

# UMA NOVA GENERALIZAÇÃO DA DISTRIBUIÇÃO LOMAX COM APLICAÇÃO EM ANÁLISE DE SOBREVIVÊNCIA

**Jailson Araujo Rodrigues**

Instituto Federal da Bahia - IFBA  
[jailsondearaujo@yahoo.com.br](mailto:jailsondearaujo@yahoo.com.br)

**Angela Lima da Silva**

Universidade Federal da Bahia - UFBA  
[angelalimadasilva@yahoo.com.br](mailto:angelalimadasilva@yahoo.com.br)

**Giovana Oliveira Silva**

Universidade Federal da Bahia-UFBA  
[giovanaa@ufba.br](mailto:giovanaa@ufba.br)

## RESUMO

Neste trabalho, é introduzida uma nova distribuição denominada Kumaraswamy exponenciada-Lomax que inclui os importantes submodelos: Lomax, Lomax exponenciada, Lomax generalizada exponenciada e Kumaraswamy Lomax. Algumas das principais propriedades estruturais dessa distribuição são apresentadas. A estimação dos parâmetros do modelo é discutida utilizando os métodos dos momentos e da máxima verossimilhança. O potencial da distribuição Kumaraswamy exponenciada-Lomax é demonstrado na modelagem de um conjunto de dados de sobrevivência.

## ABSTRACT

In this work, is introduced a new distribution called exponentiated Kumaraswamy Lomax that includes important submodels: Lomax, exponentiated Lomax, exponentiated generalized Lomax, Kumaraswamy Lomax. Some of the main structural properties of this distribution are presented. The estimation of the model parameters considering the methods of moments and maximum likelihood is also discussed. The potential of exponentiated Kumaraswamy Lomax distribution is shown in modeling a set of survival data.

**Palavras-chave:** Critério de Informação de Akaike, distribuição Lomax distribuição Kumaraswamy, momentos, método da máxima verossimilhança.

## 1 INTRODUÇÃO

A distribuição Lomax ou Pareto tipo II como também é conhecida foi originalmente proposta por Lomax [1] para analisar dados de insucesso empresarial. Logo após, diversos autores passaram a estudar e aplicar esse modelo em diversos contextos. Hassan e Al-Ghamdi [2] utilizaram a distribuição Lomax para determinar o tempo ótimo de mudança de nível para os planos de estresse simples em um modelo de exposição cumulativa. Holland et al. [3] utilizaram a distribuição Lomax para modelar tamanhos de arquivos em computadores utilizados como servidores. O modelo Lomax foi empregado por Vidondo et al. [4] para analisar dados de ecologia aquática.

Na última década, foram propostas diversas generalizações da distribuição Lomax. Entre os trabalhos mais recentes, pode-se destacar: Abdul-Moniem e Abdel-Hameed [5] com a distribuição Lomax exponenciada, Batal e Kareem [6] com a distribuição Kumaraswamy Lomax exponenciada, Coderiro et al. [7] com a distribuição Lomax generalizada exponenciada e Lemonte e Cordeiro [8] com a distribuição Lomax estendida.

Neste trabalho, é introduzida e estudada uma distribuição denominada Kumaraswamy exponenciada-Lomax. Essa nova distribuição possui como submodelos as distribuições Lomax, Lomax exponenciada e Lomax generalizada exponenciada. São apresentadas algumas das principais propriedades estruturais do modelo proposto. Entre essas propriedades inclui-se os momentos, os desvios médios e a medida da entropia de Rényi. A estimação dos parâmetros é discutida utilizando os métodos dos momentos e da máxima verossimilhança. O ajuste do novo modelo a um conjunto de dados de sobrevivência demonstra sua viabilidade.

O texto é organizado da seguinte forma: na Seção 2 a distribuição Kumaraswamy exponenciada-Lomax é definida e alguns submodelos são enumerados. Diversas propriedades estruturais do modelo são estudadas na Seção 3. A estimação dos parâmetros é discutida na Seção 4. Finalmente, na Seção 5 é reportada uma aplicação da distribuição Kumaraswamy exponenciada-Lomax.

## 2 O MODELO

A função de distribuição acumulada (FDA) da distribuição Lomax com parâmetro de forma  $a > 0$  e parâmetro de forma  $b > 0$  é definida por:

$$G(x; a, b) = 1 - \left(1 + \frac{x}{b}\right)^{-a} \quad (1)$$

sendo que  $x > 0$ . A função densidade de probabilidade (FDP) correspondente é dada por:

$$g(x; a, b) = \frac{ab^a}{(b+x)^{a+1}} = \frac{a}{b} \left(1 + \frac{x}{b}\right)^{-(a+1)} \quad (2)$$

sendo que  $x > 0$ ,  $a > 0$  e  $b > 0$ .

A distribuição Kumaraswamy tem sido extensivamente utilizada como uma alternativa para a distribuição beta pois, ambas possuem importantes características em comum. No entanto, a distribuição Kumaraswamy possui a vantagem de sua FDP (4) não depender da função beta. Esse fato, torna o tratamento matemático das propriedades dessa distribuição mais simples.

A FDA da distribuição Kumaraswamy [9] com parâmetros de forma  $\alpha > 0$  e  $\beta > 0$  é dada por

$$F(x; \alpha, \beta) = 1 - (1 - x^\alpha)^\beta \quad (3)$$

sendo que  $0 < x < 1$ . A FDP correspondente a (3) é dada por:

$$f(x; \alpha, \beta) = \alpha\beta x^{\alpha-1} (1 - x^\alpha)^{\beta-1} \quad (4)$$

Recentemente, Lemonte et al. [10] propôs uma generalização de (3) denominada distribuição Kumaraswamy exponenciada. A FDA e a FDP da distribuição Kumaraswamy exponenciada são dadas, respectivamente, por:

$$F(x; \alpha, \beta, \theta) = \left[1 - (1 - x^\alpha)^\beta\right]^\theta \quad (5)$$

e

$$f(x; \alpha, \beta, \theta) = \alpha\beta\theta x^{\alpha-1} (1 - x^\alpha)^{\beta-1} \left[1 - (1 - x^\alpha)^\beta\right]^{\theta-1} \quad (6)$$

sendo que  $\theta > 0$  é um parâmetro de forma.

Se  $G(x)$  é uma FDA de uma variável aleatória  $X$ . Então, a FDA de uma classe de funções de distribuições generalizadas é dada por:

$$F(x; \alpha, \beta, \theta) = \left[1 - (1 - G(x)^\alpha)^\beta\right]^\theta \tag{7}$$

A correspondente FDP para (7) é

$$f(x; \alpha, \beta, \theta) = \alpha\beta\theta g(x)G(x)^{\alpha-1} (1 - G(x)^\alpha)^{\beta-1} \left[1 - (1 - G(x)^\alpha)^\beta\right]^{\theta-1} \tag{8}$$

Huang e Oluyede [11] aplicaram a FDA da distribuição Dagum em (8) para definir a distribuição Kumaraswamy exponenciada-Dagum. Pode-se observar que a generalização (8) pode ser empregada para definir outras distribuições baseadas no modelo Kumaraswamy exponenciado.

Substituindo (1) em (7), obtém-se uma nova distribuição chamada de Kumaraswamy exponenciada-Lomax (KEL) com FDA

$$F(x; a, b, \alpha, \beta, \theta) = \left\{1 - \left[1 - \left(1 - \left(1 + \frac{x}{b}\right)^{-a}\right)^\alpha\right]^\beta\right\}^\theta \tag{9}$$

sendo que  $x > 0, a > 0, b > 0, \alpha > 0, \beta > 0$  e  $\theta > 0$ . A FDP associada ao modelo KEL é dada por:

$$f(x; a, b, \alpha, \beta, \theta) = \frac{a\alpha\beta\theta}{b} \left(1 + \frac{x}{b}\right)^{-(a+1)} \left[1 - \left(1 + \frac{x}{b}\right)^{-a}\right]^{\alpha-1} \left\{1 - \left[1 - \left(1 + \frac{x}{b}\right)^{-a}\right]^\alpha\right\}^{\beta-1} \times \left\{1 - \left[1 - \left(1 - \left(1 + \frac{x}{b}\right)^{-a}\right)^\alpha\right]^\beta\right\}^{\theta-1} \tag{10}$$

Na Figura 1 são exibidas algumas formas possíveis da FDP (10) para diferentes valores dos parâmetros  $a, b, \alpha, \beta$  e  $\theta$ .

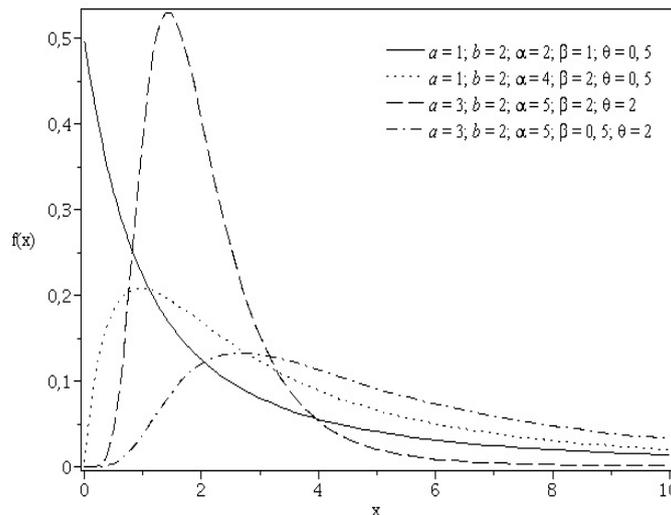


FIGURA 1: Gráficos da FDP da distribuição KEL para diferentes valores paramétricos.

### 2.1 SUBMODELOS

Nesta subseção, são enumerados alguns importantes submodelos da distribuição KEL.

1. Quando  $\alpha = \beta = \theta = 1$ , a distribuição KEL corresponde ao modelo Lomax com FDP

$$f(x; a, b) = \frac{a}{b} \left(1 + \frac{x}{b}\right)^{-(a+1)} \quad (11)$$

2. Se  $\alpha = \beta = 1$ , a distribuição KEL coincide com a distribuição Lomax exponenciada com FDP

$$f(x; a, b, \theta) = \frac{a\theta}{b} \left(1 + \frac{x}{b}\right)^{-(a+1)} \left[1 - \left(1 + \frac{x}{b}\right)^{-a}\right]^{\theta-1} \quad (12)$$

3. Caso  $\beta = \theta = 1$ , obtém-se a distribuição Lomax exponenciada como submodelo da distribuição KEL. A FDP correspondente é dada por:

$$f(x; a, b, \alpha) = \frac{a\alpha}{b} \left(1 + \frac{x}{b}\right)^{-(a+1)} \left[1 - \left(1 + \frac{x}{b}\right)^{-a}\right]^{\alpha-1} \quad (13)$$

4. Caso  $\beta = 1$ , a distribuição KEL reduz-se a outra distribuição Lomax exponenciada com parâmetros  $a$ ,  $b$  e  $\alpha\theta$ . A FDP correspondente é dada por:

$$f(x; a, b, \alpha\theta) = \frac{a\alpha\theta}{b} \left(1 + \frac{x}{b}\right)^{-(a+1)} \left[1 - \left(1 + \frac{x}{b}\right)^{-a}\right]^{\alpha\theta-1} \quad (14)$$

5. Se  $\alpha = 1$ , o modelo KEL representa a distribuição Lomax generalizada exponenciada com FDP

$$f(x; a, b, \beta, \theta) = \frac{a\beta\theta}{b} \left(1 + \frac{x}{b}\right)^{-(a\beta+1)} \left[1 - \left(1 + \frac{x}{b}\right)^{-a\beta}\right]^{\theta-1} \quad (15)$$

6. Se  $\theta = 1$ , a Equação 10 representa o modelo Kumaraswamy Lomax com FDP

$$f(x; a, b, \alpha, \beta) = \frac{a\alpha\beta}{b} \left(1 + \frac{x}{b}\right)^{-(a+1)} \left[1 - \left(1 + \frac{x}{b}\right)^{-a}\right]^{\alpha-1} \left\{1 - \left[1 - \left(1 + \frac{x}{b}\right)^{-a}\right]^\alpha\right\}^{\beta-1} \quad (16)$$

### 3 PROPRIEDADES DO MODELO KEL

#### 3.1 EXPANSÕES EM SÉRIE PARA A FDA E A FDP

Para todo número real não inteiro  $\nu > 0$ , tem-se a seguinte expansão em série de potências

$$(1 - \omega)^{\nu-1} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\nu) \omega^j}{\Gamma(\nu - j) j!} \quad (17)$$

sendo que  $|\omega| < 1$ . Utilizando a representação (17) na Equação 9, tem-se

$$F(x; a, b, \alpha, \beta, \theta) = \beta\theta \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(-1)^{j+k} j \Gamma(\theta) \Gamma(\beta j)}{\Gamma(\theta + 1 - j) \Gamma(\beta j + 1 - k) (j-1)! k!} \left[1 - \left(1 + \frac{x}{b}\right)^{-a}\right]^{\alpha k} \quad (18)$$

com  $\alpha$  e  $\beta$  reais não inteiros. Se  $\theta$  é um número inteiro, então, o índice  $j$  na soma (18) para em  $\theta$ . Se  $\beta$  é um inteiro, então, o índice  $k$  para em  $\beta j$ . De forma análoga, aplicando a representação (17) na Equação 10, pode-se escrever:

$$f(x; a, b, \alpha, \beta, \theta) = \frac{a\alpha\beta\theta}{b} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{j+k+l} \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)]}{\Gamma(\theta-j) \Gamma[\beta(j+1)-k] \Gamma[\alpha(k+1)-l] j!k!l!} \times \left(1 + \frac{x}{b}\right)^{-a(l+1)-1} \tag{19}$$

Se  $\theta$  é um número inteiro, então, o índice  $j$  na soma (19) para em  $\theta - 1$ . Se  $\beta$  é um inteiro, então, o índice  $k$  para em  $\beta(j+1) - 1$ . Caso  $\alpha$  seja inteiro, o índice  $l$  estaciona em  $\alpha(k+1) - 1$ .

### 3.2 FUNÇÃO DE RISCO

Para uma variável aleatória contínua  $X$  com FDP  $f(x)$  e CDF  $F(x)$ , a função de risco é dada por:

$$h(x) = \lim_{\Delta x \rightarrow 0} \frac{P(X < x + \Delta x | X > x)}{\Delta x} = \frac{f(x)}{1 - F(x)} \tag{20}$$

A função de risco desempenha um papel importante na análise de dados de sobrevivência. Para a distribuição KEL, a função de risco é dada por:

$$h(x; a, b, \alpha, \beta, \theta) = \frac{\frac{a\alpha\beta\theta}{b} \left(1 + \frac{x}{b}\right)^{-(a+1)} \left[1 - \left(1 + \frac{x}{b}\right)^{-a}\right]^{\alpha-1} \left\{1 - \left[1 - \left(1 + \frac{x}{b}\right)^{-a}\right]^{\alpha}\right\}^{\beta-1}}{1 - \left\{1 - \left[1 - \left(1 + \frac{x}{b}\right)^{-a}\right]^{\alpha}\right\}^{\beta}}^{\theta} \times \left\{1 - \left[1 - \left(1 + \frac{x}{b}\right)^{-a}\right]^{\alpha}\right\}^{\beta} \right\}^{\theta-1} \tag{21}$$

Na Figura 2 são exibidas algumas formas possíveis da função de risco (21) para diferentes valores dos parâmetros  $a, b, \alpha, \beta$  e  $\theta$ .

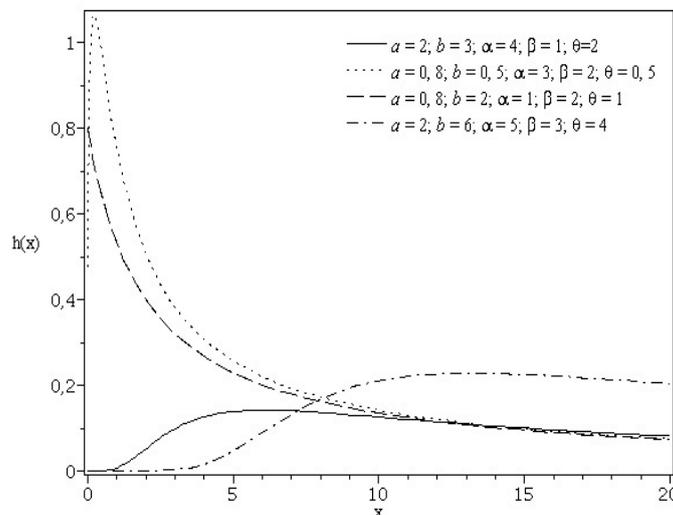


FIGURA 2: Gráficos da função de risco da distribuição KEL para diferentes valores paramétricos.

### 3.3 FUNÇÃO QUANTIL

A função quantil  $Q(p)$ ,  $0 < p < 1$ , para a distribuição KEL é dada por:

$$Q(p) = \frac{b}{\left\{1 - \left[1 - (1 - p^{1/\theta})^{1/\beta}\right]^{1/\alpha}\right\}^{1/a}} - b \tag{22}$$

A mediana da distribuição é obtida ao substituir  $p = 0,5$  em (22). Uma amostra aleatória também pode ser gerada a partir de (22), basta utilizar  $p$  como sendo as observações de uma variável aleatória uniforme no intervalo  $(0, 1)$ .

### 3.4 MOMENTOS

Diversas características importantes de uma distribuição podem ser estudadas através dos seus momentos. Dessa forma, é comum determinar os momentos quando uma distribuição é proposta.

Utilizando a Equação 19, o momento de ordem  $r$  da distribuição KEL pode ser escrito como:

$$\begin{aligned} E(X^r) &= \int_0^\infty x^r f(x; a, b, \alpha, \beta, \theta) dx \\ &= \int_0^\infty \frac{a\alpha\beta\theta}{b} \sum_{j=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \frac{(-1)^{j+k+l} \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)]}{\Gamma(\theta-j) \Gamma[\beta(j+1)-k] \Gamma[\alpha(k+1)-l] j!k!l!} \\ &\quad \times x^r \left(1 + \frac{x}{b}\right)^{-a(l+1)-1} dx \\ &= \frac{a\alpha\beta\theta}{b} \sum_{j=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \frac{(-1)^{j+k+l} \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)]}{\Gamma(\theta-j) \Gamma[\beta(j+1)-k] \Gamma[\alpha(k+1)-l] j!k!l!} \\ &\quad \times \int_0^\infty x^r \left(1 + \frac{x}{b}\right)^{-a(l+1)-1} dx \\ &= a\alpha\beta\theta \sum_{j=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \frac{(-1)^{j+k+l} b^{a(l+1)} \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)]}{\Gamma(\theta-j) \Gamma[\beta(j+1)-k] \Gamma[\alpha(k+1)-l] j!k!l!} \\ &\quad \times \int_0^\infty x^r (b+x)^{-a(l+1)-1} dx \end{aligned} \tag{23}$$

Segundo Prudnikov et al. [12], se  $|\arg z| < \pi$  e  $0 < \operatorname{Re} \alpha < \operatorname{Re} \rho$ , então,

$$\int_0^\infty x^{\alpha-1} (x+z)^{-\rho} dx = z^{\alpha-\rho} B(\alpha, \rho-\alpha) \tag{24}$$

sendo que  $B(a, b)$  representa a função beta,

$$B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt, \tag{25}$$

Então, a aplicação da integral (24) em (23) resulta que o momento de ordem  $r$  da distribuição KEL pode ser escrito como

$$E(X^r) = a\alpha\beta\theta \sum_{j=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \frac{(-1)^{j+k+l} b^r \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)] B(r+1, a(l+1)-r)}{\Gamma(\theta-j) \Gamma[\beta(j+1)-k] \Gamma[\alpha(k+1)-l] j!k!l!}$$

sempre que  $a(l+1) > r$ . Em particular, a média da distribuição KEL é dada por

$$E(X) = a\alpha\beta\theta \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{j+k+l} b \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)] B(2, a(l+1) - 1)}{\Gamma(\theta - j) \Gamma[\beta(j+1) - k] \Gamma[\alpha(k+1) - l] j! k! l!}$$

A partir dos quatro primeiros momentos da distribuição KEL, pode-se obter a variância, os coeficientes de assimetria e curtose dessa distribuição utilizando as respectivas expressões:

$$Var(X) = E(X^2) - E^2(X) \tag{26}$$

$$Ass(X) = \frac{E(X^3) - 3E(X)E^2(X) + 2E^3(X)}{Var^{3/2}(X)} \tag{27}$$

$$Cur(X) = \frac{E(X^4) - 4E(X)E^3(X) + 6E(X^2)E^2(X) - 3E^4(X)}{Var^2(X)} \tag{28}$$

A Tabela 1 contém a média da distribuição KEL para diferentes valores dos parâmetros  $a, b, \alpha, \beta$  e  $\theta$ .

Tabela 1. Média da distribuição KEL para diferentes valores paramétricos.

$a$	$b$	$\alpha$	$\beta$	$\theta$	Média
0.8	2.5	3.0	4.0	0.8	5,169
1.5	1.5	3.0	4.0	1.8	1,678
2.0	1.0	1.0	4.0	1.0	0,142
2.0	1.0	3.0	4.0	1.0	0,565
2.0	2.0	1.0	4.0	1.0	0,286
2.0	2.0	1.0	4.0	5.0	0,478
3.0	2.0	3.0	4.0	5.0	0,408
3.0	2.2	2.5	1.5	1.8	5,459
3.5	2.0	3.0	4.0	5.0	0,321
3.5	2.5	3.0	4.0	2.0	0,932
3.5	2.5	3.0	4.0	5.0	0,402
3.5	4.5	3.0	4.0	6.0	0,279
3.5	4.5	3.0	4.0	1.5	1,336
4.2	1.6	3.0	2.0	2.4	0.781
5.6	2.0	1.0	2.0	3.8	0,340
6.5	4.5	3.0	4.0	2.5	0,815

### 3.5 DESVIOS MÉDIOS

Seja  $X$  uma variável aleatória com distribuição KEL com média  $\mu = E(X)$  e mediana  $m$ . O desvio médio em relação a média é dado por:

$$\begin{aligned} \delta_1(X) &= E(|X - \mu|) \\ &= \int_0^\infty |x - \mu| f(x; a, b, \alpha, \beta, \theta) dx \\ &= 2\mu F(\mu; a, b, \alpha, \beta, \theta) - 2\mu + 2 \int_\mu^\infty x f(x; a, b, \alpha, \beta, \theta) dx \\ &= 2\mu \left\{ 1 - \left[ 1 - \left( 1 - \left( 1 + \frac{\mu}{b} \right)^{-a} \right)^\alpha \right]^\beta \right\}^\theta - 2\mu \\ &\quad + \frac{2a\alpha\beta\theta}{b} \sum_{j=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \frac{(-1)^{j+k+l} \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)]}{\Gamma(\theta-j) \Gamma[\beta(j+1)-k] \Gamma[\alpha(k+1)-l] j! k! l!} \\ &\quad \times \int_\mu^\infty x \left( 1 + \frac{x}{b} \right)^{-a(l+1)-1} dx \end{aligned} \tag{29}$$

Segundo Gradshteyn e Ryzhik [13], se  $Re \alpha < Re \nu$ , então,

$$\int_\mu^\infty x^{\alpha-1} (1 + \beta x)^{-\nu} dx = \frac{\mu^{\alpha-\nu}}{\beta^\nu (\nu - \alpha)} {}_2F_1 \left( \nu, \nu - \alpha; \nu - \alpha + 1; -\frac{1}{\beta\mu} \right) \tag{30}$$

sendo que  ${}_2F_1(a, b; c; x)$  representa a função hipergeométrica de Gauss,

$${}_2F_1(a, b; c; x) = \sum_{k=0}^\infty \frac{(a)_k (b)_k x^k}{(c)_k k!} \tag{31}$$

em que  $(d)_k = d(d+1) \dots (d+k-1)$  denota o fatorial ascendente.

Aplicando a integral (30) em (29), obtém-se o desvio médio em relação a média da distribuição KEL,

$$\begin{aligned} \delta_1(X) &= 2\mu \left\{ 1 - \left[ 1 - \left( 1 - \left( 1 + \frac{x}{b} \right)^{-a} \right)^\alpha \right]^\beta \right\}^\theta - 2\mu \\ &\quad + 2a\alpha\beta\theta \sum_{j=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \frac{(-1)^{j+k+l} b^{a(l+1)} \mu^{1-a(l+1)} \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)]}{[a(l+1)-1] \Gamma(\theta-j) \Gamma[\beta(j+1)-k] \Gamma[\alpha(k+1)-l] j! k! l!} \\ &\quad \times {}_2F_1 \left( a(l+1) + 1, a(l+1) - 1; a(l+1); -\frac{b}{\mu} \right) \end{aligned} \tag{32}$$

Por outro lado, o desvio médio em relação a mediana da distribuição KEL é dado por:

$$\begin{aligned} \delta_1(X) &= E(|X - m|) \\ &= \int_0^\infty |x - m| f(x; a, b, \alpha, \beta, \theta) dx \\ &= -\mu + 2 \int_m^\infty x f(x; a, b, \alpha, \beta, \theta) dx \\ &= -\mu + 2a\alpha\beta\theta \sum_{j=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \frac{(-1)^{j+k+l} b^{a(l+1)} m^{1-a(l+1)} \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)]}{[a(l+1)-1] \Gamma(\theta-j) \Gamma[\beta(j+1)-k] \Gamma[\alpha(k+1)-l] j! k! l!} \\ &\quad \times {}_2F_1 \left( a(l+1) + 1, a(l+1) - 1; a(l+1); -\frac{b}{m} \right) \end{aligned} \tag{33}$$

### 3.6 CURVAS DE BONFERRONI E LORENZ

As curvas de Bonferroni e Lorenz são muito utilizadas em diversas áreas do conhecimento incluindo economia, demografia, seguros e medicina. As curvas de Bonferroni e Lorenz são dadas, respectivamente, por:

$$B(p) = \frac{1}{p\mu} \int_0^q x f(x) dx \tag{34}$$

e

$$L(p) = \frac{1}{\mu} \int_0^q x f(x) dx \tag{35}$$

sendo que  $\mu = E(X)$  e  $q = F^{-1}(p)$ . Em particular, para a distribuição KEL, a curva de Bonferroni é dada por:

$$\begin{aligned} B(p) &= \frac{1}{p\mu} \int_0^q f(x; a, b, \alpha, \beta, \theta) dx \\ &= \frac{a\alpha\beta\theta}{bp\mu} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{j+k+l} \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)]}{\Gamma(\theta-j) \Gamma[\beta(j+1)-k] \Gamma[\alpha(k+1)-l] j!k!l!} \\ &\quad \times \int_0^q x \left(1 + \frac{x}{b}\right)^{-a(l+1)-1} dx \end{aligned} \tag{36}$$

Segundo Gradshteyn e Ryzhik [13], se  $Re \alpha > 0$  e  $|\arg(1 + \beta q)| < \pi$ , então,

$$\int_0^q x^{\alpha-1} (1 + \beta x)^{-\nu} dx = \frac{q^\alpha}{\alpha} {}_2F_1(\nu, \alpha; \alpha + 1; -\beta q) \tag{37}$$

Aplicando a integral (37) em (36), obtém-se a curva de Bonferroni para a distribuição KEL,

$$\begin{aligned} B(p) &= \frac{a\alpha\beta\theta q^2}{2bp\mu} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{j+k+l} \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)]}{\Gamma(\theta-j) \Gamma[\beta(j+1)-k] \Gamma[\alpha(k+1)-l] j!k!l!} \\ &\quad \times {}_2F_1\left(a(l+1) + 1, 2; 3; -\frac{q}{b}\right) \end{aligned} \tag{38}$$

Por outro lado, a curva de Lorenz para a distribuição KEL é,

$$\begin{aligned} L(p) &= \frac{a\alpha\beta\theta q^2}{2b\mu} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{j+k+l} \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)]}{\Gamma(\theta-j) \Gamma[\beta(j+1)-k] \Gamma[\alpha(k+1)-l] j!k!l!} \\ &\quad \times {}_2F_1\left(a(l+1) + 1, 2; 3; -\frac{q}{b}\right) \end{aligned} \tag{39}$$

### 3.7 ENTROPIA

Uma entropia de uma variável aleatória  $X$  é uma medida da variação da incerteza. É um importante conceito em diversos domínios científicos, especialmente em teoria da comunicação, física e probabilidade. Uma medida de entropia bastante popular é a entropia de Rényi. Se  $X$  possui FDP  $f(\cdot)$ , então, a entropia de Rényi entropy é definida como

$$H_R(\nu) = \frac{1}{1-\nu} \log \left[ \int f^\nu(x) dx \right] \tag{40}$$

sendo que  $\nu > 0$  e  $\nu \neq 1$ . Utilizando a definição (40), pode-se concluir que a entropia de Rényi para a distribuição KEL é

$$\begin{aligned}
 H_R(\nu) &= \frac{\nu}{1-\nu} (\log a + \log \alpha + \log \beta + \log \theta - \log b) \\
 &+ \frac{1}{1-\nu} \log \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{j+k+l} b \Gamma[\nu(\theta-1)+1] \Gamma[\beta j + \nu(\beta-1)+1]}{\Gamma[\nu(\theta-1)+1-j] \Gamma[\beta j + \nu(\beta-1)+1-k]} \\
 &\times \frac{\Gamma[\alpha k + \nu(\alpha-1)+1] \Gamma[al + \nu(a+1)-1]}{\Gamma[\alpha k + \nu(\alpha-1)+1-l] \Gamma[al + \nu(a+1)] j! k! l!}
 \end{aligned}
 \tag{41}$$

### 3.8 ESTATÍSTICA DE ORDEM

Estatísticas de ordem são muito importantes ao se trabalhar com estatística não paramétrica e inferência em testes de hipóteses. Dessa forma, é importante a discussão das propriedades da estatística de ordem para a distribuição KEL. Suponha que  $X_1, X_2, \dots, X_n$  é uma amostra aleatória da distribuição KEL. Se  $X_{1:n} < X_{2:n} < \dots < X_{n:n}$  denota a estatística de ordem correspondente. Segundo Arnold et al. [14], a FDP e a FDA da  $r$ -ésima estatística de ordem, denotada por  $Y = X_{r:n}$ , são dadas, respectivamente, por:

$$\begin{aligned}
 f_Y(y) &= \frac{n!}{(r-1)!(n-r)!} F^{r-1}(y) [1-F(y)]^{n-r} f(y) \\
 &= \frac{n!}{(r-1)!(n-r)!} \sum_{l=0}^{n-r} \binom{n-r}{l} (-1)^l F^{l+r-1}(y) f(y)
 \end{aligned}
 \tag{42}$$

e

$$\begin{aligned}
 F_Y(y) &= \sum_{j=r}^n \binom{n}{j} F^j(y) [1-F(y)]^{n-j} \\
 &= \sum_{j=r}^n \sum_{l=0}^{n-j} \binom{n}{j} \binom{n-j}{l} (-1)^l F^{j+l}(y)
 \end{aligned}
 \tag{43}$$

sendo que  $f(\cdot)$  e  $F(\cdot)$  representam a FDP e FDA da distribuição KEL. Das Equações 9–10, tem-se:

$$\begin{aligned}
 f_Y(y) &= \frac{a\alpha\beta\theta n!}{b(r-1)!(n-r)!} \left(1 + \frac{y}{b}\right)^{-(a+1)} \left[1 - \left(1 + \frac{y}{b}\right)^{-a}\right]^{\alpha-1} \left\{1 - \left[1 - \left(1 + \frac{y}{b}\right)^{-a}\right]^{\alpha}\right\}^{\beta-1} \\
 &\times \sum_{l=0}^{n-r} \binom{n-r}{l} (-1)^l \left\{1 - \left[1 - \left(1 - \left(1 + \frac{y}{b}\right)^{-a}\right)^{\alpha}\right]^{\beta}\right\}^{\theta(l+r)-1}
 \end{aligned}
 \tag{44}$$

e

$$F_Y(y) = \sum_{j=r}^n \sum_{l=0}^{n-j} \binom{n}{j} \binom{n-j}{l} (-1)^l \left\{1 - \left[1 - \left(1 - \left(1 + \frac{y}{b}\right)^{-a}\right)^{\alpha}\right]^{\beta}\right\}^{\theta(j+l)}
 \tag{45}$$

## 4 ESTIMAÇÃO DOS PARÂMETROS DO MODELO

Nesta seção, é analisada a estimação dos cinco parâmetros da distribuição proposta utilizando os métodos dos momentos e da máxima verossimilhança. Suponha que  $x_1, \dots, x_n$  é uma amostra aleatória de tamanho  $n$  da distribuição KEL. Considerando o método dos

momentos, igualando-se os momentos teóricos  $E(X^r)$  com os respectivos momentos amostrais,

$$M_r = \frac{1}{n} \sum_{l=1}^n x_l^r, \quad r = 1, \dots, 5. \tag{46}$$

obtém-se o seguinte sistema de equações:

$$M_r = a\alpha\beta\theta \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{j+k+l} b^r \Gamma(\theta) \Gamma[\beta(j+1)] \Gamma[\alpha(k+1)] B(r+1, a(l+1) - r)}{\Gamma(\theta - j) \Gamma[\beta(j+1) - k] \Gamma[\alpha(k+1) - l] j! k! l!} \tag{47}$$

que pode ser resolvido numericamente e cujas raízes são as esmimativas dos parâmetros  $a, b, \alpha, \beta$  e  $\theta$ .

Considere agora a estimação dos parâmetros pelo método da máxima verossimilhança. A log-verossimilhança para uma amostra aleatória  $x_1, \dots, x_n$  da distribuição KEL é dada por:

$$\begin{aligned} \log L(a, b, \alpha, \beta, \theta) &= n \log a + n \log \alpha + n \log \beta + n \log \theta - n \log b - (a+1) \sum_{i=1}^n \log \left(1 + \frac{x_i}{b}\right) \\ &+ (\alpha - 1) \sum_{i=1}^n \log \left[1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right] + (\beta - 1) \sum_{i=1}^n \log \left\{1 - \left[1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right]^{\alpha}\right\} \\ &+ (\theta - 1) \sum_{i=1}^n \log \left\{1 - \left[1 - \left(1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right)^{\alpha}\right]^{\beta}\right\} \end{aligned} \tag{48}$$

Derivando a log-verossimilhança em relação a cada um dos parâmetros  $a, b, \alpha, \beta, \theta$  e igualando os resultados a zero, tem-se:

$$\begin{aligned} \frac{\partial \log L}{\partial a} &= \frac{n}{a} - \sum_{i=1}^n \log \left(1 + \frac{x_i}{b}\right) + \alpha(1 - \beta) \sum_{i=1}^n \frac{\left[1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right]^{\alpha-1} \left(1 + \frac{x_i}{b}\right)^{-a} \log \left(1 + \frac{x_i}{b}\right)}{1 - \left[1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right]^{\alpha}} \\ &+ \alpha\beta(\theta - 1) \sum_{i=1}^n \frac{\left\{1 - \left[1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right]^{\alpha}\right\}^{\beta-1} \left[1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right]^{\alpha-1} \left(1 + \frac{x_i}{b}\right)^{-a} \log \left(1 + \frac{x_i}{b}\right)}{1 - \left\{1 - \left[1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right]^{\alpha}\right\}^{\beta}} \\ &+ (\alpha - 1) \sum_{i=1}^n \frac{\log \left(1 + \frac{x_i}{b}\right)}{\left(1 + \frac{x_i}{b}\right)^a - 1} \end{aligned} \tag{49}$$

$$\begin{aligned} \frac{\partial \log L}{\partial b} &= -\frac{n}{b} + \frac{(a+1)}{b} \sum_{i=1}^n \frac{x_i}{b + x_i} + \frac{a\alpha(\beta - 1)}{b^2} \sum_{i=1}^n \frac{x_i \left[1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right]^{\alpha-1} \left(1 + \frac{x_i}{b}\right)^{-(a+1)}}{1 - \left[1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right]^{\alpha}} \\ &+ \frac{a\alpha\beta(1 - \theta)}{b^2} \sum_{i=1}^n \frac{x_i \left(1 + \frac{x_i}{b}\right)^{-(a+1)} \left[1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right]^{\alpha-1} \left\{1 - \left[1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right]^{\alpha}\right\}^{\beta-1}}{1 - \left\{1 - \left[1 - \left(1 + \frac{x_i}{b}\right)^{-a}\right]^{\alpha}\right\}^{\beta}} \\ &+ \frac{a(1 - \alpha)}{b^2} \sum_{i=1}^n \frac{x_i \left(1 + \frac{x_i}{b}\right)^{-(a+1)}}{1 - \left(1 + \frac{x_i}{b}\right)^{-a}} \end{aligned} \tag{50}$$

$$\begin{aligned} \frac{\partial \log L}{\partial \alpha} &= \frac{n}{\alpha} + \sum_{i=1}^n \log \left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right] + (1 - \beta) \sum_{i=1}^n \frac{\left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right]^\alpha \log \left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right]}{1 - \left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right]^\alpha} \\ &+ \beta (\theta - 1) \sum_{i=1}^n \frac{\left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right]^\alpha \log \left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right] \left\{ 1 - \left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right]^\alpha \right\}^{\beta-1}}{1 - \left\{ 1 - \left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right]^\alpha \right\}^\beta} \\ \frac{\partial \log L}{\partial \beta} &= \frac{n}{\beta} + (1 - \theta) \sum_{i=1}^n \frac{\log \left\{ 1 - \left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right]^\alpha \right\} \left\{ 1 - \left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right]^\alpha \right\}^\beta}{1 - \left\{ 1 - \left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right]^\alpha \right\}^\beta} \\ &+ \sum_{i=1}^n \log \left\{ 1 - \left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right]^\alpha \right\} \end{aligned} \quad (51)$$

$$\frac{\partial \log L}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^n \log \left\{ 1 - \left[ 1 - \left( 1 + \frac{x_i}{b} \right)^{-a} \right]^\alpha \right\}^\beta \quad (52)$$

As estimativas de máximo verossimilhança  $\hat{a}$ ,  $\hat{b}$ ,  $\hat{\alpha}$ ,  $\hat{\beta}$  e  $\hat{\theta}$  dos parâmetros desconhecidos  $a$ ,  $b$ ,  $\alpha$ ,  $\beta$  e  $\theta$  são obtidas resolvendo simultaneamente as Equações 49–52.

## 5 APLICAÇÃO

Nesta seção, é utilizado um conjunto de dados de sobrevivência para comparar o ajuste da distribuição KEL com outras quatro distribuições: gama, log-normal, Lomax e Pareto. As FDPs das distribuições gama, log-normal e Pareto são dadas por:

1. Gama:

$$f(x; a, b) = \frac{b^a x^{a-1}}{\Gamma(a)} \exp(-bx) \quad (53)$$

2. Log-normal (LN):

$$f(x; a, b) = \frac{1}{xb\sqrt{2\pi}} \exp \left[ -\frac{(\log x - a)^2}{2b^2} \right] \quad (54)$$

3. Pareto:

$$f(x; a, b) = ab^a x^{-(a+1)} \quad (55)$$

A base de dados considerada corresponde ao tempo de remissão (em meses) do câncer de bexiga diagnosticado em cento e vinte e oito pacientes. Esses dados foram retirados de Lee e Wang [15].

As estimativas de máxima verossimilhança dos parâmetros dos cinco modelos ajustados são reportadas na Tabela 2. Na mesma tabela, são apresentados os valores do Critério de Informação de Akaike (AIC). Os resultados indicam que o modelo KEL apresenta um ajuste melhor se comparado com os demais modelos.

Na Figura 3 são apresentados os gráficos de probabilidade para os cinco modelos avaliados. Ao observar a figura, pode-se notar que a distribuição Pareto foi a com pior o ajuste. Os gráficos são inconclusivos em relação a distribuição que apresenta melhor ajuste. Dessa forma, os gráficos de probabilidade corroboram parcialmente com os resultados obtidos anteriormente pelos testes analíticos.

Tabela 2. Estimativas dos parâmetros e AIC dos modelos ajustados.

Modelos	$a$	$b$	$\alpha$	$\beta$	$\theta$	AIC
Gama	1,173	0,125	...	...	...	830,736
KEL	5,497	22,980	2,170	0,758	0,783	830,433
LN	1,753	1,152	...	...	...	834,189
Lomax	13,938	121,022	...	...	...	831,670
Pareto	0,080	0,234	...	...	...	1081,046

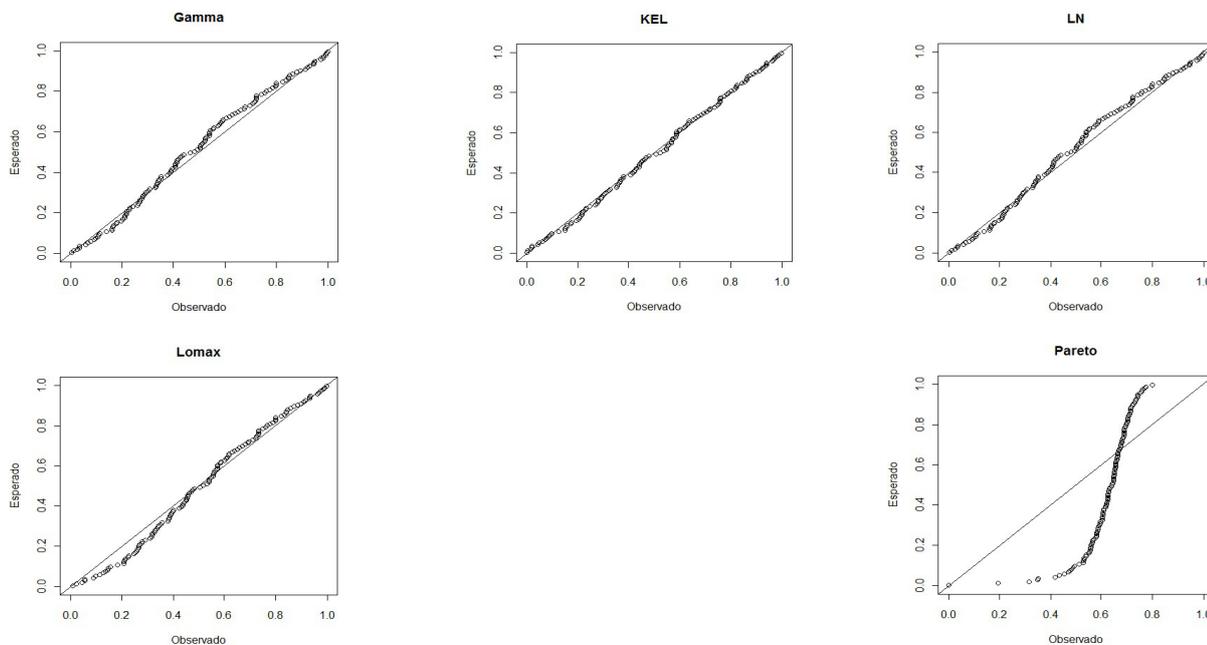


FIGURA 3: Gráficos de probabilidade das distribuições ajustadas.

Na Figura 4 pode-se comparar o histograma de frequências dos dados observados sobreposto pelas FDPs gama, KEL, LN, Lomax e Pareto que foram ajustadas. O que se pode notar é que as curvas de ajustes das distribuições gama, KEL, LN e Lomax acompanham as condições de moda e assimetria, o que não ocorre com o modelo Pareto. No entanto, não é possível determinar, visualmente, a que melhor se ajustou aos dados observados.

## 6 CONCLUSÕES

Neste trabalho, foi introduzida uma nova distribuição com cinco parâmetros. Essa distribuição, denominada de Kumaraswamy exponenciada-Lomax inclui como submodelos as distribuições Lomax, Lomax exponenciada, Lomax generalizada exponenciada e Kumaraswamy Lomax. Diversas propriedades estruturais do modelo foram deduzidas. Foram apresentados sistemas de equações não lineares para determinação das estimativas dos parâmetros do modelo. Essas equações fornecem estimativas pelos métodos dos momentos e da máxima verossimilhança. Uma aplicação da distribuição proposta a dados reais demonstrou a flexibilidade e o potencial dessa distribuição na análise de dados de sobrevivência.

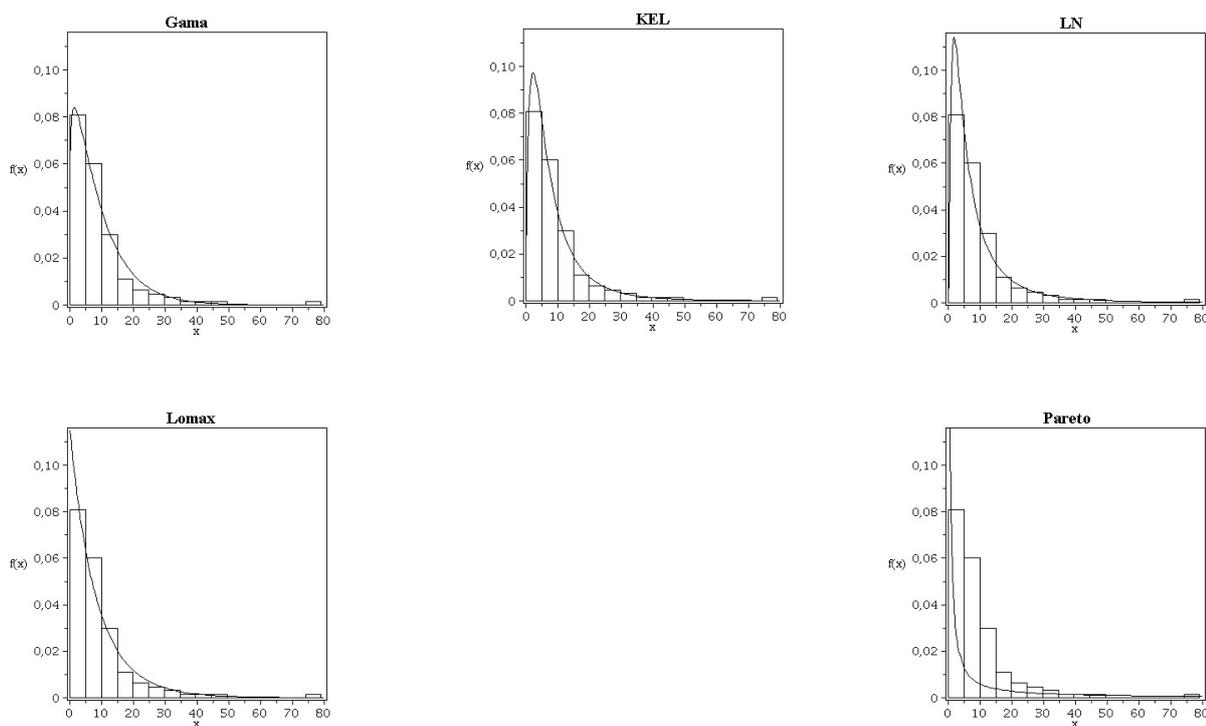


FIGURA 4: Histograma dos dados e FDPs ajustadas.

## REFERÊNCIAS

- [1] H. S. LOMAX, "Business failures: Another example of the analysis of failure data," *Journal of the American Statistical Association*, vol. 49(268), pp. 847–852, 1954.
- [2] A. S. HASSAN and A. S. AL-GAMDI, "Optimum step stress accelerated life testing for lomax distribution," *Journal of Applied Sciences Research*, vol. 5, pp. 2153–2164, 2009.
- [3] O. HOLLAND, A. GOLAU, and A. H. AGHVAMI, "Traffic characteristics of aggregated module downloads for mobile terminal reconfiguration," *IEEE Proceedings - Communications*, vol. 153(5), pp. 683–690, 2006.
- [4] B. VIDONDO, Y. PRAIRIE, J. M. BLANCO, and J. M. DUARTE, "Some aspects of the analysis of size spectra in aquatic ecology," *Limnology and Oceanography*, vol. 42, pp. 184–192, 1997.
- [5] I. B. ABDUL-MONIEM and H. F. ABDEL-MAMEED, "On exponentiated lomax distribution," *International Journal of Mathematical Archive*, vol. 3(5), pp. 1–7, 2012.
- [6] I. E. BATAL and A. KAREEM, "Statistical properties of kumaraswamy exponentiated lomax distribution," *Journal of Modern Mathematics and Statistics*, vol. 8(5), pp. 1–7, 2014.
- [7] G. M. CORDEIRO, E. M. M. ORTEGA, and D. C. C. CUNHA, "The exponentiated generalized class of distributions," *Journal of data science*, vol. 11, pp. 1–27, 2013.
- [8] A. J. LEMONTE and G. M. CORDEIRO, "An extended lomax distribution," *Statistics*, vol. 47(4), pp. 1–17, 2011.
- [9] P. KUMARASWAMY, "A generalized probability density function for double-bounded random processes," *Journal of Hydrology*, vol. 46(1), pp. 79–88, 1980.

- [10] A. J. LEMONTE, W. B. SOUZA, and G. M. CORDEIRO, “The exponentiated kumaraswamy distribution and its log-transform,” *Brazilian Journal of Probability and Statistics*, vol. 27(1), pp. 31–53, 2013.
- [11] S. HUANG and B. OLUYEDE, “Exponentiated kumaraswamy-dagum distribution with applications to income and lifetime data,” *Journal of Statistical Distributions and Applications*, vol. 1, pp. 1–18, 2014.
- [12] A. P. PRUDNIKOV, Y. A. BRYCHKOV, and O. I. MARICHEV, *Integrals and series*, vol. 1. Gordon and Breach, Amsterdam, 1998.
- [13] I. S. GRADSHTEYN and I. M. RYZHIK, *Table of integrals, series and products*. Academic Press, 2000.
- [14] B. C. ARNOLD, A. N. BALAKRISHNAN, and H. N. NAGARAJA, *A first course in order statistics*. Wiley-Interscience, 1992.
- [15] E. T. LEE and J. W. WANG, *Statistical Methods for Survival Data Analysis*. Wiley, New York, 2003.