} - \sigma^2)^2, (\hat{\sigma}^2_{3i} - \sigma^2)^2, \\
& \sum_{k=1}^{50} \frac{[f(x_k, \hat{\theta}_{1i}) - f(x_k, \theta)]^2}{50}, \sum_{k=1}^{100} \frac{[f(x_k, \hat{\theta}_{2i}) - f(x_k, \theta)]^2}{100}, \\
& \sum_{k=1}^{150} \frac{[f(x_k, \hat{\theta}_{3i}) - f(x_k, \theta)]^2}{150} \Bigg].
\end{aligned}
$$

Preencha as posições do vetor $m_i^\top$, que dependem das estimativas $\theta_{ji}$ não válidas, com NA;
**Passo 2.** $M(\theta) = [m_1 | m_2 | \ldots | m_{1000}]^\top$;
**Passo 3.** Retorne o vetor de médias das colunas da matriz $M(\theta)$ e os valores de CONT$j$.

Vale ressaltar que nessa simulação cada réplica das estimativas foi obtida aplicando os algoritmos (1) e (3) para o método dos momentos e para o método MV via EM com chute inicial igual ao parâmetro verdadeiro, respectivamente. Além disso, utilizamos as estimativas de momento como chute inicial para o MV via EM nas estimações do método MV via EM com momentos. No método MV via EM foi utilizado um nível de precisão $\epsilon = 10^{-4}$. Adotamos 1000 réplicas Monte Carlo, mas a simulação no método dos momentos apresentou problemas na estimação dos parâmetros para um certo número de réplicas, então acrescentamos um contador para que desconsiderasse a réplica que não gerasse estimativas, mas que computasse todos esses casos (Passo 1.4).

### 3.5.1 Análise da viabilidade do método dos momentos

A Tabela 3.5 contém o número total de amostras para as quais não conseguiu-se obter estimativas pelo método dos momentos. Observe que quanto maior o $\lambda$ mais difícil é de obter estimativas por este método. Note que o vetor 2, que possui o menor valor de $\lambda$, apresenta os melhores resultados para todos os tamanhos de amostra. Na variação do parâmetro $\mu$, quando comparamos os vetores 1 e 3, o total de amostras para as quais conseguiu-se obter estimativas é maior para tamanhos grandes de amostras quando o $\mu$ diminui. Já nas simulações com pequenas amostras, esse número é melhor quando o $\mu$ aumenta, ou seja, no vetor 1. Comparando os vetores 1 e 4, quando aumentamos o valor de $\sigma$ o número de amostras para as quais não conseguiu-se obter estimativas diminui nas simulações com grandes amostras e nas realizadas com pequenas amostras, esse número aumenta. Ainda pode ser conferido, nos vetores 3 e 4, que foram obtidos os mesmos resultados quando fixamos o valor de $\lambda$ e adotamos os mesmos valores para $\mu$ e $\sigma$.

**Tabela 3.5:** Total de amostras sem estimativas pelo MM na simulação Monte Carlo.

| VETOR | **Réplicas (Total** 1000) | | | | | |
|---|---|---|---|---|---|---|
| | N=50 | N=100 | N=150 | N=500 | N=1000 | N=2000 |
| **1.** $(1, 1, 0.5)$ | 335 | 357 | 361 | 435 | 423 | 443 |
| **2.** $(0.5, 1, 0.5)$ | 274 | 263 | 272 | 268 | 219 | 147 |
| **3.** $(1, 0.5, 0.5)$ | 376 | 375 | 367 | 383 | 381 | 375 |
| **4.** $(1, 1, 1)$ | 376 | 375 | 367 | 383 | 381 | 375 |

### 3.5.2 Pequenas amostras

As Tabelas 3.6, 3.7 e 3.8 apresentam informações sobre dois vetores de parâmetros cada uma, as médias das estimativas obtidas para cada parâmetro, as médias dos seus erros quadráticos médios (EQMs) e a média do erro de ajuste (ERRO FIT) em cada situação. No entanto, para o método de máxima verossimilhança via algoritmo EM utilizando as

estimativas de momentos como chute inicial, só obtivemos resultados para $N = 150$ como pode ser conferido nestas tabelas. Isto sugere que o método precisa de um bom chute inicial para convergir o que não é o caso das estimativas obtidas por momentos com pequenas amostras. Considere, ainda, $C = 10^{-3}$ uma constante para auxiliar na visualização dos resultados.

- **Efeitos da variação de $\lambda$**

A Tabela 3.6 mostra que em todos os métodos e para todos os tamanhos de amostra o erro de ajuste foi menor no caso do maior valor de $\lambda$. Neste vetor de parâmetros, as médias das estimativas de $\sigma$ são mais próximas do verdadeiro valor do parâmetro. No método dos momentos, observe para $N = 100$, que a média resultante das estimativas para $\sigma = 0.5$ foi de 0.4318 no primeiro vetor e 0.4250 no segundo. Apesar disto, a estimativa de $\mu$ apresentou menor viés e EQM e a de $\sigma$ menor EQM no caso de menor tamanho do parâmetro $\lambda$.

Os ajustes e as estimativas do método MV via EM utilizando como chute inicial o verdadeiro parâmetro foram sempre os melhores, sugerindo que com um bom chute inicial o método MV via algoritmo EM deve funcionar bem, favorecendo o uso de uma busca intensiva para encontrar o chute inicial. Ao comparar momentos com o método MV via EM com momentos, vemos que este último apresenta menor erro de ajuste e EQM, apesar de produzir estimativas um pouco mais viesadas e apresentar uma particularidade no EQM($\hat{\lambda}$), que é menor no método dos momentos do segundo vetor. Por fim, temos que tanto para momentos quanto para o MV via EM, o viés das estimativas diminui com o aumento do tamanho da amostra.

- **Efeitos da variação de $\mu$**

Na Tabela 3.7, tanto no método dos momentos quanto no MV via EM com momentos, observa-se que com a diminuição de $\mu$ no segundo vetor de parâmetros, as estimações do EQM($\hat{\mu}$), do EQM($\hat{\sigma}$) e o erro de ajuste estão melhores do que no primeiro caso, ou seja, apresentam os menores valores. Ainda assim, no caso do maior valor de $\mu$ a estimativa de $\lambda$ apresentou menor viés e maior EQM. Percebe-se também, que o MV via EM com momentos apresenta menor EQM e erro de ajuste do que o método dos momentos, mesmo obtendo a maioria das estimativas mais viesadas.

O método dos momentos mostra que a média das estimativas melhoram à medida que o tamanho da amostra aumenta, ou seja, podemos notar que para a média das estimativas dos parâmetros, os resultados obtidos se aproximam dos seus verdadeiros valores e que tanto os erros quadráticos médios quanto o erro de ajuste diminuem com o aumento da amostra. No método MV via EM, os resultados obtidos foram sempre os melhores, confirmando o

que foi analisado anteriormente. Verificamos que o EQM($\hat{\lambda}$) é mais satisfatório e a média das estimativas para $\lambda$ se aproximam do verdadeiro valor do parâmetro nos tamanhos amostrais $N = 100$ e $N = 150$ para o caso de menor valor de $\mu$. Quando aumentamos o valor de $\mu$, nota-se resultados melhores na média das estimativas de $\lambda$ para $N = 50$, na média das estimativas de $\sigma$, no EQM($\hat{\sigma}$) e no erro fit.

- **Efeitos da variação de $\sigma$**

Segundo a Tabela 3.8, o erro de ajuste foi menor em todos os métodos e para todos os tamanhos de amostra no caso de maior valor de $\sigma$. Mesmo assim, no caso de menor valor de $\sigma$, o método dos momentos mostra que a estimativa de $\mu$ apresentou menor viés, referente aos tamanhos de amostra $N = 50$ e $N = 150$, e menor EQM. E mostra ainda, o menor viés e o menor erro quadrático médio da estimativa de $\lambda$.

Os ajustes e as estimativas produzidas pelo método MV via EM apresentaram resultados mais satisfatórios quando comparados aos obtidos para os outros métodos. Verifique que os melhores resultados da média das estimativas para o parâmetro $\mu$ e da média do EQM($\hat{\mu}$) estão no primeiro vetor. Podemos observar que como o esperado, o método MV via EM utilizando as estimativas de momentos como chute inicial apresenta menor erro quadrático médio e melhor ajuste que o método dos momentos em ambos os casos. E no segundo vetor de parâmetros, apresenta a melhor média das estimativas de $\lambda$.

**Tabela 3.6:** Simulações para os métodos dos momentos e MV via EM variando o parâmetro $\lambda$.

| **Estimações** | **Momentos** | | | **Método MV via EM** | | | **MV via EM com momentos** |
|---|---|---|---|---|---|---|---|
| | 50 | 100 | 150 | 50 | 100 | 150 | 150 |
| $(\lambda, \mu, \sigma)$=$(1, 1, 0.5)$ | | | | | | | |
| $\overline{\hat{\lambda}}_1$ | 3.7537 | 2.5625 | 1.9805 | 1.0004 | 1.0035 | 1.0028 | 1.9835 |
| $\overline{\hat{\mu}}_1$ | 0.7119 | 0.7639 | 0.8198 | 0.9992 | 1.0004 | 1.0000 | 0.8176 |
| $\overline{\hat{\sigma}}_1$ | 0.4215 | 0.4318 | 0.4486 | 0.4982 | 0.4993 | 0.4997 | 0.4439 |
| EQM ($\hat{\lambda}$) | 31429.24×C | 8715.63×C | 3176.90×C | 18.25×C | 9.24×C | 6.38×C | 3169.20×C |
| EQM ($\hat{\mu}$) | 296.13×C | 212.06×C | 160.90×C | 1.22×C | 0.60×C | 0.41×C | 159.29×C |
| EQM ($\hat{\sigma}$) | 46.33×C | 37.49×C | 32.21×C | 0.89×C | 0.45×C | 0.32×C | 29.55×C |
| ERRO FIT | 6.38×C | 4.56×C | 3.82×C | 1.04×C | 0.51×C | 0.36×C | 3.74×C |
| $(\lambda, \mu, \sigma)$=$(0.5, 1, 0.5)$ | | | | | | | |
| $\overline{\hat{\lambda}}_1$ | 3.1773 | 2.2076 | 1.4996 | 0.4943 | 0.4970 | 0.4978 | 1.5076 |
| $\overline{\hat{\mu}}_1$ | 0.7693 | 0.8079 | 0.8575 | 0.9974 | 0.9989 | 0.9990 | 0.8545 |
| $\overline{\hat{\sigma}}_1$ | 0.4130 | 0.4250 | 0.4472 | 0.4940 | 0.4966 | 0.4977 | 0.4431 |
| EQM ($\hat{\lambda}$) | 41989.26×C | 23338.42×C | 6261.44×C | 9.67×C | 4.78×C | 3.39×C | 6277.76×C |
| EQM ($\hat{\mu}$) | 232.15×C | 176.82×C | 136.60×C | 2.39×C | 1.15×C | 0.80×C | 136.24×C |
| EQM ($\hat{\sigma}$) | 43.05×C | 36.01×C | 29.31×C | 1.31×C | 0.67×C | 0.46×C | 27.04×C |
| ERRO FIT | 7.52×C | 5.28×C | 4.33×C | 2.11×C | 1.00×C | 0.69×C | 4.07×C |

**Tabela 3.7:** Simulações para os métodos dos momentos e MV via EM variando o parâmetro $\mu$.

| **Estimações** | **Momentos** | | | **Método MV via EM** | | | **MV via EM com momentos** |
|---|---|---|---|---|---|---|---|
| | 50 | 100 | 150 | 50 | 100 | 150 | 150 |
| $(\lambda, \mu, \sigma)$=(1, 1, 0.5) | | | | | | | |
| $\overline{\hat{\lambda}}_1$ | 3.7537 | 2.5625 | 1.9805 | 1.0004 | 1.0035 | 1.0028 | 1.9835 |
| $\overline{\hat{\mu}}_1$ | 0.7119 | 0.7639 | 0.8198 | 0.9992 | 1.0004 | 1.0000 | 0.8176 |
| $\overline{\hat{\sigma}}_1$ | 0.4215 | 0.4318 | 0.4486 | 0.4982 | 0.4993 | 0.4997 | 0.4439 |
| EQM ($\hat{\lambda}$) | 31429.24×C | 8715.63×C | 3176.90×C | 18.25×C | 9.24×C | 6.38×C | 3169.20×C |
| EQM ($\hat{\mu}$) | 296.13×C | 212.06×C | 160.90×C | 1.22×C | 0.60×C | 0.41×C | 159.29×C |
| EQM ($\hat{\sigma}$) | 46.33×C | 37.49×C | 32.21×C | 0.89×C | 0.45×C | 0.32×C | 29.55×C |
| ERRO FIT | 6.38×C | 4.56×C | 3.82×C | 1.04×C | 0.51×C | 0.36×C | 3.74×C |
| $(\lambda, \mu, \sigma)$=(1, 0.5, 0.5) | | | | | | | |
| $\overline{\hat{\lambda}}_1$ | 5.2139 | 3.5661 | 3.0519 | 1.0016 | 1.0029 | 1.0023 | 3.0515 |
| $\overline{\hat{\mu}}_1$ | 0.3409 | 0.3893 | 0.3902 | 0.4995 | 0.5008 | 0.5001 | 0.3898 |
| $\overline{\hat{\sigma}}_1$ | 0.3847 | 0.4212 | 0.4262 | 0.4957 | 0.4978 | 0.4992 | 0.4241 |
| EQM ($\hat{\lambda}$) | 63842.76×C | 28472.62×C | 17855.82×C | 9.83×C | 5.02×C | 3.45×C | 17838.22×C |
| EQM ($\hat{\mu}$) | 96.05×C | 74.75×C | 66.61×C | 2.09×C | 1.05×C | 0.72×C | 66.41×C |
| EQM ($\hat{\sigma}$) | 43.53×C | 32.78×C | 28.24×C | 1.40×C | 0.67×C | 0.48×C | 28.11×C |
| ERRO FIT | 4.11×C | 2.53×C | 1.92×C | 1.78×C | 0.87×C | 0.60×C | 1.90×C |

**Tabela 3.8:** Simulações para os métodos dos momentos e MV via EM variando o parâmetro $\sigma$.

| **Estimações** | **Momentos** | | | **Método MV via EM** | | | **MV via EM com momentos** |
|---|---|---|---|---|---|---|---|
| | 50 | 100 | 150 | 50 | 100 | 150 | 150 |
| $(\lambda, \mu, \sigma)$=(1, 1, 0.5) | | | | | | | |
| $\overline{\hat{\lambda}}_1$ | 3.7537 | 2.5625 | 1.9805 | 1.0004 | 1.0035 | 1.0028 | 1.9835 |
| $\overline{\hat{\mu}}_1$ | 0.7119 | 0.7639 | 0.8198 | 0.9992 | 1.0004 | 1.0000 | 0.8176 |
| $\overline{\hat{\sigma}}_1$ | 0.4215 | 0.4318 | 0.4486 | 0.4982 | 0.4993 | 0.4997 | 0.4439 |
| EQM ($\hat{\lambda}$) | 31429.24×C | 8715.63×C | 3176.90×C | 18.25×C | 9.24×C | 6.38×C | 3169.20×C |
| EQM ($\hat{\mu}$) | 296.13×C | 212.06×C | 160.90×C | 1.22×C | 0.60×C | 0.41×C | 159.29×C |
| EQM ($\hat{\sigma}$) | 46.33×C | 37.49×C | 32.21×C | 0.89×C | 0.45×C | 0.32×C | 29.55×C |
| ERRO FIT | 6.38×C | 4.56×C | 3.82×C | 1.04×C | 0.51×C | 0.36×C | 3.74×C |
| $(\lambda, \mu, \sigma)$=(1, 1, 1) | | | | | | | |
| $\overline{\hat{\lambda}}_1$ | 5.1427 | 3.5661 | 3.0519 | 1.0016 | 1.0029 | 1.0023 | 3.0515 |
| $\overline{\hat{\mu}}_1$ | 0.6830 | 0.7785 | 0.7804 | 0.9991 | 1.0016 | 1.0002 | 0.7797 |
| $\overline{\hat{\sigma}}_1$ | 0.7704 | 0.8424 | 0.8525 | 0.9914 | 0.9956 | 0.9984 | 0.8482 |
| EQM ($\hat{\lambda}$) | 60370.11×C | 28472.62×C | 17855.82×C | 9.83×C | 5.02×C | 3.45×C | 17838.22×C |
| EQM ($\hat{\mu}$) | 383.69×C | 299.00×C | 266.45×C | 8.35×C | 4.20×C | 2.87×C | 265.63×C |
| EQM ($\hat{\sigma}$) | 173.56×C | 131.12×C | 112.94×C | 5.60×C | 2.70×C | 1.91×C | 112.46×C |
| ERRO FIT | 1.03×C | 0.63×C | 0.48×C | 0.45×C | 0.22×C | 0.15×C | 0.47×C |

### 3.5.3 Grandes amostras

Dando continuidade à análise das simulações, as Tabelas 3.9, 3.10 e 3.11 apresentam informações, citadas anteriormente, para os tamanhos de amostra 500, 1000 e 2000.

- **Efeitos da variação de $\lambda$**

Os resultados mais satisfatórios da Tabela 3.9 comparando as estimações de cada vetor por método, são comentados a seguir. No primeiro vetor: os melhores resultados para o método dos momentos e o MV via EM com momentos, foram em $N = 2000$, as médias das estimativas de $\mu$ e de $\sigma$; para o MV via EM, as médias das estimativas de $\mu$, das estimativas de $\sigma$ nos tamanhos $N = 500$ e $N = 1000$, as médias do EQM($\hat{\mu}$), do EQM($\hat{\sigma}$) e do erro de ajuste. No segundo vetor: para o método MV via EM, a média das estimativas de $\sigma$ para o tamanho $N = 2000$; enquanto que para momentos e MV via EM com momentos, quando o valor de $\lambda$ diminui o erro de ajuste e o EQM das estimativas diminuem.

Além disso, quando aumentamos o tamanho da amostra, no segundo vetor, o método dos momentos apresenta estimativas mais viesadas para $N = 2000$. O método MV via EM apresenta a mesma situação para as estimativas de $\lambda$ e de $\mu$. E o MV via EM com momentos para as estimativas de $\mu$ e de $\sigma$. No caso de maior valor de $\lambda$, quando o tamanho da amostra aumenta, o método MV via EM apresenta esse tipo de resultado para $\lambda$ em $N = 1000$ e para $\sigma$ no tamanho $N = 2000$. Já no método dos momentos e MV via EM com momentos, à medida que o tamanho da amostra aumenta as estimativas melhoram. Ainda assim, o método MV via EM utilizando como chute inicial o verdadeiro valor dos parâmetros produziu os resultados mais satisfatórios.

Comparando o método dos momentos com o MV via EM com momentos, o primeiro vetor mostra que as estimativas do MV via EM com momentos são mais viesadas, porém o ajuste e o EQM são melhores, pois apresentam menores valores. Como esperado, o MV via EM com momentos no segundo caso apresenta menor erro de ajuste e menor erro quadrático médio para quase todos os casos, pois o EQM ($\hat{\lambda}$) para $N = 500$ e $N = 1000$ são melhores no método dos momentos.

- **Efeitos da variação de $\mu$**

A Tabela 3.10 exibe os resultados obtidos nas simulações quando variamos o parâmetro $\mu$. Observa-se que as melhores estimações para as médias das estimativas de $\lambda$, das estimativas de $\sigma$ e do EQM($\hat{\lambda}$) no método dos momentos e no MV via EM com momentos, encontram-se no primeiro vetor. Sendo assim, a média do EQM($\hat{\sigma}$) e do erro de ajuste são melhores no segundo.

Analisando o método MV via EM, o primeiro vetor proporcionou melhores médias das estimativas de $\lambda$ nos casos $N = 500$ e $N = 2000$, melhores médias das estimativas de $\sigma$, do EQM($\hat{\sigma}$) e do erro de ajuste. Ao mesmo tempo que as médias restantes referentes ao parâmetro $\lambda$ são melhores no segundo vetor. Ao comparar o método MV via EM com momentos com o método dos momentos, o EQM e o erro de ajuste são menores no primeiro método para ambos os vetores, mas apresenta dois casos particulares: para $N = 1000$ e $N = 2000$, a média do EQM($\hat{\lambda}$) é menor no método dos momentos. Ainda pode ser notado, que conforme o tamanho da amostra aumenta, o viés das estimativas diminui.

- **Efeitos da variação de $\sigma$**

A Tabela 3.11 mostra que o erro de ajuste é menor para todos os métodos no caso de maior valor de $\sigma$. Note que nesse caso, o viés da estimativa de $\lambda$ para $N = 1000$, a média das estimativas de $\mu$ (nos tamanhos de amostra $N = 500$ e $N = 2000$) e do EQM($\hat{\lambda}$) melhoraram no método MV via EM. Apesar disso, o método dos momentos e o MV via EM com momentos produziram melhores médias das estimativas para $\lambda$ e $\mu$, diminuição do EQM($\hat{\lambda}$) e do EQM($\hat{\mu}$) no primeiro vetor. Enquanto que para o MV via EM, este vetor traz melhores resultados para as médias das estimativas de $\lambda$ (nos tamanhos amostrais $N = 500$ e $N = 2000$), das estimativas de $\mu$ quando $N = 1000$ e da média do EQM($\hat{\mu}$).

No caso de menor valor de $\sigma$, os valores obtidos no EQM($\hat{\lambda}$) são menores para o método dos momentos do que os obtidos para o método MV via EM com momentos nos tamanhos de amostra $N = 1000$ e $N = 2000$. Observe que com o aumento do tamanho amostral, temos melhores estimativas tanto no método dos momentos quanto no MV via EM com momentos .

Por fim, notamos que:

a) Como esperado para uma análise de simulação, ocorre diminuição do erro quadrático médio e do erro de ajuste em todos os métodos e em cada caso, quando o tamanho da amostra aumenta;

b) Utilizar as estimativas de momentos como ponto de partida para a estimação por máxima verossimilhança via algoritmo EM, os resultados foram, na sua maioria, aceitáveis.

**Tabela 3.9:** Simulações para os métodos dos momentos e MV via EM variando o parâmetro $\lambda$.

| **Estimações** | **Momentos** | | | **Método MV via EM** | | | **MV via EM com momentos** | | |
|---|---|---|---|---|---|---|---|---|---|
| | 500 | 1000 | 2000 | 500 | 1000 | 2000 | 500 | 1000 | 2000 |
| $(\lambda, \mu, \sigma)$=$(1, 1, 0.5)$ | | | | | | | | | |
| $\overline{\hat{\lambda}}_1$ | 1.3945 | 1.2244 | 1.0908 | 1.0002 | 0.9994 | 0.9999 | 1.3993 | 1.2290 | 1.0955 |
| $\overline{\hat{\mu}}_1$ | 0.9199 | 0.9536 | 0.9925 | 0.9998 | 0.9998 | 1.0001 | 0.9174 | 0.9513 | 0.9901 |
| $\overline{\hat{\sigma}}_1$ | 0.4803 | 0.4884 | 0.5060 | 0.5002 | 0.5000 | 0.4998 | 0.4767 | 0.4865 | 0.5034 |
| EQM $(\hat{\lambda})$ | 884.80×C | 441.69×C | 256.94×C | 2.05×C | 0.92×C | 0.46×C | 882.87×C | 439.54×C | 254.08×C |
| EQM $(\hat{\mu})$ | 83.75×C | 53.72×C | 39.08×C | 0.14×C | 0.06×C | 0.03×C | 82.58×C | 52.74×C | 38.10×C |
| EQM $(\hat{\sigma})$ | 23.54×C | 17.17×C | 14.50×C | 0.10×C | 0.05×C | 0.02×C | 20.27×C | 14.33×C | 11.83×C |
| ERRO FIT | 2.66×C | 2.06×C | 1.72×C | 0.11×C | 0.05×C | 0.03×C | 2.58×C | 1.98×C | 1.64×C |
| $(\lambda, \mu, \sigma)$=$(0.5, 1, 0.5)$ | | | | | | | | | |
| $\overline{\hat{\lambda}}_1$ | 0.6404 | 0.5146 | 0.4810 | 0.4992 | 0.4999 | 0.4990 | 0.6474 | 0.5192 | 0.4836 |
| $\overline{\hat{\mu}}_1$ | 0.9860 | 1.0072 | 1.0114 | 0.9996 | 0.9996 | 0.9994 | 0.9831 | 1.0052 | 1.0101 |
| $\overline{\hat{\sigma}}_1$ | 0.5014 | 0.5088 | 0.5096 | 0.4989 | 0.4998 | 0.5000 | 0.4972 | 0.5058 | 0.5078 |
| EQM $(\hat{\lambda})$ | 616.08×C | 139.29×C | 33.25×C | 0.96×C | 0.49×C | 0.25×C | 622.28×C | 141.13×C | 32.98×C |
| EQM $(\hat{\mu})$ | 36.56×C | 13.58×C | 4.49×C | 0.24×C | 0.12×C | 0.06×C | 36.27×C | 13.38×C | 4.32×C |
| EQM $(\hat{\sigma})$ | 11.37×C | 5.86×C | 2.29×C | 0.13×C | 0.07×C | 0.03×C | 9.55×C | 4.67×C | 1.66×C |
| ERRO FIT | 2.09×C | 1.31×C | 0.67×C | 0.19×C | 0.10×C | 0.05×C | 1.89×C | 1.13×C | 0.55×C |

**Tabela 3.10:** Simulações para os métodos dos momentos e MV via EM variando o parâmetro $\mu$.

| **Estimações** | **Momentos** | | | **Método MV via EM** | | | **MV via EM com momentos** | | |
|---|---|---|---|---|---|---|---|---|---|
| | 500 | 1000 | 2000 | 500 | 1000 | 2000 | 500 | 1000 | 2000 |
| $(\lambda, \mu, \sigma)$=$(1, 1, 0.5)$ | | | | | | | | | |
| $\overline{\hat{\lambda}}_1$ | 1.3945 | 1.2244 | 1.0908 | 1.0002 | 0.9994 | 0.9999 | 1.3993 | 1.2290 | 1.0955 |
| $\overline{\hat{\mu}}_1$ | 0.9199 | 0.9536 | 0.9925 | 0.9998 | 0.9998 | 1.0001 | 0.9174 | 0.9513 | 0.9901 |
| $\overline{\hat{\sigma}}_1$ | 0.4803 | 0.4884 | 0.5060 | 0.5002 | 0.5000 | 0.4998 | 0.4767 | 0.4865 | 0.5034 |
| EQM $(\hat{\lambda})$ | 884.80×C | 441.69×C | 256.94×C | 2.05×C | 0.92×C | 0.46×C | 882.87×C | 439.54×C | 254.08×C |
| EQM $(\hat{\mu})$ | 83.75×C | 53.72×C | 39.08×C | 0.14×C | 0.06×C | 0.03×C | 82.58×C | 52.74×C | 38.10×C |
| EQM $(\hat{\sigma})$ | 23.54×C | 17.17×C | 14.50×C | 0.10×C | 0.05×C | 0.02×C | 20.27×C | 14.33×C | 11.83×C |
| ERRO FIT | 2.66×C | 2.06×C | 1.72×C | 0.11×C | 0.05×C | 0.03×C | 2.58×C | 1.98×C | 1.64×C |
| $(\lambda, \mu, \sigma)$=$(1, 0.5, 0.5)$ | | | | | | | | | |
| $\overline{\hat{\lambda}}_1$ | 1.9688 | 1.5860 | 1.2466 | 1.0005 | 0.9995 | 0.9996 | 1.9701 | 1.5875 | 1.2482 |
| $\overline{\hat{\mu}}_1$ | 0.4274 | 0.4525 | 0.4860 | 0.5000 | 0.4998 | 0.5000 | 0.4270 | 0.4522 | 0.4857 |
| $\overline{\hat{\sigma}}_1$ | 0.4553 | 0.4715 | 0.4921 | 0.4997 | 0.4996 | 0.4995 | 0.4536 | 0.4704 | 0.4909 |
| EQM $(\hat{\lambda})$ | 3981.37×C | 1967.46×C | 863.60×C | 1.05×C | 0.51×C | 0.26×C | 3979.34×C | 1968.23×C | 864.67×C |
| EQM $(\hat{\mu})$ | 39.06×C | 26.81×C | 15.82×C | 0.23×C | 0.10×C | 0.05×C | 38.97×C | 26.75×C | 15.78×C |
| EQM $(\hat{\sigma})$ | 15.60×C | 10.30×C | 6.02×C | 0.15×C | 0.08×C | 0.04×C | 15.34×C | 10.03×C | 5.78×C |
| ERRO FIT | 0.87×C | 0.60×C | 0.41×C | 0.18×C | 0.09×C | 0.05×C | 0.86×C | 0.60×C | 0.41×C |

**Tabela 3.11:** Simulações para os métodos dos momentos e MV via EM variando o parâmetro $\sigma$.

| **Estimações** | **Momentos** | | | **Método MV via EM** | | | **MV via EM com momentos** | | |
|---|---|---|---|---|---|---|---|---|---|
| | 500 | 1000 | 2000 | 500 | 1000 | 2000 | 500 | 1000 | 2000 |
| $(\lambda, \mu, \sigma)$=(1, 1, 0.5) | | | | | | | | | |
| $\overline{\hat{\lambda}}_1$ | 1.3945 | 1.2244 | 1.0908 | 1.0002 | 0.9994 | 0.9999 | 1.3993 | 1.2290 | 1.0955 |
| $\overline{\hat{\mu}}_1$ | 0.9199 | 0.9536 | 0.9925 | 0.9998 | 0.9998 | 1.0001 | 0.9174 | 0.9513 | 0.9901 |
| $\overline{\hat{\sigma}}_1$ | 0.4803 | 0.4884 | 0.5060 | 0.5002 | 0.5000 | 0.4998 | 0.4767 | 0.4865 | 0.5034 |
| EQM ($\hat{\lambda}$) | 884.80×C | 441.69×C | 256.94×C | 2.05×C | 0.92×C | 0.46×C | 882.87×C | 439.54×C | 254.08×C |
| EQM ($\hat{\mu}$) | 83.75×C | 53.72×C | 39.08×C | 0.14×C | 0.06×C | 0.03×C | 82.58×C | 52.74×C | 38.10×C |
| EQM ($\hat{\sigma}$) | 23.54×C | 17.17×C | 14.50×C | 0.10×C | 0.05×C | 0.02×C | 20.27×C | 14.33×C | 11.83×C |
| ERRO FIT | 2.66×C | 2.06×C | 1.72×C | 0.11×C | 0.05×C | 0.03×C | 2.58×C | 1.98×C | 1.64×C |
| $(\lambda, \mu, \sigma)$=(1, 1, 1) | | | | | | | | | |
| $\overline{\hat{\lambda}}_1$ | 1.9688 | 1.5860 | 1.2466 | 1.0005 | 0.9995 | 0.9996 | 1.9701 | 1.5875 | 1.2482 |
| $\overline{\hat{\mu}}_1$ | 0.8548 | 0.9050 | 0.9720 | 1.0000 | 0.9996 | 1.0000 | 0.8540 | 0.9043 | 0.9713 |
| $\overline{\hat{\sigma}}_1$ | 0.9106 | 0.9431 | 0.9841 | 0.9995 | 0.9992 | 0.9990 | 0.9072 | 0.9407 | 0.9818 |
| EQM ($\hat{\lambda}$) | 3981.37×C | 1967.46×C | 863.60×C | 1.05×C | 0.51×C | 0.26×C | 3979.34×C | 1968.23×C | 864.66×C |
| EQM ($\hat{\mu}$) | 156.26×C | 107.22×C | 63.28×C | 0.93×C | 0.42×C | 0.22×C | 155.88×C | 107.01×C | 63.12×C |
| EQM ($\hat{\sigma}$) | 62.39×C | 41.19×C | 24.08×C | 0.61×C | 0.31×C | 0.15×C | 61.36×C | 40.11×C | 23.13×C |
| ERRO FIT | 0.22×C | 0.15×C | 0.10×C | 0.05×C | 0.02×C | 0.01×C | 0.22×C | 0.15×C | 0.10×C |

## 3.6 Aplicação

Nesta seção, apresentamos uma aplicação a uma imagem SAR (*Synthetic Aperture Radar* - SAR) extraída da imagem de Foulum (Dinamarca) e obtida pelo sensor EMISAR (LEE; POTTIER, 2009), construído pelo *ElectroMagnetics Institute* (EMI) na *Technical University of Denmark*. A distribuição proposta, CPTN, foi utilizada para descrever um atributo associado ao um pixel de uma imagem SAR, que é chamado de *intensidade* como explicado na discussão a seguir.

O sistema SAR tem sido indicado como uma ferramenta eficiente de sensoriamento remoto. Este fato se deve: (i) à capacidade de tais sistemas de operarem independentemente da luminosidade e das condições climáticas e (ii) à capacidade de produzir imagens com alta resolução espacial (LEE; POTTIER, 2009). Entretanto, as imagens produzidas pelo sistema SAR são fortemente contaminadas por um tipo particular de interferência, chamada de *ruído speckle*. O ruído *speckle* danifica visualmente a imagem de modo a comprometer a interpretabilidade da imagem SAR. Assim, a proposta de uma modelagem especializada consiste em uma importante etapa de pré-processamento.

O dado associado a um *pixel* de uma imagem SAR, diga-se $X(i,j)$ representando a entrada "$(i,j)$" de uma imagem, é formado obedecendo à seguinte dinâmica: *Um pulso polarizado linearmente (nas direções vertical, 'V', e horizontal, 'H') é emitido a uma região geográfica que se quer sensoriar e o seu retorno é capturado segundo uma direção (V ou H). Ao final do processo, um* pixel *de uma imagem SAR concentra a informação de quatro canais de polarização na forma de quatro números complexos, diga-se* $(Z_{HH}, Z_{HV}, Z_{VH}, Z_{VV})^\top$.

A área que explora o tratamento de todas as informações conjuntamente é conhecida como processamento de dados PolSAR (*Polarimetric SAR*-SAR) (LEE; POTTIER, 2009). Nesta aplicação, objetiva-se trabalhar com a modelagem de um atributo de um canal de polarização, o que é conhecido na literatura como processamento de dados SAR. Um dado SAR pode ser representado como $X(i,j) = |X(i,j)| \exp\{i\,\phi(i,j)\}$, em que $|X(i,j)| = \sqrt{\Re[X(i,j)]^2 + \Im[X(i,j)]^2}$ e $\phi(i,j)$ representam a amplitude e a fase associadas a um pixel $X(i,j)$. Nesta aplicação, os dados são da forma $|X(i,j)|^2$ e denotados como *intensidades*, atributo muito utilizado na literatura (NASCIMENTO *et al.*, 2010).

O conjunto de dados apresentado na Tabela 3.12 é composto por 131 valores de intensidades referentes ao canal de polarização HH, de uma imagem da região de Foulum (Dinamarca). A Tabela 3.13 apresenta medidas descritivas para os dados de intensidade. A diferença aproximada absoluta entre média e mediana de 3.648% do intervalo de variação dos dados é suficiente para caracterizar os dados como assimétricos e a desigualdade mediana $<$ média aponta para uma densidade populacional com cauda a direita ou assimetria positiva, o que é confirmado pela assimetria positiva. O coeficiente de variação

(em %) de 50.04% pode ser entendido como um indicativo de um conjunto de dados fortemente heterogêneo. Finalmente, a curtose positiva aponta para uma densidade com cauda longa e pesada, o que é conhecido como distribuições *leptocúrticas*.

**Tabela 3.12:** Conjunto de dados de Foulum.

| | | | | | |
|---|---|---|---|---|---|
| 0.01894287 | 0.01935418 | 0.02396332 | 0.02647153 | 0.02082554 | 0.01911615 |
| 0.02293175 | 0.03228633 | 0.03474018 | 0.03384829 | 0.02407404 | 0.02296903 |
| 0.02721631 | 0.03141554 | 0.03752478 | 0.02784315 | 0.04168270 | 0.04891223 |
| 0.02856023 | 0.02839673 | 0.03598676 | 0.04629390 | 0.05547253 | 0.05395810 |
| 0.03879451 | 0.02941171 | 0.02338895 | 0.02777823 | 0.04783208 | 0.03824020 |
| 0.03232158 | 0.03402714 | 0.05611064 | 0.07915612 | 0.03754972 | 0.03101604 |
| 0.03501506 | 0.03453897 | 0.05705985 | 0.05515739 | 0.04647379 | 0.06612180 |
| 0.05260086 | 0.02942552 | 0.03763689 | 0.05229843 | 0.06204657 | 0.08059592 |
| 0.05190274 | 0.02861129 | 0.04111884 | 0.06560329 | 0.06084920 | 0.03424702 |
| 0.02792585 | 0.04164984 | 0.05508701 | 0.09429808 | 0.09665179 | 0.08470160 |
| 0.08028217 | 0.05686707 | 0.04585500 | 0.04230256 | 0.03445653 | 0.04063699 |
| 0.04832320 | 0.06187658 | 0.06816631 | 0.05449910 | 0.05385335 | 0.04675854 |
| 0.02833025 | 0.01992964 | 0.02185369 | 0.02363688 | 0.02051137 | 0.02317232 |
| 0.02260664 | 0.02931118 | 0.02775543 | 0.01802599 | 0.01570864 | 0.01149687 |
| 0.01250427 | 0.01528162 | 0.02017052 | 0.02358412 | 0.02585229 | 0.03893850 |
| 0.04768521 | 0.07643031 | 0.05274564 | 0.02667463 | 0.03024757 | 0.03427167 |
| 0.04638059 | 0.08566783 | 0.10034920 | 0.07746060 | 0.07138552 | 0.05092829 |
| 0.10605280 | 0.12579180 | 0.09340851 | 0.07589816 | 0.04526951 | 0.06653712 |
| 0.06442861 | 0.03224484 | 0.04694617 | 0.03803634 | 0.02665159 | 0.02892729 |
| 0.03578667 | 0.05177016 | 0.04639203 | 0.05562469 | 0.06370583 | 0.05983544 |
| 0.05941767 | 0.04971199 | 0.04367883 | 0.04341392 | 0.03989288 | 0.04538165 |
| 0.04278920 | 0.06403044 | 0.12067840 | 0.10924980 | 0.05840242 | |

Fonte: https://earth.esa.int/web/polsar/home

**Tabela 3.13:** Descrição do conjunto de dados de Foulum.

| Média | Desvio Padrão | Mediana | Mínimo | Máximo | $C_V$ | Assimetria | Curtose |
|---|---|---|---|---|---|---|---|
| 0.04582 | 0.02293 | 0.04165 | 0.01150 | 0.12580 | 0.50045 | 1.16726 | 4.32383 |

Para esse conjunto de dados, comparamos o ajuste da distribuição prosposta, CPTN, aos ajustes das seguintes distribuições:

- Normal, com fdp
$$f(x) = (2\pi\sigma^2)^{-1/2} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right],$$
em que $x \in \mathbb{R}$;

- Weibull, com fdp
$$f(x) = \frac{kx^{k-1}}{\lambda^k} \exp[-(x/\lambda)^k],$$
em que $x > 0$;

- Weibull-Poisson (LU; SHI, 2012) com fdp

$$f(x) = \frac{\theta k x^{k-1} \exp[\theta(1 - \exp[-(x/\lambda)^k]) - (x/\lambda)^k]}{\lambda^k(e^\theta - 1)},$$

em que $x > 0$;

- Kumaraswamy-Weibull Poisson (RAMOS *et al.*, 2015) com fdp

$$f(x) = \frac{\lambda abc\beta^c}{e^\lambda - 1} x^{c-1}[1 - e^{-(\beta x)^c}]^{a-1}\{1 - [1 - e^{-(\beta x)^c}]^a\}^{b-1} \times \exp[\lambda\{1 - [1 - e^{-(\beta x)^c}]^a\}^b - (\beta x)^c],$$

em que $x > 0$.

As estimativas de máxima verossimilhança (MV) para os parâmetros da distribuição Normal e da Weibull foram obtidas através da função `fitdistr()` do pacote `MASS` do `R`. Enquanto que para as distribuições Weibull-Poisson (WPoisson) e Kumaraswamy-Weibull Poisson (Kw-WPoisson) foi utilizada a função `goodness.fit()` do pacote `AdequacyModel` também presente na plataforma `R`.

A estimação dos parâmetros da distribuição proposta foi feita pelo método dos momentos (CPTN-MM), MV via EM com momentos (CPTN-EM.MM) e pelo método MV via EM com busca intensiva (CPTN-EM.BI). Para este, fizemos uso de duas situações: Na primeira busca intensiva (CPTN-EM.BI-1), condicionamos a estimação ao critério de informação de Akaike (AIC), em que primeiro chutou-se um valor de $\mu$ de modo que a primeira moda dos dados fosse aproximadamente igual a $\mu$ e a segunda moda aproximadamente igual a $2 \cdot \mu$. Depois, testou-se valores de $\lambda$ e $\sigma$ nos intervalos $[0.1, 10]$ e $[0.007, 0.008]$ com variação entre os pontos de 0.01 e 0.0001, respectivamente; na segunda busca intensiva (CPTN-EM.BI-2), a estimação foi condicionada ao teste Kolmogorov-Sminorv (KS). Essa medida se refere a distância entre a densidade empírica e a densidade ajustada aos dados. Nesse caso, adotamos para o parâmetro $\mu$ valores entre as suas estimativas anteriores, ou seja, entre 0.02 e 0.04. Em seguida, testamos valores de $\lambda$ no intervalo $[0.001, 0.1]$ variando 0.0001 entre os pontos e para $\sigma$, valores entre $[0.1, 5]$ com variação de 0.1.

A Tabela 3.14 mostra as estimativas dos parâmetros de cada modelo e seus respectivos erros padrão. Além disso, a média e o desvio padrão dos modelos são computados. Para a distribuição CPTN, como a matriz de informação observada ficou em função de muitos somatórios com infinitos termos (o que inviabilizou seu uso para obtenção da variância assintótica), os erros padrão foram obtidos utilizando a técnica de leave-one-out (JAIN *et al.*, 1987). Deste modo, podemos verificar que todas as estimativas foram estatisticamente diferentes de zero por teste para parâmetros individuais (i.e., do tipo $\mathcal{H}_0 \colon \theta_i = 0$) via intervalos de confiança assintóticos.

**Tabela 3.14:** Parâmetros estimados ao conjunto de dados de Foulum.

| MODELO | PARÂMETRO | ESTIMATIVA | ERRO PADRÃO | MÉDIA | DESVIO PADRÃO |
|---|---|---|---|---|---|
| CPTN-MM | $\lambda$ | 0.2454 | $0.84\times10^{-2}$ | 0.0457 | 0.0229 |
| | $\mu$ | 0.0406 | $0.02\times10^{-2}$ | | |
| | $\sigma$ | 0.0164 | $0.02\times10^{-2}$ | | |
| CPTN-EM.MM | $\lambda$ | 0.2656 | $0.41\times10^{-2}$ | 0.0458 | 0.0227 |
| | $\mu$ | 0.0402 | $0.09\times10^{-2}$ | | |
| | $\sigma$ | 0.0157 | $0.05\times10^{-3}$ | | |
| CPTN-EM.BI-1 | $\lambda$ | 1.2314 | $0.86\times10^{-2}$ | 0.0458 | 0.0261 |
| | $\mu$ | 0.0263 | $0.02\times10^{-3}$ | | |
| | $\sigma$ | 0.0072 | $0.02\times10^{-3}$ | | |
| CPTN-EM.BI-2 | $\lambda$ | 0.4959 | $0.01\times10^{-5}$ | 0.0458 | 0.0251 |
| | $\mu$ | 0.0361 | $0.02\times10^{-2}$ | | |
| | $\sigma$ | 0.0141 | $0.02\times10^{-2}$ | | |
| Normal-MV | $\mu$ | 0.0458 | 0.0020 | 0.0458 | 0.0229 |
| | $\sigma$ | 0.0229 | 0.0014 | | |
| Weibull-MV | k | 2.1397 | 0.1370 | 0.0461 | 0.0227 |
| | $\lambda$ | 0.0520 | 0.0022 | | |
| WPoisson-MV | $\theta$ | $1.0\times10^{-10}$ | $2.74\times10^{-17}$ | 17.0439 | 8.3871 |
| | $k$ | 2.1390 | $1.14\times10^{-2}$ | | |
| | $\lambda$ | 19.2452 | $7.49\times10^{-2}$ | | |
| Kw-WPoisson-MV | a | 13.8932 | 0.7668 | 0.0163 | 0.0304 |
| | b | 31.9847 | 5.2516 | | |
| | c | 0.3865 | 0.0875 | | |
| | $\lambda$ | 2.2244 | 0.2064 | | |
| | $\beta$ | 43.4708 | 5.8499 | | |

A fim de comparar quantitativamente os modelos elencados na Tabela 3.14, utilizamos como figuras de mérito as seguintes medidas: Critério de Informação de Akaike (AIC), Versão corrigida do AIC (AICc), Critério de Informação Bayesiano (BIC) e o teste Kolmogorov-Sminorv (KS).

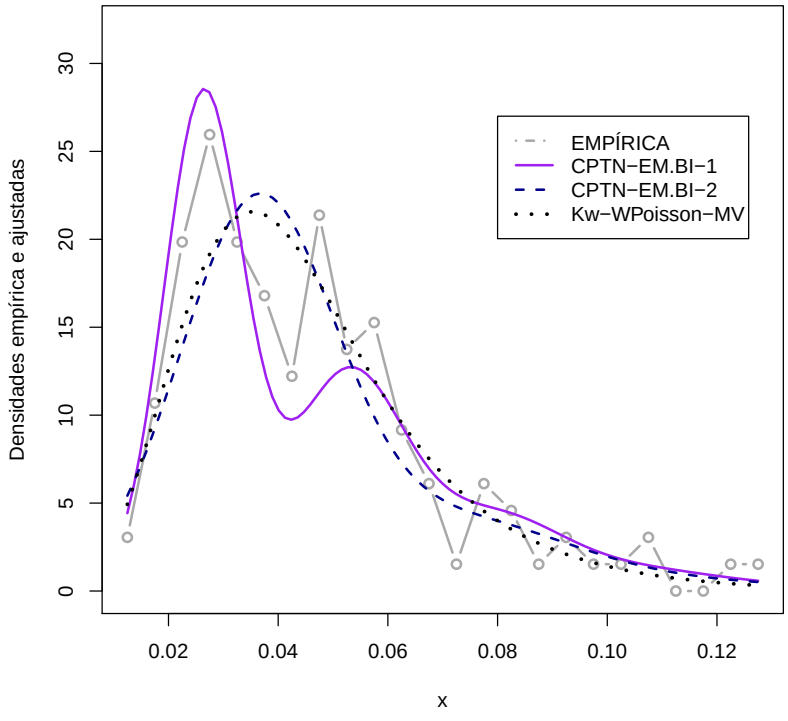
Para estas medidas, vale-se a regra padrão "os melhores modelos são os associados às menores medidas". A Tabela 3.15 apresenta os resultados dos ajustes por estas medidas para os dados de intensidade. Os resultados apresentam evidências de que nas medidas de bondade de ajuste envolvendo densidades (AICc e BIC) o modelo CPTN-EM.BI-1 superou todos os outros modelos, contudo no AIC, a distribuição Kw-WPoisson se destacou. No teste KS, representado com as medidas da distância entre as densidades e seus respectivos p-valores, podemos notar que os modelos CPTN-EM.BI-2 e Kw-WPoisson-MV apresentaram os menores valores para essas distâncias. No entanto, a distribuição Composta Poisson-Truncada Normal apresenta a melhor adequação do ajuste. Outros pontos para comparação entre essas duas distribuições, são apresentados nas Tabelas 3.13 e 3.14. Observe que para todos os métodos de estimação da distribuição CPTN, os valores estimados para a média e o desvio padrão são mais próximos dos valores empíricos do que os obtidos para a distribuição Kw-WPoisson.

**Tabela 3.15:** Bondade de ajuste.

| | AIC | AICc | BIC | KS (p-valor) |
|---|---|---|---|---|
| CPTN-MM | -637.8932 | -637.7042 | -629.2676 | 0.0793 (0.3823) |
| CPTN-EM.MM | -639.0794 | -638.8905 | -630.4539 | 0.0824 (0.3367) |
| CPTN-EM.BI-1 | -643.4457 | **-643.2567** | **-634.8201** | 0.0985 (0.1570) |
| CPTN-EM.BI-2 | -640.9153 | -640.7264 | -632.2897 | **0.0494** (0.9060) |
| N-MV | -614.3466 | -614.2528 | -608.5962 | 0.0913 (0.2252) |
| W-MV | -635.7022 | -635.6085 | -629.9518 | 0.0708 (0.5272) |
| WPoisson-MV | -633.7022 | -633.5133 | -625.0766 | 0.0709 (0.5252) |
| Kw-WPoisson-MV | **-643.5756** | -643.0956 | -629.1996 | 0.0560 (0.8061) |

Diante disso, apresentamos na Figura 3.2 as densidades da CPTN-EM.BI-1, CPTN-EM.BI-2 e da Kw-WPoisson-MV ajustadas ao banco de dados. Note que a distribuição Composta Poisson-Truncada Normal, utilizando o resultado da busca intensiva 1 como chute inicial para a estimação dos parâmetros, tem um ajuste melhor que a Kumaraswamy-Weibull Poisson, mesmo esta possuindo 2 parâmetros a mais que a distribuição proposta.

**Figura 3.2:** Densidades ajustadas aos dados.



# CAPÍTULO 4

## Conclusões e sugestões para trabalhos futuros

## 4.1 Conclusões

Nesta dissertação, foi apresentado um ensaio teórico aplicado da família composta $N$. Em particular, um novo gerador de distribuições através da sub-família Composta Poisson-Truncada (CPT) foi introduzido e o caso especial CPT Normal (CPTN) foi detalhado. Analisamos algumas propriedades e derivamos que o modelo CPTN consiste de uma mistura com apenas três parâmetros que pode descrever distribuições empíricas multimodais de alta complexidade.

Apresentamos um capítulo com três métodos de estimação para os parâmetros da distribuição Composta Poisson-Truncada Normal, propondo algoritmos de estimação dos parâmetros para cada método. Deste modo, um estudo de simulação Monte Carlo foi realizado a fim de quantificar a performance do método dos momentos, da estimação por função característica empírica e do método de máxima verossimilhança via algoritmo EM.

Dentre os métodos abordados, realizamos um estudo de casos sobre a estimação dos parâmetros da CPTN utilizando o método FCE e o MV via algoritmo EM para dois tamanhos amostrais. O método da FCE requer uma função de ponderação específica para produzir um estimador eficiente. No caso da distribuição proposta, utilizamos a função de ponderação aplicada para a distribuição Normal em Yu (2004). As estimativas dos parâmetros obtidas pelos métodos foram próximas aos seus verdadeiros valores, mas a estimação por FCE apresentou menos de 50% de amostras com estimativas em alguns cenários da simulação. Além disso, o seu custo computacional é menos vantajoso que o do método de máxima verossimilhança via algoritmo EM. A partir desses resultados, realizamos simulações com tamanhos pequenos e grandes de amostra utilizando, apenas, o método dos momentos e o MV via algoritmo EM.

Uma análise da viabilidade do método dos momentos foi feita, na qual verificamos que quanto maior o valor do parâmetro $\lambda$ mais difícil obtermos as estimativas por este método. Os resultados mostraram que analisando a variação de $\lambda$, nas simulações com grandes amostras, o total de amostras sem estimativas passou de 40% com o aumento do valor deste parâmetro.

No método de máxima verossimilhança, aplicamos dois chutes iniciais. Como primeiro chute, utilizamos os valores encontrados na estimação por momentos, para comparar com as estimativas obtidas pelo método dos momentos. Na simulação com pequenas amostras, não foi possível obter estimativas para alguns tamanhos amostrais pelo método MV via algoritmo EM neste cenário. O segundo chute foi o verdadeiro vetor de parâmetros, com o intuito de analisar a variabilidade desses parâmetros. Em geral, o método de máxima verossimilhança via algoritmo EM condicionado a um bom chute inicial se mostrou o melhor procedimento de estimação.

Finalmente, uma aplicação a processamento de imagens SAR foi realizada. O conjunto de dados requer uma certa flexibilidade, sendo visualmente caracterizado por um mínimo de três modas. Nesta aplicação, o modelo CPTN foi comparado com outros quatro modelos, Normal, Weibull, Weibull-Poisson e Kumaraswamy-Weibull Poisson. Dentre as medidas de bondade de ajuste consideradas, o modelo CPTN utilizando o método MV via algoritmo EM, com o mecanismo de busca intensiva para encontrar um chute inicial, superou os demais no AICc, no BIC e no teste Kolmogorov-Sminorv. Como uma conclusão, o modelo CPTN parece ser um bom descritor de intensidade em cenários de imagens de radar, caracterizados por várias regiões.

## 4.2 Sugestões para trabalhos futuros

Como trabalhos futuros, destacamos duas frentes:

- Estimação por função característica por discretização de seu suporte;
- Derivação de mais propriedades estatísticas para o modelo CPTN e de um modelo de regressão CPTN.

# Referências


BARNETT, V. *Comparative Statistical Inference*. Baffins Lane, Chichester, West Sussex, England: John Wiley & Sons, 1999.

BOLFARINE, H.; SANDOVAL, M. C. *Introdução à inferência estatística*. Rio de Janeiro, RJ, Brasil: SBM, 2010.

CARRASCO, M. H.; FLORENS, J.-P. *et al.* Estimation of a mixture via the empirical characteristic function. In: ECONOMETRIC SOCIETY. *Econometric Society World Congress 2000 Contributed Papers*. [S.l.], 2000.

CASELLA, G.; BERGER, R. *Statistical Inference*. United States of America: Duxbury Pacific Grove, CA, 2002.

COBB, L. Stochastic catastrophe models and multimodal distributions. *Behavioral Science*, Wiley Online Library, v. 23, n. 4, p. 360–374, 1978.

DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, JSTOR, p. 1–38, 1977.

EL-ZAART, A.; ZIOU, D. Statistical modelling of multimodal sar images. *Int. J. Remote Sens.*, Taylor & Francis, Inc., Bristol, PA, USA, v. 28, n. 10, p. 2277–2294, maio 2007. ISSN 0143-1161.

EVERITT, B. S.; HAND, D. J. *et al. Finite mixture distributions*. London: Chapman and Hall, 1981.

FEUERVERGER, A.; MCDUNNOUGH, P. On the efficiency of empirical characteristic function procedures. *Journal of the Royal Statistical Society. Series B (Methodological)*, JSTOR, p. 20–27, 1981.

FINE, T. L. *Probability and Probabilistic Reasoning for Electrical Engineering*. [S.l.]: Prentice Hall, 2006.

GOLUBEV, A. Exponentially modified gaussian (emg) relevance to distributions related to cell proliferation and differentiation. *Journal of theoretical biology*, Elsevier, v. 262, n. 2, p. 257–266, 2010.

GRUSHKA, E. Characterization of exponentially modified gaussian peaks in chromatography. *Analytical Chemistry*, ACS Publications, v. 44, n. 11, p. 1733–1738, 1972.

HEATHCOTE, C. The integrated squared error estimation of parameters. *Biometrika*, Biometrika Trust, v. 64, n. 2, p. 255–264, 1977.

JAIN, A. K.; DUBES, R. C.; CHEN, C.-C. Bootstrap techniques for error estimation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-9, n. 5, p. 628–633, Sept 1987.

KARLIS, D.; XEKALAKI, E. Mixed poisson distributions. *International Statistical Review*, Wiley Online Library, v. 73, n. 1, p. 35–58, 2005.

KNIGHT, J. L.; YU, J. Empirical characteristic function in time series estimation. *Econometric Theory*, Cambridge Univ Press, v. 18, n. 03, p. 691–721, 2002.

LAWLESS, J. F. *Statistical models and methods for lifetime data*. Hoboken, New Jersey: John Wiley & Sons, 2003.

LEE, J.-S.; POTTIER, E. *Polarimetric radar imaging: from basics to applications*. Boca Ranton, London, New York: CRC press, 2009.

LU, W.; SHI, D. A new compounding life distribution: the weibull–poisson distribution. *Journal of applied statistics*, Taylor & Francis, v. 39, n. 1, p. 21–38, 2012.

NASCIMENTO, A. D. C.; CINTRA, R. J.; FRERY, A. C. Hypothesis testing in speckled data with stochastic distances. *IEEE Transactions on Geoscience and Remote Sensing*, v. 48, p. 373–385, 2010.

NAVARRETE, V. E. d. L. S. *Um novo método para compor distribuições: uma análise da classe G-Poisson*. Dissertação (Mestrado) — Universidade Federal de Pernambuco, 2013.

PANJER, H. H. Recursive evaluation of a family of compound distributions. *Astin Bulletin*, Cambridge Univ Press, v. 12, n. 01, p. 22–26, 1981.

PATIL, G. On certain compound poisson and compound binomial distributions. *Sankhya*, v. 26, p. 293–294, 1964.

PAULSON, A. S.; HOLCOMB, E. W.; LEITCH, R. A. The estimation of the parameters of the stable laws. *Biometrika*, Biometrika Trust, v. 62, n. 1, p. 163–170, 1975.

PRESS, S. J. Estimation in univariate and multivariate stable distributions. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 67, n. 340, p. 842–846, 1972.

QUANDT, R. E.; RAMSEY, J. B. Estimating mixtures of normal distributions and switching regressions. *Journal of the American statistical Association*, Taylor & Francis, v. 73, n. 364, p. 730–738, 1978.

RAMOS, M. W. A. *et al.* "the kumaraswamy-g poisson family of distributions". *A aparecer em Journal of Statistical Theory and Applications*, 2015.

RESNICK, S. I. *Adventures in stochastic processes*. Boston: Birkhauser, 1992.

REVFEIM, K. An initial model of the relationship between rainfall events and daily rainfalls. *Journal of Hydrology*, Elsevier, v. 75, n. 1, p. 357–364, 1984.

SHANNON, C. E. A mathematical theory of communication. *Bell System Technical Journal*, v. 27, p. 379–432, 1948.

TEICH, M. C.; DIAMENT, P. Multiply stochastic representations for k distributions and their poisson transforms. *JOSA A*, Optical Society of America, v. 6, n. 1, p. 80–91, 1989.

THOMPSON, C. Homogeneity analysis of rainfall series: an application of the use of a realistic rainfall model. *Journal of climatology*, Wiley Online Library, v. 4, n. 6, p. 609–619, 1984.

WIRJANTO, T. S.; XU, D. *The Applications of Mixtures of Normal Distributions in Empirical Finance: A Selected Survey*. [S.l.], 2009.

YU, J. Empirical characteristic function estimation and its applications. *Econometric reviews*, Taylor & Francis, v. 23, n. 2, p. 93–123, 2004.

# APÊNDICE A

Apêndice

## A.1 Cálculo dos momentos através da função característica

Recorde que a função característica de uma v.a. $S$ com distribuição Composta $N$ é dada por $\varphi_S(t) = \varphi_N(-i \log \varphi_X(t))$. Nesta seção do apêndice, iremos calcular $E(S)^j = \frac{\varphi_S^{(j)}(0)}{i^j}$ em termos dos momentos de $N$ e $X$, para $j = 1, 2 \text{ e } 3$.

A seguir, serão apresentadas as equações para o cálculo dos 3 primeiros momentos.

1. **Primeiro Momento**

$$\begin{aligned}
E(S) &= \frac{\varphi_S^{(1)}(0)}{i} \\
&= \frac{\varphi_N^{(1)}(-i \log \varphi_X(t))}{i} \left( \frac{-i\varphi_X^{(1)}(t)}{\varphi_X(t)} \right) \Bigg|_{t=0} \\
&= \frac{\varphi_N^{(1)}(-i \log \varphi_X(0))}{i} \left( \frac{-i\varphi_X^{(1)}(0)}{\varphi_X(0)} \right) \\
&= \frac{\varphi_N^{(1)}(-i \log 1)}{i} \cdot \left( \frac{-i\varphi_X^{(1)}(0)}{1} \right) \\
&= \frac{\varphi_N^{(1)}(0)}{i} \cdot (-i\varphi_X^{(1)}(0)) \\
&= E(N) \cdot E(X).
\end{aligned}$$

2. **Segundo Momento**

$$
\begin{aligned}
E(S^2) &= \frac{\varphi_S^{(2)}(0)}{i^2} \\
&= \frac{1}{i^2}\left\{\left[\varphi_N^{(2)}(-i\log\,\varphi_X(t))\cdot\left(\frac{-i\varphi_X^{(1)}(t)}{\varphi_X(t)}\right)\cdot(-i\varphi_X^{(1)}(t))\right.\right. \\
&\quad \left.+(-i\varphi_X^{(2)}(t)\cdot\varphi_N^{(1)}(-i\log\,\varphi_X(t))\right]\times\varphi_X(t) \\
&\quad \left.-\varphi_X^{(1)}(t)\cdot(\varphi_N^{(1)}(-i\log\varphi_X(t))\cdot(-i\varphi_X^{(1)}(t)))\right\}\times\frac{1}{(\varphi_X(t))^2}\Bigg|_{t=0} \\
&= \frac{1}{(i\varphi_X(0))^2}\cdot\left\{\left[\varphi_N^{(2)}(-i\log\,\varphi_X(0))\cdot\left(\frac{-i\varphi_X^{(1)}(0)}{\varphi_X(0)}\right)\cdot(-i\varphi_X^{(1)}(0))\right.\right. \\
&\quad \left.+(-i\varphi_X^{(2)}(0)\cdot\varphi_N^{(1)}(-i\log\,\varphi_X(0))\right]\times\varphi_X(0) \\
&\quad \left.-\varphi_X^{(1)}(0)\cdot(\varphi_N^{(1)}(-i\log\varphi_X(0))\cdot(-i\varphi_X^{(1)}(0)))\right\} \\
&= \frac{\varphi_N^{(2)}(0)}{i^2}\cdot\left(-i\varphi_X^{(1)}(0)\right)^2+\frac{\varphi_X^{(2)}(0)}{i^2}\cdot(-i\varphi_N^{(1)}(0))+\frac{[\varphi_X^{(1)}(0)]^2}{i^2}\cdot(i\varphi_N^{(1)}(0)) \\
&= E(N^2)\cdot[E(X)]^2+E(X^2)\cdot E(N)-[E(X)]^2\cdot E(N) \\
&= E(N^2)\cdot[E(X)]^2+E(N)\cdot[E(X^2)-[E(X)]^2] \\
&= E(N^2)\cdot[E(X)]^2+E(N)\cdot Var(X).
\end{aligned}
$$

3. **Terceiro Momento**

$$E(S^3) = \frac{\varphi_S^{(3)}(0)}{i^3}$$

Para isso, precisamos calcular $\varphi_S^{(3)}(t)$, como:

$$
\begin{aligned}
\varphi_S^{(2)}(t) &= \left[\varphi_N^{(2)}(-i\log\,\varphi_X(t))\cdot\left(\frac{i\varphi_X^{(1)}(t)}{\varphi_X(t)}\right)^2\right] \\
&\quad +\left[\frac{(-i\varphi_X^{(2)}(t)\cdot\varphi_N^{(1)}(-i\log\,\varphi_X(t))}{\varphi_X(t)}\right] \\
&\quad +\left[\frac{i\varphi_N^{(1)}(-i\log\varphi_X(t))\cdot(\varphi_X^{(1)}(t))^2}{(\varphi_X(t))^2}\right]
\end{aligned}
$$

Desta forma $\varphi_S^{(3)}(t) = \mathrm{A} + \mathrm{B} + \mathrm{C}$, em que:

$$\mathrm{A} = \frac{\partial\left[\varphi_N^{(2)}(-i\log\,\varphi_X(t))\cdot\left(\frac{i\varphi_X^{(1)}(t)}{\varphi_X(t)}\right)^2\right]}{\partial t},\ \mathrm{B} = \frac{\partial\left[\frac{(-i\varphi_X^{(2)}(t)\cdot\varphi_N^{(1)}(-i\log\,\varphi_X(t))}{\varphi_X(t)}\right]}{\partial t}$$

$$\text{e } \mathrm{C} = \frac{\partial\left[\frac{i\varphi_N^{(1)}(-i\log\varphi_X(t))\cdot(\varphi_X^{(1)}(t))^2}{(\varphi_X(t))^2}\right]}{\partial t}.$$

Portanto,

$$E(S^3) = \frac{\varphi_S^{(3)}(0)}{i^3} = \frac{[\mathrm{A}+B+\mathrm{C}]|_{t=0}}{i^3}.$$

Calculando as parcelas, temos:

$$\begin{aligned}
\mathrm{A} &= \left\{\left[\varphi_N^{(3)}(-i\log\,\varphi_X(t))\cdot\left(\frac{-i\varphi_X^{(1)}(t)}{\varphi_X(t)}\right)\cdot(i\varphi_X^{(1)}(t))^2\right.\right. \\
&\quad \left.+2i^2\varphi_X^{(1)}(t)\cdot\varphi_X^{(2)}(t)\cdot\varphi_N^{(2)}(-i\log\,\varphi_X(t))\right]\times[\varphi_X(t)]^2 \\
&\quad \left.-2\varphi_X(t)\cdot\varphi_X^{(1)}(t)\cdot\varphi_N^{(2)}(-i\log\,\varphi_X(t))\cdot[i\varphi_X^{(1)}(t)]^2\right\}\times\frac{1}{(\varphi_X(t))^4} \\
&= \frac{\varphi_N^{(3)}(-i\log\,\varphi_X(t))\cdot(-i\varphi_X^{(1)}(t))^3}{(\varphi_X(t))^3} \\
&\quad +\frac{2i^2\varphi_N^{(2)}(-i\log\,\varphi_X(t))\cdot\varphi_X^{(1)}(t)\cdot\varphi_X^{(2)}(t)}{(\varphi_X(t))^2} \\
&\quad -\frac{2i^2\varphi_N^{(2)}(-i\log\,\varphi_X(t))\cdot(\varphi_X^{(1)}(t))^3}{(\varphi_X(t))^3}
\end{aligned}$$

Então,

$$\begin{aligned}
\frac{\mathrm{A}|_{t=0}}{i^3} &= \frac{\varphi_N^{(3)}(0))\cdot(-i\varphi_X^{(1)}(0))^3}{i^3} \\
&\quad +\frac{2i^2\varphi_N^{(2)}(0)\cdot\varphi_X^{(1)}(0)\cdot\varphi_X^{(2)}(0)}{i^3} \\
&\quad -\frac{2i^2\varphi_N^{(2)}(0))\cdot(\varphi_X^{(1)}(0))^3}{i^3} \\
&= E(N^3)\cdot[E(X)]^3 + 2\cdot E(N^2)\cdot E(X)\cdot E(X^2) - 2\cdot E(N^2)\cdot[E(X)]^3
\end{aligned}$$

Calculando B:

$$\begin{aligned}
\mathrm{B} &= \left\{ -i\left[\varphi_X^{(3)}(t)\cdot\varphi_N^{(1)}(-i\log\,\varphi_X(t))\right.\right. \\
&\quad +\varphi_N^{(2)}(-i\log\,\varphi_X(t))\cdot\left(\frac{-i\varphi_X^{(1)}(t)}{\varphi_X(t)}\right)\cdot\varphi_X^{(2)}(t)\Bigg]\times[\varphi_X(t)] \\
&\quad \left.\left.-\varphi_X^{(1)}(t)[-i\varphi_X^{(2)}(t)\cdot\varphi_N^{(1)}(-i\log\,\varphi_X(t))]\right\}\times\frac{1}{(\varphi_X(t))^2}\right. \\
&= \frac{-i\varphi_N^{(1)}(-i\log\,\varphi_X(t))\cdot\varphi_X^{(3)}(t)}{\varphi_X(t)} \\
&\quad +\frac{i^2\varphi_N^{(2)}(-i\log\,\varphi_X(t))\cdot\varphi_X^{(2)}(t)\cdot\varphi_X^{(1)}(t)}{(\varphi_X(t))^2} \\
&\quad -\frac{(-i\varphi_N^{(1)}(-i\log\,\varphi_X(t))\cdot\varphi_X^{(2)}(t)\cdot(\varphi_X^{(1)}(t))}{(\varphi_X(t))^2}
\end{aligned}$$

Segue,

$$\begin{aligned}
\frac{\mathrm{B}|_{t=0}}{i^3} &= \frac{-i\varphi_N^{(1)}(0)\cdot\varphi_X^{(3)}(0)}{i^3} \\
&\quad +\frac{i^2\varphi_N^{(2)}(0)\cdot\varphi_X^{(2)}(0)\cdot\varphi_X^{(1)}(0)}{i^3} \\
&\quad -\frac{(-i\varphi_N^{(1)}(0))\cdot\varphi_X^{(2)}(0)\cdot(\varphi_X^{(1)}(0))}{i^3} \\
&= E(N)\cdot E(X^3)+E(N^2)\cdot E(X^2)\cdot E(X)-E(N)\cdot E(X^2)\cdot E(X)
\end{aligned}$$

Analogamente para C, temos:

$$\begin{aligned}
\mathrm{C} &= \left\{ i\left[\varphi_N^{(2)}(-i\log\,\varphi_X(t))\cdot\left(\frac{-i\varphi_X^{(1)}(t)}{\varphi_X(t)}\right)\cdot(\varphi_X^{(1)}(t))^2\right.\right. \\
&\quad \left.+2\varphi_X^{(1)}(t)\cdot\varphi_X^{(2)}(t)\cdot(\varphi_N^{(1)}(-i\log\,\varphi_X(t)))\right]\times[\varphi_X(t)]^2 \\
&\quad \left.-2i\varphi_X(t)\cdot\varphi_X^{(1)}(t)\cdot(\varphi_N^{(1)}(-i\log\,\varphi_X(t))\cdot(\varphi_X^{(1)}(t))^2\right\}\times\frac{1}{(\varphi_X(t))^4}
\end{aligned}$$

Então,

$$\begin{aligned}\frac{\mathrm{C}|_{t=0}}{i^3} &= \frac{-i^2\varphi_N^{(2)}(0)\cdot(\varphi_X^{(1)}(0))^3}{i^3} \\ &\quad +\frac{2i\varphi_N^{(1)}(0)\cdot\varphi_X^{(2)}(0)\cdot\varphi_X^{(1)}(0)}{i^3} \\ &\quad -\frac{2i\varphi_N^{(1)}(0)\cdot(\varphi_X^{(1)}(0))^3}{i^3} \\ &= -E(N^2)\cdot[E(X)]^3 - 2\cdot E(N)\cdot E(X^2)\cdot E(X) + 2\cdot E(N)[E(X)]^3\end{aligned}$$

Logo,

$$\begin{aligned}E(S^3) &= E(N^3)\cdot[E(X)]^3 + 2\cdot E(N^2)\cdot E(X)\cdot E(X^2) - 2\cdot E(N^2)\cdot[E(X)]^3 \\ &\quad +E(N)\cdot E(X^3) + E(N^2)\cdot E(X)\cdot E(X^2) - E(N)\cdot E(X)\cdot E(X^2) \\ &\quad -E(N^2)\cdot[E(X)]^3 - 2\cdot E(N)\cdot E(X^2)\cdot E(X) + 2\cdot E(N)[E(X)]^3 \\ &= E(N^3)\cdot[E(X)]^3 - 3\cdot E(N^2)\cdot[E(X)]^3 + 3\cdot E(N^2)\cdot E(X)\cdot E(X^2) \\ &\quad -3\cdot E(N)\cdot E(X)\cdot E(X^2) + E(N)\cdot E(X^3) + 2\cdot E(N)[E(X)]^3\end{aligned}$$

## A.2 Cálculo para a estimação dos parâmetros por FCE

- Distribuição CPTN.

Na Seção 3.2, vimos que $D(\theta; x)$ é a distância entre a função característica empírica e a função característica teórica. Nesta seção do apêndice, iremos encontrar uma expressão que será utilizada para obtermos as estimativas dos parâmetros pelo método da função característica empírica. Recorde que utilizamos

$$D(\theta; x) = \int_{-\infty}^{\infty} |\varphi_n(r) - \varphi(r)|^2 \, \exp(-r^2)\, dr$$

Deste modo, considerando a função característica da distribuição CPTN, temos que:

$$D(\theta; x) = \int_{-\infty}^{\infty} \left| \frac{1}{n} \sum_{j=1}^{n} e^{irX_j} - \left( \frac{e^{\exp[i\mu r - \sigma^2 r^2/2]\lambda} - 1}{e^{\lambda} - 1} \right) \right|^2 \exp(-r^2)\, dr.$$

$$= \int_{-\infty}^{\infty} \left[ \left( \frac{1}{n} \sum_{j=1}^{n} e^{irX_j} - \frac{e^{\lambda \exp[i\mu r - \sigma^2 r^2/2]} - 1}{e^{\lambda} - 1} \right) \cdot \left( \frac{1}{n} \sum_{j=1}^{n} e^{-irX_t} - \frac{e^{\lambda \exp[-i\mu r - \sigma^2 r^2/2]} - 1}{e^{\lambda} - 1} \right) \right] \times \exp(-r^2)\, dr$$

$$= \int_{-\infty}^{\infty} \left[ \frac{1}{n^2} \sum_{j=1}^{n} \sum_{t=1}^{n} e^{ir(X_j - X_t)} - \frac{1}{n} \sum_{j=1}^{n} e^{irX_j} \left( \frac{e^{\lambda \exp[-i\mu r - \sigma^2 r^2/2]} - 1}{e^{\lambda} - 1} \right) - \frac{1}{n} \sum_{t=1}^{n} e^{-irX_t} \left( \frac{e^{\lambda \exp[-i\mu r - \sigma^2 r^2/2]} - 1}{e^{\lambda} - 1} \right) + \frac{(e^{\lambda \exp[i\mu r - \sigma^2 r^2/2]} - 1) \cdot (e^{\lambda \exp[-i\mu r - \sigma^2 r^2/2]} - 1)}{(e^{\lambda} - 1)^2} \right] \times \exp(-r^2)\, dr$$

$$
\begin{aligned}
=&\frac{1}{n^2}\sum_{j=1}^{n}\sum_{t=1}^{n}\left(\int_{-\infty}^{\infty} e^{ir(X_j-X_t)-r^2}\,dr\right)\\
&-\frac{1}{n(e^{\lambda}-1)}\times\left[\sum_{j=1}^{n}\left(\int_{-\infty}^{\infty} e^{irX_j}\cdot e^{\lambda\exp[-i\mu r-\sigma^2r^2/2]}\cdot e^{-r^2}\,dr\right)\right.\\
&-\sum_{j=1}^{n}\left(\int_{-\infty}^{\infty} e^{irX_j-r^2}\,dr\right)+\sum_{t=1}^{n}\left(\int_{-\infty}^{\infty} e^{-irX_t}\cdot e^{\lambda\exp[i\mu r-\sigma^2r^2/2]}\cdot e^{-r^2}\,dr\right)\\
&\left.-\sum_{t=1}^{n}\left(\int_{-\infty}^{\infty} e^{-irX_t-r^2}\,dr\right)\right]\\
&+\frac{1}{(e^{\lambda}-1)^2}\times\left(\int_{-\infty}^{\infty}\left[e^{\lambda\exp[i\mu r-\sigma^2r^2/2]}\cdot e^{\lambda\exp[-i\mu r-\sigma^2r^2/2]}\right.\right.\\
&\left.\left.-e^{\lambda\exp[i\mu r-\sigma^2r^2/2]}-e^{\lambda\exp[-i\mu r-\sigma^2r^2/2]}+1\right]e^{-r^2}\,dr\right)
\end{aligned}
$$

$$
\begin{aligned}
=&\frac{\pi^{1/2}}{n^2}\sum_{j=1}^{n}\sum_{t=1}^{n}\left(e^{-\frac{1}{4}(X_j-X_t)^2}\right)-\frac{1}{n(e^{\lambda}-1)}\times\\
&\left[\sum_{j=1}^{n}\sum_{k=0}^{\infty}\left(\int_{-\infty}^{\infty} e^{irX_j}\cdot\frac{[\lambda\exp(-i\mu r-\sigma^2r^2/2)]^k}{k!}\cdot e^{-r^2}\,dr\right)\right.\\
&+\sum_{j=1}^{n}\sum_{k=0}^{\infty}\left(\int_{-\infty}^{\infty} e^{-irX_t}\cdot\frac{[\lambda\exp(i\mu r-\sigma^2r^2/2)]^k}{k!}\cdot e^{-r^2}\,dr\right)\\
&\left.-\sum_{j=1}^{n}\left(\sqrt{\pi}\exp\left\{-\frac{1}{4}X_j^2\right\}\right)-\sum_{t=1}^{n}\left(\sqrt{\pi}\exp\left\{-\frac{1}{4}X_t^2\right\}\right)\right]\\
&+\frac{1}{(e^{\lambda}-1)^2}\times\int_{-\infty}^{\infty}\left[e^{\lambda(\exp[i\mu r-\sigma^2r^2/2]+\exp[-i\mu r-\sigma^2r^2/2])}\right.\\
&\left.-\sum_{k=0}^{\infty}\frac{[\lambda\exp(i\mu r-\sigma^2r^2/2)]^k}{k!}-\sum_{k=0}^{\infty}\frac{[\lambda\exp(-i\mu r-\sigma^2r^2/2)]^k}{k!}+1\right]\cdot e^{-r^2}\,dr
\end{aligned}
$$

$$
\begin{aligned}
=&\frac{\pi^{1/2}}{n^2}\sum_{j=1}^{n}\sum_{t=1}^{n}\left(e^{-\frac{1}{4}(X_j-X_t)^2}\right)-\frac{1}{n(e^{\lambda}-1)}\times\\
&\left[\sum_{j=1}^{n}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left(\int_{-\infty}^{\infty}\exp[irX_j-r^2-(i\mu r+\sigma^2r^2/2)k]\,dr\right)\right.\\
&+\sum_{t=1}^{n}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left(\int_{-\infty}^{\infty}\exp[-irX_t-r^2+(i\mu r-\sigma^2r^2/2)k]\,dr\right)\\
&\left.-2\sqrt{\pi}\sum_{j=1}^{n}\left(\exp\left\{-\frac{1}{4}X_j^2\right\}\right)\right]\\
&+\frac{1}{(e^{\lambda}-1)^2}\int_{-\infty}^{\infty}\left[\sum_{k=0}^{\infty}\left(\frac{[\lambda(\exp[i\mu r-\sigma^2r^2/2]+\exp[-i\mu r-\sigma^2r^2/2])]^k}{k!}\cdot e^{-r^2}\right)dr\right.\\
&-\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left(\int_{-\infty}^{\infty}\exp\left[(i\mu r-\sigma^2r^2/2)k-r^2\right]dr\right)\\
&\left.-\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left(\int_{-\infty}^{\infty}\exp\left[(-i\mu r-\sigma^2r^2/2)k-r^2\right]dr\right)+\int_{-\infty}^{\infty}\exp(-r^2)\,dr\right]
\end{aligned}
$$

$$
\begin{aligned}
=&\frac{\pi^{1/2}}{n^2}\sum_{j=1}^{n}\sum_{t=1}^{n}e^{-\frac{1}{4}(X_j-X_t)^2}\\
&-\frac{1}{n(e^{\lambda}-1)}\times\left[\sum_{j=1}^{n}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left(\frac{2\pi}{2+k\sigma^2}\right)^{1/2}\exp\left\{-\frac{(X_j-k\mu)^2}{4+2k\sigma^2}\right\}\right.\\
&\left.+\sum_{t=1}^{n}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left(\frac{2\pi}{2+k\sigma^2}\right)^{1/2}\exp\left\{-\frac{(X_t-k\mu)^2}{4+2k\sigma^2}\right\}-2\sqrt{\pi}\sum_{j=1}^{n}\left(\exp\left\{-\frac{1}{4}X_j^2\right\}\right)\right]\\
&+\frac{1}{(e^{\lambda}-1)^2}\left[\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\sum_{p=0}^{k}\binom{k}{p}\right.\\
&\int_{-\infty}^{\infty}[\exp(i\mu r-\sigma^2r^2/2)]^{k-p}\cdot[\exp(-i\mu r-\sigma^2r^2/2)]^p\cdot\exp(-r^2)\,dr\\
&-\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left(\frac{2\pi}{2+k\sigma^2}\right)^{1/2}\exp\left\{-\frac{k^2\mu^2}{4+2k\sigma^2}\right\}\\
&\left.-\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left(\frac{2\pi}{2+k\sigma^2}\right)^{1/2}\exp\left\{-\frac{k^2\mu^2}{4+2k\sigma^2}\right\}+\sqrt{\pi}\right]
\end{aligned}
$$

$$
\begin{aligned}
=&\frac{\pi^{1/2}}{n^2}\sum_{j=1}^{n}\sum_{t=1}^{n} e^{-\frac{1}{4}(X_j-X_t)^2} - \frac{2}{n(e^{\lambda}-1)}\left[\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left(\frac{2\pi}{2+k\sigma^2}\right)^{1/2}\right.\\
&\left.\times\sum_{j=1}^{n}\exp\left\{-\frac{(X_j-k\mu)^2}{4+2k\sigma^2}\right\} - \pi^{1/2}\sum_{j=1}^{n}\left(\exp\left\{-\frac{1}{4}X_j^2\right\}\right)\right]\\
&-\left(\frac{2}{(e^{\lambda}-1)^2}\right)\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left(\frac{2\pi}{2+k\sigma^2}\right)^{1/2}\exp\left\{-\frac{k^2\mu^2}{4+2k\sigma^2}\right\} + \frac{\pi^{1/2}}{(e^{\lambda}-1)^2}\\
&+\frac{1}{(e^{\lambda}-1)^2}\left[\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\sum_{p=0}^{k}\binom{k}{p}\right.\\
&\left.\int_{-\infty}^{\infty}\exp\left\{[i\mu r-\sigma^2r^2/2]\cdot(k-p)+[-i\mu r-\sigma^2r^2/2]\cdot p - r^2\right\}\,dr\right]
\end{aligned}
$$

$$
\begin{aligned}
=&\frac{\pi^{1/2}}{n^2}\sum_{j=1}^{n}\sum_{t=1}^{n} e^{-\frac{1}{4}(X_j-X_t)^2} - \left(\frac{2}{n(e^{\lambda}-1)}\right)\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left(\frac{2\pi}{2+k\sigma^2}\right)^{1/2}\sum_{j=1}^{n}\left(\exp\left\{\frac{-(X_j-k\mu)^2}{4+2k\sigma^2}\right\}\right)\\
&+\left(\frac{2\pi^{1/2}}{n(e^{\lambda}-1)}\right)\sum_{j=1}^{n}\exp\left\{-\frac{1}{4}X_j^2\right\} - \left(\frac{2}{(e^{\lambda}-1)}\right)\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left(\frac{2\pi}{2+k\sigma^2}\right)^{1/2}\exp\left\{\frac{-k^2\mu^2}{4+2k\sigma^2}\right\}\\
&+\left(\frac{\pi}{(e^{\lambda}-1)^4}\right)^{1/2} + \left(\frac{1}{(e^{\lambda}-1)^2}\right)\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\sum_{p=0}^{k}\binom{k}{p}\left(\frac{2\pi}{2+k\sigma^2}\right)^{1/2}\exp\left\{\frac{-(k-2p)^2\mu^2}{4+2k\sigma^2}\right\}.
\end{aligned}
$$