

DISTRIBUIÇÃO DAGUM TRANSMUTADA: PROPRIEDADES E APLICAÇÃO

Jailson de Araujo Rodrigues¹, Ana Paula Coelho Madeira Silva²,
Ângela Lima da Silva

Resumo: Neste artigo, é introduzida uma distribuição com quatro parâmetros denominada Dagum transmutada. Algumas das principais propriedades estruturais dessa distribuição são apresentadas. Em particular, são encontradas expressões para os momentos, os desvios médios e as curvas de Bonferroni e Lorenz. A estimação dos parâmetros é discutida utilizando os métodos dos momentos e da máxima verossimilhança. Uma aplicação do modelo a um conjunto de dados de sobrevivência demonstrou que o mesmo pode apresentar um ajuste melhor que o de distribuições clássicas.

Palavras-chave: Distribuição Dagum, Momentos, Estimadores de Máxima Verossimilhança.

1 Introdução

As distribuições de probabilidade tem sido extensivamente utilizadas na modelagem de dados em diferentes domínios científicos. Esse fato faz com que sua teoria esteja sempre em pauta e novas distribuições sejam desenvolvidas.

O modelo Dagum [4, 5] foi originalmente proposto para estudar dados de renda. Ao longo dos anos, diversos autores passaram a estudar e aplicar esse modelo em diferentes contextos. Isso levou ao surgimento de generalizações e formas modificadas dessa distribuição. Por exemplo, [3] definiu e estudou as principais propriedades da distribuição beta-Dagum. [8] introduziu a distribuição Kumaraswamy exponenciada-Dagum. [13] propôs a distribuição Mc-Dagum com seis parâmetros e apresentou suas várias propriedades.

A função de distribuição acumulada (FDA) e a função densidade de probabilidade (FDP) da distribuição Dagum são dadas, respectivamente, por:

$$G(x) = (1 + \lambda x^{-\delta})^{-\beta} \tag{1}$$

e

$$g(x) = \beta \lambda \delta x^{-\delta-1} (1 + \lambda x^{-\delta})^{-\beta-1} \tag{2}$$

sendo que $x > 0$, $\beta > 0$, $\lambda > 0$ e $\delta > 0$. O parâmetro λ é um parâmetro de escala, β e δ são parâmetros de forma. O r -ésimo momento da distribuição Dagum é dado por:

$$E(X^r) = \beta \lambda^{\frac{r}{\delta}} B\left(\beta + \frac{r}{\delta}, 1 - \frac{r}{\delta}\right) \tag{3}$$

¹Instituto Federal da Bahia-IFBA, Departamento de Ensino. e-mail: jailsondearaujo@yahoo.com.br

²Universidade Federal de São João Del Rei - UFSJ, Departamento de Ciências Exatas e Biológicas.

sendo que $B(\cdot, \cdot)$ representa a função beta,

$$B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt \quad (4)$$

Neste trabalho, é introduzida e estudada uma distribuição denominada Dagum transmutada. São apresentadas algumas das principais propriedades estruturais desse modelo. 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 demonstra que o mesmo pode ser utilizado como alternativa no ajuste de dados de sobrevivência.

O texto é organizado da seguinte forma: na Seção 2 a distribuição Dagum transmutada 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 do modelo.

Os cálculos envolvidos no trabalho incluem o uso da função hipergeométrica de Gauss ${}_2F_1(a, b; c; x)$ dada por:

$${}_2F_1(a, b; c; x) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k x^k}{(c)_k k!} \quad (5)$$

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

2 O Modelo

Uma variável aleatória X possui distribuição transmutada quando sua FDA, $F(x)$, é dada por:

$$F(x) = (1 + \alpha)G(x) - \alpha G^2(x) \quad (6)$$

sendo que $|\alpha| \leq 1$ e $G(x)$ é a FDA da distribuição base. A FDP correspondente a (6) é dada por:

$$f(x) = g(x) [(1 + \alpha) - 2\alpha G(x)] \quad (7)$$

com $g(x) = dG(x)/dx$.

A generalização (6) está sendo utilizada por muito autores para construir novas classes de distribuições que são extensões dos modelos usuais. A inclusão do parâmetro adicional, α , fornece a distribuição base maior flexibilidade. Entre as distribuições transmutadas já estudadas, pode-se destacar: a distribuição Weibull transmutada [2], a distribuição Weibull transmutada exponenciada [7], a distribuição gama exponenciada transmutada [9], a distribuição Frechét transmutada [11] e a distribuição exponencial exponenciada transmutada [12].

Combinando (1) com (6), obtém-se a distribuição Dagum transmutada (DT) com FDA dada por:

$$F(x) = \left[(\alpha + 1) - \alpha (1 + \lambda x^{-\delta})^{-\beta} \right] (1 + \lambda x^{-\delta})^{-\beta} \quad (8)$$

A FDP da distribuição Dagum transmutada com parâmetros $|\alpha| \leq 1$, $\beta > 0$, $\lambda > 0$ e $\delta > 0$ é dada por:

$$f(x) = \beta \lambda \delta x^{-\delta-1} \left[(\alpha + 1) - 2\alpha (1 + \lambda x^{-\delta})^{-\beta} \right] (1 + \lambda x^{-\delta})^{-\beta-1} \quad (9)$$

Claramente, se $\alpha = 0$ a distribuição DT reduz-se ao modelo Dagum. A função de risco correspondente ao modelo (8) é dada por:

$$h(x) = \lim_{\Delta x \rightarrow 0} \frac{P(X < x + \Delta x | X > x)}{\Delta x} = \frac{f(x)}{1 - F(x)} \quad (10)$$

ou seja,

$$h(x) = \frac{\beta \lambda \delta x^{-\delta-1} \left[(\alpha + 1) - 2\alpha (1 + \lambda x^{-\delta})^{-\beta} \right] (1 + \lambda x^{-\delta})^{-1}}{(1 + \lambda x^{-\delta})^\beta - \left[(\alpha + 1) - \alpha (1 + \lambda x^{-\delta})^{-\beta} \right]} \quad (11)$$

Na Figura 1 são exibidas algumas formas possíveis da FDP (9) para diferentes valores dos parâmetros α , β , λ e δ .

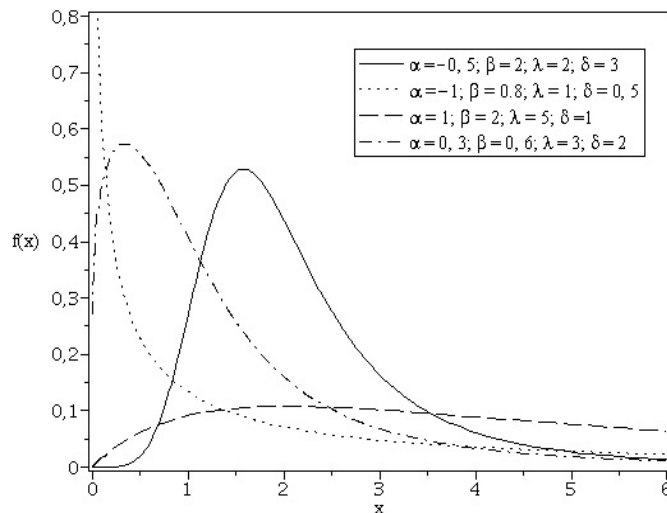


Figura 1: Gráficos da FDP da distribuição DT para diferentes valores paramétricos.

3 Propriedades do modelo

3.1 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 9, o momento de ordem r da distribuição DT pode ser escrito como:

$$\begin{aligned} E(X^r) &= \int_0^\infty x^r f(x) dx \\ &= \beta \lambda \delta \int_0^\infty x^{r-\delta-1} \left[(\alpha + 1) - 2\alpha (1 + \lambda x^{-\delta})^{-\beta} \right] (1 + \lambda x^{-\delta})^{-\beta-1} dx \\ &= \beta \lambda \delta (\alpha + 1) \int_0^\infty x^{r-\delta-1} (1 + \lambda x^{-\delta})^{-\beta-1} dx - 2\alpha \beta \lambda \delta \int_0^\infty x^{r-\delta-1} (1 + \lambda x^{-\delta})^{-2\beta-1} dx \quad (12) \end{aligned}$$

Efetuando a mudança de variável $t = \lambda x^{-\delta}$ na soma de integrais (12), pode-se escrever:

$$E(X^r) = \beta(\alpha + 1)\lambda^{\frac{r}{\delta}} \int_0^\infty t^{-\frac{r}{\delta}}(1+t)^{-\beta-1} dt - 2\alpha\beta\lambda^{\frac{r}{\delta}} \int_0^\infty t^{-\frac{r}{\delta}}(1+t)^{-2\beta-1} dt \quad (13)$$

Segundo [14], se $|\arg z| < \pi$ e $0 < \operatorname{Re} a < \operatorname{Re} \rho$, então,

$$\int_0^\infty x^{a-1}(x+z)^{-\rho} dx = z^{a-\rho} B(a, \rho-a) \quad (14)$$

Então, a aplicação da integral (14) em (13) resulta que o momento de ordem r da distribuição DT pode ser escrito como

$$E(X^r) = \beta\lambda^{\frac{r}{\delta}} \left[(\alpha + 1) B\left(1 - \frac{r}{\delta}, \beta + \frac{r}{\delta}\right) - 2\alpha\beta B\left(1 - \frac{r}{\delta}, 2\beta + \frac{r}{\delta}\right) \right] \quad (15)$$

sempre que $\delta > r$. Em particular, a média da distribuição DT é dada por:

$$\mu = E(X) = \beta\lambda^{\frac{1}{\delta}} \left[(\alpha + 1) B\left(1 - \frac{1}{\delta}, \beta + \frac{1}{\delta}\right) - 2\alpha\beta B\left(1 - \frac{1}{\delta}, 2\beta + \frac{1}{\delta}\right) \right] \quad (16)$$

A Tabela 1 contém a média da distribuição DT para diferentes valores dos parâmetros α , β , λ e δ .

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

α	β	λ	δ	Média
-1,0	0,1	1,5	1,1	3,1356
-0,7	0,5	0,2	1,1	1,7523
-0,1	0,1	0,2	1,1	0,2562
-0,1	1,0	1,5	1,5	3,0951
0,1	0,8	0,6	1,5	1,5101
0,2	0,5	0,3	2,0	0,6166
0,3	1,8	4,0	3,0	3,1917
0,3	8,0	0,5	3,0	2,7754
0,4	5,0	0,3	8,0	1,7244
0,5	7,5	2,3	6,0	3,0307
0,6	1,9	9,0	2,3	8,7363
0,6	5,0	0,3	4,5	2,3723
0,7	2,0	0,9	1,5	6,5725
0,8	0,6	4,0	7,0	2,5744
0,9	2,8	0,1	2,3	2,0503
0,9	2,0	3,0	2,3	7,6019
1,0	0,3	6,0	1,7	5,0851

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

$$\operatorname{Var}(X) = E(X^2) - E^2(X) \quad (17)$$

$$\operatorname{Ass}(X) = \frac{E(X^3) - 3E(X)E^2(X) + 2E^3(X)}{\operatorname{Var}^{3/2}(X)} \quad (18)$$

$$\operatorname{Cur}(X) = \frac{E(X^4) - 4E(X)E^3(X) + 6E(X^2)E^2(X) - 3E^4(X)}{\operatorname{Var}^2(X)} \quad (19)$$

3.2 Desvios médios

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

$$\begin{aligned}
\delta_1(X) &= E(|X - \mu|) \\
&= \int_0^\infty |x - \mu| f(x) dx \\
&= 2\mu F(\mu) - 2\mu + 2 \int_\mu^\infty x f(x) dx \\
&= 2\mu \left[(\alpha + 1) - \alpha (1 + \lambda\mu^{-\delta})^{-\beta} \right] (1 + \lambda\mu^{-\delta})^{-\beta} - 2\mu + 2 \int_\mu^\infty x f(x) dx \\
&= 2\mu \left[(\alpha + 1) - \alpha (1 + \lambda\mu^{-\delta})^{-\beta} \right] (1 + \lambda\mu^{-\delta})^{-\beta} - 2\mu \\
&\quad + 2\beta\lambda\delta \int_\mu^\infty x^{-\delta} \left[(\alpha + 1) - 2\alpha (1 + \lambda x^{-\delta})^{-\beta} \right] (1 + \lambda x^{-\delta})^{-\beta-1} dx \\
&= 2\mu \left[(\alpha + 1) - \alpha (1 + \lambda\mu^{-\delta})^{-\beta} \right] (1 + \lambda\mu^{-\delta})^{-\beta} - 2\mu \\
&\quad + 2\beta\lambda\delta (\alpha + 1) \int_\mu^\infty x^{-\delta} (1 + \lambda x^{-\delta})^{-\beta-1} dx - 4\alpha\beta\lambda\delta \int_\mu^\infty x^{-\delta} (1 + \lambda x^{-\delta})^{-2\beta-1} dx \quad (20)
\end{aligned}$$

Efetuada a mudança de variável $t = \lambda x^{-\delta}$ na soma de integrais (20), pode-se escrever:

$$\begin{aligned}
\delta_1(X) &= 2\mu \left[(\alpha + 1) - \alpha (1 + \lambda\mu^{-\delta})^{-\beta} \right] (1 + \lambda\mu^{-\delta})^{-\beta} - 2\mu \\
&\quad + 2\beta\lambda^{\frac{1}{\delta}} (\alpha + 1) \int_0^{\lambda\mu^{-\delta}} t^{-\frac{1}{\delta}} (1+t)^{-\beta-1} dx - 4\alpha\beta\lambda^{\frac{1}{\delta}} \int_0^{\lambda\mu^{-\delta}} t^{-\frac{1}{\delta}} (1+t)^{-2\beta-1} dx \quad (21)
\end{aligned}$$

Segundo [6], se $0 < \operatorname{Re} a$ e $|\arg(1 + bk)| < \pi$, então,

$$\int_0^k x^{a-1} (1 + bx)^{-\nu} dx = \frac{k^a}{a} {}_2F_1(\nu, a; a + 1; -bk) \quad (22)$$

Aplicando a integral (22) em (21), obtém-se o desvio médio em relação a média da distribuição DT,

$$\begin{aligned}
\delta_1(X) &= 2\mu \left[(\alpha + 1) - \alpha (1 + \lambda\mu^{-\delta})^{-\beta} \right] (1 + \lambda\mu^{-\delta})^{-\beta} - 2\mu \\
&\quad + \frac{2\beta\lambda\delta (\alpha + 1) \mu^{1-\delta}}{\delta - 1} {}_2F_1\left(\beta + 1, 1 - \frac{1}{\delta}; 2 - \frac{1}{\delta}; -\lambda\mu^{-\delta}\right) \\
&\quad - \frac{4\alpha\beta\lambda\delta \mu^{1-\delta}}{\delta - 1} {}_2F_1\left(2\beta + 1, 1 - \frac{1}{\delta}; 2 - \frac{1}{\delta}; -\lambda\mu^{-\delta}\right) \quad (23)
\end{aligned}$$

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

$$\begin{aligned}
\delta_1(X) &= E(|X - m|) \\
&= \int_0^\infty |x - m| f(x) dx \\
&= -\mu + 2 \int_m^\infty x f(x) dx \\
&= -\mu + \frac{2\beta\lambda\delta (\alpha + 1) m^{1-\delta}}{\delta - 1} {}_2F_1\left(\beta + 1, 1 - \frac{1}{\delta}; 2 - \frac{1}{\delta}; -\lambda m^{-\delta}\right) \\
&\quad - \frac{4\alpha\beta\lambda\delta m^{1-\delta}}{\delta - 1} {}_2F_1\left(2\beta + 1, 1 - \frac{1}{\delta}; 2 - \frac{1}{\delta}; -\lambda m^{-\delta}\right) \quad (24)
\end{aligned}$$

3.3 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:

$$B(p) = \frac{1}{p\mu} \int_0^q xf(x) dx \quad (25)$$

e

$$L(p) = \frac{1}{\mu} \int_0^q xf(x) dx \quad (26)$$

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

$$\begin{aligned} B(p) &= \frac{1}{p\mu} \int_0^q xf(x) dx \\ &= \frac{\beta\lambda\delta}{p\mu} \int_0^q x^{-\delta} \left[(\alpha + 1) - 2\alpha (1 + \lambda x^{-\delta})^{-\beta} \right] (1 + \lambda x^{-\delta})^{-\beta-1} dx \\ &= \frac{\beta\lambda\delta(\alpha + 1)}{p\mu} \int_0^q x^{-\delta} (1 + \lambda x^{-\delta})^{-\beta-1} dx - \frac{2\alpha\beta\lambda\delta}{p\mu} \int_0^q x^{-\delta} (1 + \lambda x^{-\delta})^{-2\beta-1} dx \end{aligned} \quad (27)$$

Efetuada a mudança de variável $t = \lambda x^{-\delta}$ na soma de integrais (27), pode-se escrever:

$$B(p) = \frac{\beta\lambda^{\frac{1}{\delta}}}{p\mu} (\alpha + 1) \int_{\lambda q^{-\delta}}^{\infty} t^{-\frac{1}{\delta}} (1+t)^{-\beta-1} dx - \frac{2\alpha\beta\lambda^{\frac{1}{\delta}}}{p\mu} \int_{\lambda q^{-\delta}}^{\infty} t^{-\frac{1}{\delta}} (1+t)^{-2\beta-1} dx \quad (28)$$

Segundo [6], se $Re a < Re \nu$, então,

$$\int_k^{\infty} x^{a-1} (1+bx)^{-\nu} dx = \frac{k^{a-\nu}}{b^{\nu}(\nu-a)} {}_2F_1\left(\nu, \nu-a; \nu-a+1; -\frac{1}{bk}\right) \quad (29)$$

Aplicando a integral (29) em (28), obtém-se a curva de Bonferroni para a distribuição DT,

$$\begin{aligned} B(p) &= \frac{\beta\delta(\alpha+1)q^{\beta\delta+1}}{p\delta\mu(\beta\delta+1)\lambda^{\beta}} {}_2F_1\left(\beta+1, \beta+\frac{1}{\delta}; \beta+1+\frac{1}{\delta}; -\frac{q^{\delta}}{\lambda}\right) \\ &\quad - \frac{2\alpha\beta\delta q^{2\beta\delta+1}}{p\delta\mu(2\beta\delta+1)\lambda^{2\beta}} {}_2F_1\left(2\beta+1, 2\beta+\frac{1}{\delta}; 2\beta+1+\frac{1}{\delta}; -\frac{q^{\delta}}{\lambda}\right) \end{aligned} \quad (30)$$

Por outro lado, a curva de Lorenz para a distribuição DT é dada por:

$$\begin{aligned} L(p) &= \frac{\beta\delta(\alpha+1)q^{\beta\delta+1}}{\delta\mu(\beta\delta+1)\lambda^{\beta}} {}_2F_1\left(\beta+1, \beta+\frac{1}{\delta}; \beta+1+\frac{1}{\delta}; -\frac{q^{\delta}}{\lambda}\right) \\ &\quad - \frac{2\alpha\beta\delta q^{2\beta\delta+1}}{\delta\mu(2\beta\delta+1)\lambda^{2\beta}} {}_2F_1\left(2\beta+1, 2\beta+\frac{1}{\delta}; 2\beta+1+\frac{1}{\delta}; -\frac{q^{\delta}}{\lambda}\right) \end{aligned} \quad (31)$$

3.4 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 DT. Suponha que X_1, X_1, \dots, X_n é uma amostra aleatória da distribuição DT. Se $X_{1:n} < X_{2:n} < \dots < X_{n:n}$ denota a estatística de ordem correspondente. Segundo [1], a FDP e a FDA da r -ésima estatística de ordem, denotada por $Y = X_{r:n}$, são dadas, respectivamente, por:

$$\begin{aligned}
f_Y(y) &= \frac{n!}{(r-1)!(n-r)!} F^{r-1}(y) [1 - F(y)]^{n-r} f(y) \\
&= \frac{n!}{(r-1)!(n-r)!} \sum_{l=0}^{n-r} \binom{n-r}{l} (-1)^l F^{l+r-1}(y) f(y)
\end{aligned} \tag{32}$$

e

$$\begin{aligned}
F_Y(y) &= \sum_{j=r}^n \binom{n}{j} F^j(y) [1 - F(y)]^{n-j} \\
&= \sum_{j=r}^n \sum_{l=0}^{n-j} \binom{n}{j} \binom{n-j}{l} (-1)^l F^{j+l}(y)
\end{aligned} \tag{33}$$

sendo que $f(\cdot)$ e $F(\cdot)$ representam a FDP e FDA da distribuição DT. Das Equações 8 e 9, tem-se:

$$\begin{aligned}
f_Y(y) &= \frac{\beta \lambda \delta y^{-\delta-1} n!}{(r-1)!(n-r)!} \left[(\alpha + 1) - 2\alpha (1 + \lambda y^{-\delta})^{-\beta} \right] \\
&\quad \times \sum_{l=0}^{n-r} \binom{n-r}{l} (-1)^l \left[(\alpha + 1) - \alpha (1 + \lambda y^{-\delta})^{-\beta} \right]^{l+r-1} (1 + \lambda y^{-\delta})^{-\beta(l+r)-1}
\end{aligned} \tag{34}$$

e

$$F_Y(y) = \sum_{j=r}^n \sum_{l=0}^{n-j} \binom{n}{j} \binom{n-j}{l} (-1)^l \left[(\alpha + 1) - \alpha (1 + \lambda y^{-\delta})^{-\beta} \right]^{j+l} (1 + \lambda y^{-\delta})^{-\beta(j+l)} \tag{35}$$

4 Estimação dos parâmetros do modelo

Nesta seção, é analisada a estimação dos quatro parâmetros da distribuição (8) 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 DT. Considerando o método dos momentos, ao igualar-se os momentos teóricos $E(X^r)$ com os respectivos momentos amostrais,

$$M_r = \frac{1}{n} \sum_{l=1}^n x_l^r, \quad r = 1, \dots, 4. \tag{36}$$

obtém-se o seguinte sistema de equações:

$$M_r = \beta \lambda^{\frac{r}{\delta}} \left[(\alpha + 1) B \left(1 - \frac{r}{\delta}, \beta + \frac{r}{\delta} \right) - 2\alpha \beta B \left(1 - \frac{r}{\delta}, 2\beta + \frac{r}{\delta} \right) \right] \tag{37}$$

O sistema não linear obtido, pode ser resolvido numericamente e suas raízes são as estimativas dos parâmetros α , β , λ e δ .

Por outro lado, levando-se em consideração a estimação dos parâmetros pelo método da máxima verossimilhança, a log-verossimilhança para uma amostra aleatória x_1, \dots, x_n da distribuição DT é dada por:

$$\begin{aligned}
\log L(\alpha, \beta, \lambda, \delta) &= n \log \beta + n \log \lambda + n \log \delta - (\delta + 1) \sum_{i=1}^n \log x_i - (\beta + 1) \sum_{i=1}^n \log (1 + \lambda x_i^{-\delta}) \\
&\quad + \sum_{i=1}^n \log \left[(\alpha + 1) - 2\alpha (1 + \lambda x_i^{-\delta})^{-\beta} \right]
\end{aligned} \tag{38}$$

Diferenciando (38) em relação a cada um dos parâmetros α , β , λ , δ e igualando os resultados a zero, tem-se:

$$\frac{\partial \log L}{\partial \alpha} = \sum_{i=1}^n \log \frac{1 - 2(1 + \lambda x_i^{-\delta})^{-\beta}}{(\alpha + 1) - 2\alpha(1 + \lambda x_i^{-\delta})^{-\beta}} = 0 \quad (39)$$

$$\frac{\partial \log L}{\partial \beta} = \frac{n}{\beta} - \sum_{i=1}^n \log(1 + \lambda x_i^{-\delta}) + 2\alpha \sum_{i=1}^n \log \frac{(1 + \lambda x_i^{-\delta})^{-\beta} \log(1 + \lambda x_i^{-\delta})}{(\alpha + 1) - 2\alpha(1 + \lambda x_i^{-\delta})^{-\beta}} = 0 \quad (40)$$

$$\frac{\partial \log L}{\partial \lambda} = \frac{n}{\lambda} - (\beta + 1) \sum_{i=1}^n \frac{x_i^{-\delta}}{1 + \lambda x_i^{-\delta}} + 2\alpha\beta \sum_{i=1}^n \frac{x_i^{-\delta} (1 + \lambda x_i^{-\delta})^{-\beta-1}}{(\alpha + 1) - 2\alpha(1 + \lambda x_i^{-\delta})^{-\beta}} = 0 \quad (41)$$

$$\frac{\partial \log L}{\partial \delta} = \frac{n}{\delta} - \sum_{i=1}^n \log x_i + (\beta + 1) \sum_{i=1}^n \frac{\lambda x_i^{-\delta} \log x_i}{1 + \lambda x_i^{-\delta}} - 2\alpha\beta\lambda \sum_{i=1}^n \frac{x_i^{-\delta} (1 + \lambda x_i^{-\delta})^{-\beta-1} \log x_i}{(\alpha + 1) - 2\alpha(1 + \lambda x_i^{-\delta})^{-\beta}} = 0 \quad (42)$$

As estimativas de máximo verossimilhança $\hat{\alpha}$, $\hat{\beta}$, $\hat{\lambda}$ e $\hat{\delta}$ dos parâmetros desconhecidos α , β , λ e δ são obtidas resolvendo simultaneamente as Equações 39–42.

5 Aplicação

Nesta seção, é feito o ajuste da distribuição DT a um conjunto de dados de sobrevivência corresponde ao tempo de remissão (em meses) do câncer de bexiga diagnosticado em cento e vinte e oito pacientes. Esses dados foram retirados de [10].

O ajuste do modelo DT é comparado com outras quatro distribuições: gama, log-normal, Lomax e Pareto. As FDPs das distribuições gama, log-normal, Lomax e Pareto são dadas por:

1. Gama:

$$f(x) = \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} \exp(-\beta x) \quad (43)$$

2. Log-normal (LN):

$$f(x) = \frac{1}{\beta x \sqrt{2\pi}} \exp \left[-\frac{(\log x - \alpha)^2}{2\beta^2} \right] \quad (44)$$

3. Lomax:

$$f(x) = \frac{\alpha}{\beta} \left(1 + \frac{x}{\beta} \right)^{-(\alpha+1)} \quad (45)$$

4. Pareto:

$$f(x) = \alpha \beta^\alpha x^{-(\alpha+1)} \quad (46)$$

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 DT apresenta um ajuste melhor se comparado com os demais modelos.

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

Modelos	α	β	λ	δ	AIC
DT	-0,822	0,397	49,001	1,889	828,970
Gama	1,173	0,125	830,736
LN	1,753	1,152	834,189
Lomax	13,938	121,022	831,670
Pareto	0,080	0,234	1081,046

Na Figura 2 são apresentados os gráficos de probabilidade para os cinco modelos avaliados. Ao observar a figura, pode-se notar que a distribuição DT apresentou melhor ajuste em relação as demais distribuições e o pior o ajuste foi apresentado pelo modelo Pareto. Dessa forma, os gráficos de probabilidade corroboram com os resultados obtidos anteriormente na análise dos valores de AIC.

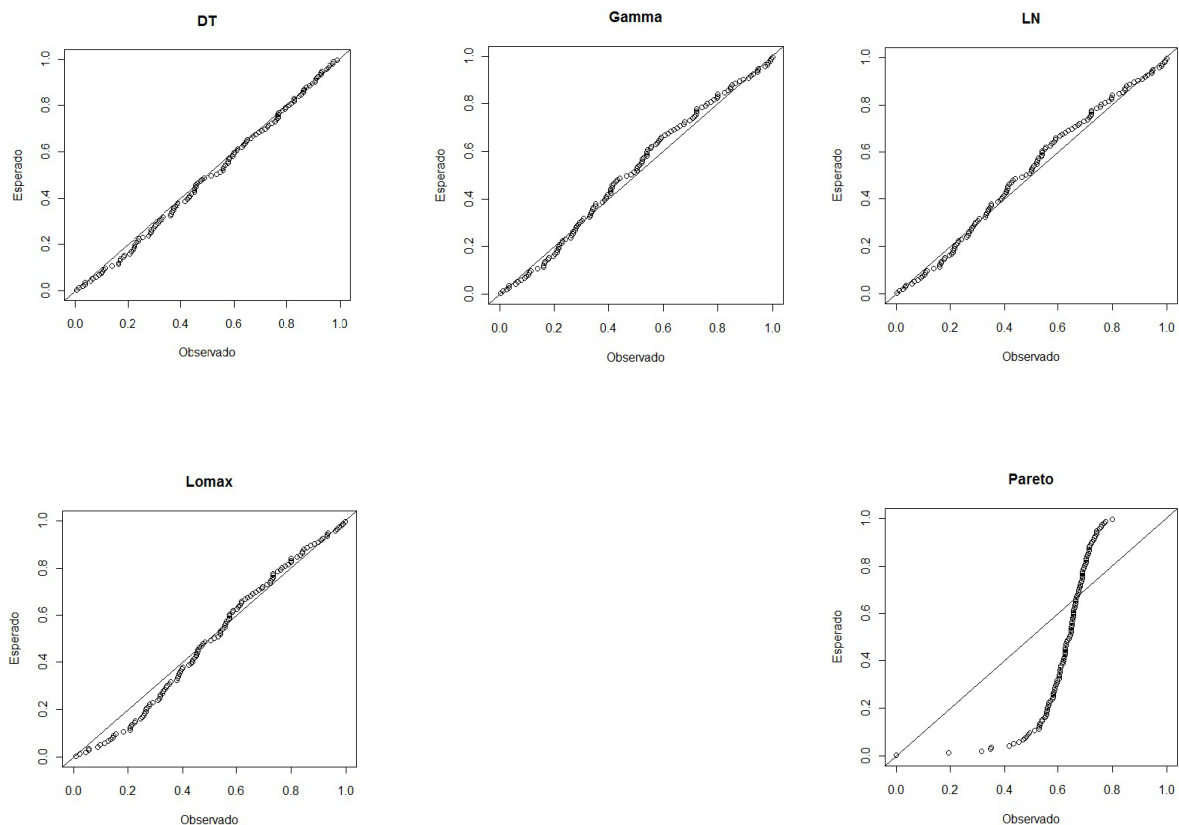


Figura 2: Gráficos de probabilidade das distribuições ajustadas.

Na Figura 3 pode-se comparar o histograma de frequências dos dados observados sobreposto pelas FDPs DT, gama, LN, Lomax e Pareto que foram ajustadas. O que se pode notar é que as curvas de ajustes das distribuições DT, gama, LN e Lomax acompanham as condições de moda e assimetria, o que não ocorre com o modelo Pareto. É possível determinar, visualmente, que o melhor ajuste foi apresentado pela distribuição DT.

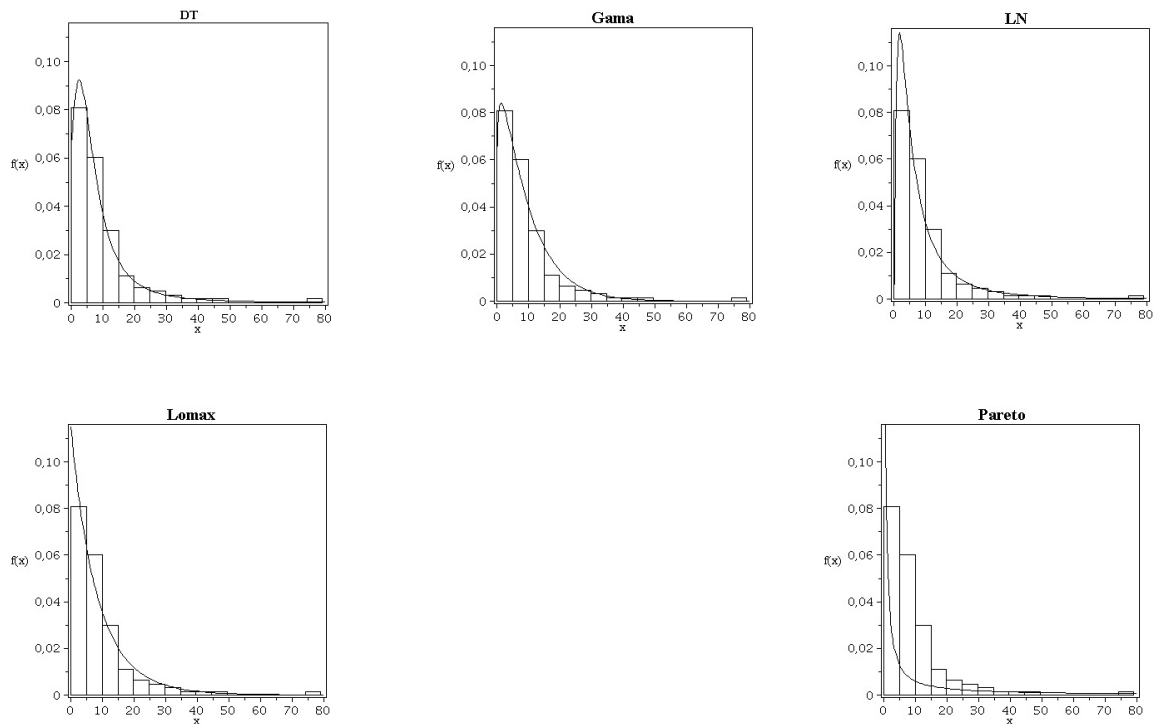


Figura 3: Histograma dos dados e FDPs ajustadas.

6 Conclusões

O artigo introduziu uma distribuição com quatro parâmetros denominada Dagum transmutada. Foram deduzidas expressões para os momentos, desvios médios, curva de Bonferroni e curva de Lorenz. Além disso, foi apresentada a estatística de ordem para o modelo proposto. A estimação dos parâmetros foi discutida utilizando os métodos dos momentos e da máxima verossimilhança. A aplicação do modelo no ajuste de um conjunto de dados corresponde ao tempo de remissão do câncer de bexiga, diagnosticado em cento e vinte e oito pacientes, revelou o potencial da distribuição Dagum transmutada na análise de dados de sobrevivência.

Referências

- [1] ARNOLD, B.C.; BALAKRISHNAN, A.N.; NAGARAJA, H.N., “A first course in order statistics,” *New York: Wiley-Interscience*, 1992. 279p.
- [2] ARYAL, G. R.; TSOKOS, C. P. “Transmuted Weibull distribution: a generalization of the Weibull probability distribution,” *European Journal of Pure and Applied Mathematics*, v. 4, n. 2, p. 89–102, 2011.
- [3] CONDINO, F.; DOMMA, F.; LEE, C., “The beta-Dagum distribution,” *Commun. Stat. - Theory and Methods*, v. 42, n. 22, p. 4070–4090, 2013.

- [4] DAGUM, C., “A new model of personal income distribution: specification and estimation,” *Economie Appliquée*, v. 30, p. 413–437, 1997.
- [5] DAGUM, C., “The generalization and distribution of income, the Lorenz curve and the Gini ratio,” *Economie Appliquée*, v. 33, p. 327–367, 1980.
- [6] GRADSHTEYN, I.S.; RYZHIK, I.M. “Table of integrals, series and products,” *San Diego: Academic Press*, 2000. 1221p.
- [7] HADY, A. E.; EBRAHEIN, N. “Exponentiated transmuted Weibull distribution: a generalization of the Weibull distribution,” *World Academy of Science, Engineering and Technology*, v. 8, n. 6, p. 831–839, 2014.
- [8] HUANG, S.; OLUEYEDE, B. O. “Exponentiated Kumaraswamy-Dagum distribution with applications to income and lifetime data,” *Journal of Statistical Distributions and Applications*, v. 1, p. 1–18, 2014.
- [9] HUSSIAN, M. A. “Transmuted exponentiated gamma distribution: a generalization of the exponentiated gamma probability distribution,” *Applied Mathematical Sciences*, v. 8, n. 27, p. 1297–1310, 2014.
- [10] LEE, E.T.; WANG, J.W. “Statistical Methods for Survival Data Analysis,” *New York: Wiley*, 2003. 512p.
- [11] MAHMOUD, M. R.; MANDOUH, R. M. “On the transmuted Frechet distribution,” *Journal of Applied Sciences Research*, v. 9, n. 10, p. 5553–5561, 2013.
- [12] MEROVCI, F. “Transmuted exponentiated exponential distribution,” *Mathematical Sciences and Applications E-Notes*, v. 1, n. 2, p. 112–122, 2013.
- [13] OLUYEDE, B. O.; RAJASOORIYA, S. “The Mc-Dagum distribution and its statistical properties with applications,” *Asian Journal of Mathematics and Applications*, v. 2013, p. 1–16, 2013.
- [14] PRUDNIKOV, A. P.; BRYCHKOV, Y. A.; MARICHEV, O. I., “Integrals and series, vol.1,” *Amsterdam: Gordon and Breach*, 1998. 808p.