PLISKA
STUDIA MATHEMATICA

ПЛИСКА
**МАТЕМАТИЧЕСКИ
СТУДИИ**

Pliska Stud. Math. **30** (2019), 117–126

PLISKA
STUDIA MATHEMATICA

A NOTE ON THE DAGUM–II SIGMOID FUNCTION WITH APPLICATIONS TO WEALTH DATA. OTHER APPLICATIONS*

Nikolay Kyurkchiev, Anton Iliev

The Dagum–II distribution is a flexible and simple model with applications to wealth data.

We prove upper and lower estimates for the Hausdorff approximation of the shifted Heaviside function $\tilde{h}_{r,t_0}(t)$ by a class of Dagum–II cumulative distribution function – (D–IICDF). Numerical examples, illustrating our results are given.

1. Introduction

Dagum (1977) [2] motivates his model from the empirical observation that the income elasticity $\eta(F,t)$ of the cumulative distribution function (CDF) F of income is a decreasing and bounded function F.

Starting from the differential equation

$$\eta(F,t) = \frac{d\,\log F(t)}{d\,\log t} = ap\left(1 - (F(t))^{\frac{1}{p}}\right), \quad t \geq 0, \tag{1}$$

subject to $p > 0$ and $ap > 0$, one obtains (see, also Kleiber [1]):

$$F(t) = \left(1 + \left(\frac{t}{b}\right)^{-a}\right)^{-p}. \tag{2}$$

2010 *Mathematics Subject Classification:* 41A46.

Key words: Dagum–II cumulative distribution function (D–IICDF), shifted Heaviside function, Hausdorff distance, upper and lower bounds.

*This work has been supported by the project FP17-FMI-008 of Department for Scientific Research, Paisii Hilendarski University of Plovdiv.

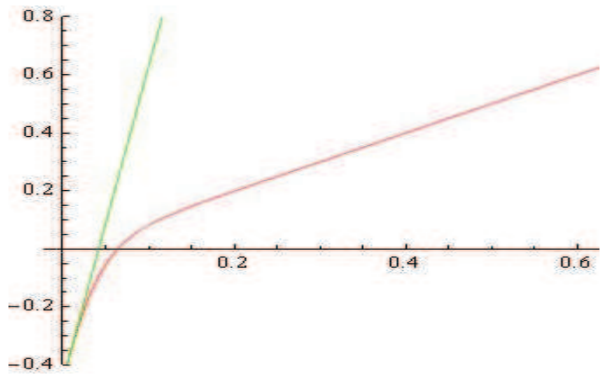

Figure 1: The functions H and G.

This approach was further developed in a series of papers on generating systems for income distribution [3]–[6].

For other results, see [7], [8], [9], [10] and [29].

The Dagum–II type distribution has the following cumulative function(D–IICDF):

$$F(t) = r + (1-r)\left(1 + \left(\frac{t}{b}\right)^{-a}\right)^{-p}; \quad t > 0, \tag{3}$$

where $a, b, p > 0$ and $r \in (0, 1)$.

Clearly, this is a mixture of a point mass at the origin with a Dagum (type I) distribution over the positive halfline.

The Dagum–II type distribution was proposed as a model for income distributions with null and negative incomes, but more particularly to fit wealth data (see Kleiber and Kotz (2000) [7] for further details).

In this paper we prove upper and lower estimates for the Hausdorff approximation of the shifted Heaviside function $\tilde{h}_{r,t_0}(t)$ by a class of Dagum–II cumulative distribution function – (D–IICDF).

2. Preliminaries

Definition 1. *The (basic) step function is:*

$$\tilde{h}_{r,t_0}(t) = \begin{cases} r, & if \quad t < t_0, \\ [r,1], & if \quad t = t_0, \\ 1, & if \quad t > t_0, \end{cases} \tag{4}$$

usually known as shifted Heaviside function.

Definition 2. ([11], [12]) *The Hausdorff distance (the H–distance)* [11] $\rho(f,g)$ *between two interval functions* f, g *on* $\Omega \subseteq \mathbb{R}$, *is the distance between their completed graphs* $F(f)$ *and* $F(g)$ *considered as closed subsets of* $\Omega \times \mathbb{R}$.

More precisely,

$$\rho(f,g) = \max\left\{ \sup_{A\in F(f)} \inf_{B\in F(g)} \|A-B\|, \sup_{B\in F(g)} \inf_{A\in F(f)} \|A-B\| \right\}, \tag{5}$$

wherein $\|\cdot\|$ *is any norm in* $\mathbb{R}^2$, *e. g. the maximum norm* $\|(t,x)\| = \max\{|t|,|x|\}$; *hence the distance between the points* $A = (t_A, x_A)$, $B = (t_B, x_B)$ *in* $\mathbb{R}^2$ *is* $\|A-B\| = \max(|t_A - t_B|, |x_A - x_B|)$.

Let us point out that the Hausdorff distance is a natural measuring criteria for the approximation of bounded discontinuous functions.

3. Main results

Let us consider the following four parametric sigmoid function

$$F^*(t) = r + (1-r)\left(1 + \left(\frac{t}{b}\right)^{-a}\right)^{-p} \tag{6}$$

with

$$F^*(t_0) = \frac{1+r}{2}, \quad t_0 = b\left(\left(\frac{1}{2}\right)^{-\frac{1}{p}} - 1\right)^{-\frac{1}{a}}. \tag{7}$$

The H-distance $d = \rho(\tilde{h}_{r,t_0}, F^*)$ between the shifted Heaviside step function $\tilde{h}_{r,t_0}$ and the sigmoidal function F^* satisfies the relation:

$$F^*(t_0+d) = r + (1-r)\left(1 + \left(\frac{t_0+d}{b}\right)^{-a}\right)^{-p} = 1 - d. \tag{8}$$

The following theorem gives upper and lower bounds for $d = d(a,b,p,r)$

Theorem 3. *Let*

$$\alpha = -\frac{1-r}{2}; \ \beta = 1 + \frac{ap(1-r)}{b}\left(\frac{1}{2}\right)^{\frac{1+p}{p}}\left(\left(\frac{1}{2}\right)^{-\frac{1}{p}} - 1\right)^{\frac{1+a}{a}}. \tag{9}$$

The H-distance d between the function $\tilde{h}_{r,t_0}$ and the function F^* can be expressed in terms of the parameters a, b, p, r for any real $\beta \geq \frac{e^{1.05}}{2.1}(1-r)$ as follows:

$$d_l = \frac{1}{2.1\frac{\beta}{1-r}} < d < \frac{\ln(2.1\frac{\beta}{1-r})}{2.1\frac{\beta}{1-r}} = d_r. \tag{10}$$

P r o o f. We define the functions

$$H(d) = F^*(t_0+d) - 1 + d = r + (1-r)\left(1+\left(\frac{t_0+d}{b}\right)^{-a}\right)^{-p} - 1 + d \tag{11}$$

$$G(d) = \alpha + \beta d. \tag{12}$$

From Taylor expansion

$$H(d) - G(d) = O(d^2)$$

we see that the function $G(d)$ approximates $H(d)$ with $d \to 0$ as $O(d^2)$ (cf. Fig. 1).

In addition $G'(d) > 0$ and for $\beta \geq \frac{e^{1.05}}{2.1}(1-r)$

$$G(d_l) < 0; \quad G(d_r) > 0.$$

This completes the proof of the inequalities (10). □

The generated sigmoidal functions $F^*(t)$ for $a = 20$; $b = 0.7$; $p = 1.1$; $r = 0.2$; $a = 25$; $b = 0.6$; $p = 1.2$; $r = 0.1$ and $a = 26$; $b = 0.5$; $p = 1.3$; $r = 0.05$ are visualized on Fig. 2–Fig. 4.

From the Fig. 2–Fig. 4 it can be seen that the "supersaturation" is fast.

Dagum (1977) [2] in a period when individual data were rarely available, minimized

$$\sum_{i=1}^{n}\left(F_n(t_i) - \left(1+\left(\frac{t_i}{b}\right)^{-a}\right)^{-p}\right)^2.$$

a non–linear least–squares criterion based on the distance between the empirical SDF F_n and the CDF of a Dagum approximation.

The appropriate least–square fitting of the real wealth data by the Dagum model yields for $a = 32.6281$, $b = 0.409036$, $p = 0.884619$ and $r \approx 0$ and is visualized on Fig. 5.

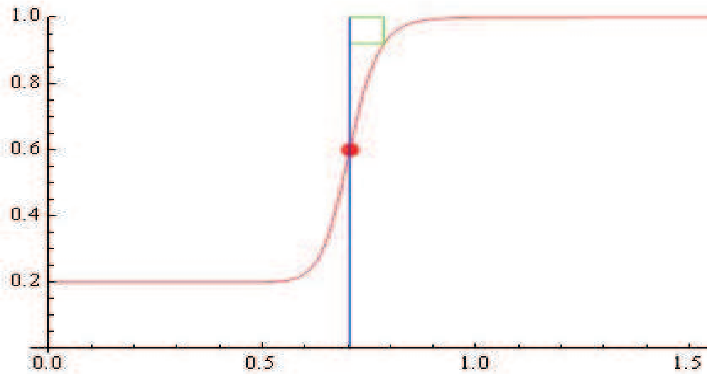

Figure 2: The function $F^*(t)$ for $a = 20$; $b = 0.7$; $p = 1.1$; $r = 0.2$; $t_0 = 0.704574$; H-distance $d = 0.0804123$; $d_l = 0.0696314$; $d_r = 0.185536$.

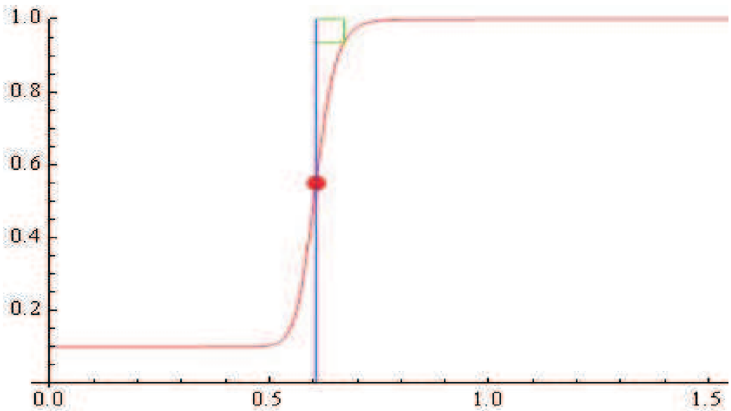

Figure 3: The function $F^*(t)$ for $a = 25$; $b = 0.6$; $p = 1.2$; $r = 0.1$; $t_0 = 0.605937$; H-distance $d = 0.0640358$; $d_l = 0.040358$; $d_r = 0.137843$.

We examine the following data presented in Table 1, was proposed in [10].

The week index is from 1 week to 18 weeks, and there are 176 cumulative failures at 18 weeks in Dataset.

The fitted model

$$F_1(t) = \omega \left(r + (1-r) \left(1 + \left(\frac{t}{b} \right)^{-a} \right)^{-p} \right)$$

based on the data of Table 1 for the estimated parameters: $\omega = 176$; $a = 4.011$; $b = 3.6415$; $p = 21.0737$; $r = 0.162082$ is plotted on Fig. 6.

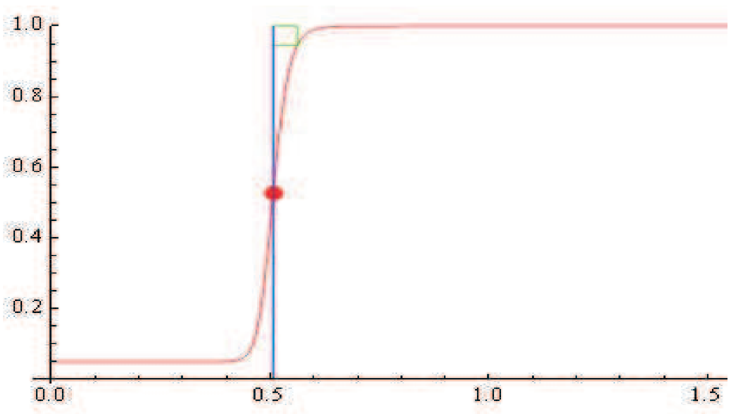

Figure 4: The function $F^*(t)$ for $a = 26$; $b = 0.5$; $p = 1.3$; $r = 0.05$; $t_0 = 0.506785$; H-distance $d = 0.0554515$; $d_l = 0.0337905$; $d_r = 0.114468$.

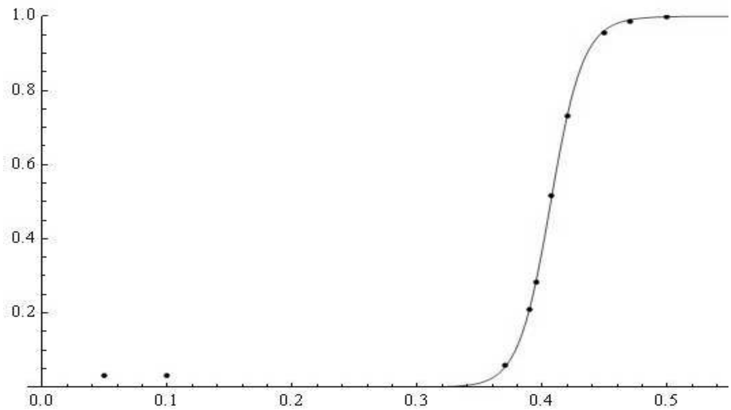

Figure 5: The appropriate least–square fitting of the real wealth data by the Dagum model yields for $a = 32.6281$, $b = 0.409036$, $p = 0.884619$ and $r \approx 0$.

The example results show a good fit to the presented model. Some of the existing modifications of the cumulative distribution of Dagum – type are considered in the light of modern debugging and test theory. For more details see monographs [27] and [28].

4. Conclusion

In this paper we prove upper and lower estimates for the Hausdorff approximation of the shifted Heaviside function $\tilde{h}_{r,t_0}(t)$ by a class of Dagum–II cumulative distribution function – (D–IICDF).

Week Index	Failures	Cumulative failures
1	28	28
2	1	29
3	0	29
4	0	29
5	0	29
6	8	37
7	26	63
8	29	92
9	24	116
10	9	125
11	14	139
12	13	152
13	12	164
14	0	164
15	1	165
16	3	168
17	2	170
18	6	176

Table 1: Dataset [10]

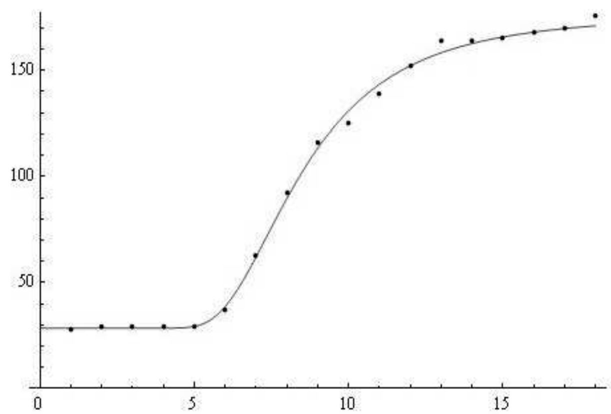

Figure 6: The fitted model $F_1(t)$ with $\omega = 176$; $a = 4.011$; $b = 3.6415$; $p = 21.0737$; $r = 0.162082$.

A family of four parametric sigmoidal functions based on Dagum–II cumulative distribution function – (DCDF) is introduced finding application in income theory.

In this note we consider dependence of supersaturation by means of this class.

Numerical examples, illustrating our results are given.

We propose a software module (intellectual property) within the programming environment *CAS Mathematica* for the analysis of the considered family of (D–IICDF) functions.

For other results, see [13]–[26].

REFERENCES

[1] Ch. Kleiber. A Guide to the Dagum Distribution. In: Modeling Income Distributions and Lorenz Curves: Essays in Memory of Camilo Dagum (Ed. D. Chotikapanich). New York, Springer–Verlag, 2008, ISBN 178-0-387-72796-7.

[2] C. Dagum. A new method of personal income distribution: Specification and estimation. *Economie Appliquée* **33** (1977), 413–437.

[3] C. Dagum. Generating systems and properties of income distribution models. *Metron* **38** (1980), 3–26.

[4] C. Dagum. Sistemas generadores de distribucion de ingreso y la ley de Pareto. *El Trimestre Economico* **47** (1980), 877–917.

[5] C. Dagum. Income distribution models. In: Encyclopedia of Statistical Sciences vol. **4** (Eds S. Kotz, N.L. Johnson, C. Read). New York, John Wiley, 1983, 27–34.

[6] C. Dagum. Generation and properties of income distribution functions. In: Income and Wealth Distribution, Inequality and Poverty: Proceedings of the Second International Conference on Income Distribution by Size: Generation, Distribution, Measurement and Applications (Eds C. Dagim, M. Zenga) New York–Berlin–Tokyo, Springer, 1990, 7–17.

[7] Ch. Kleiber, S. Kotz. Statistical size distribution in economics and actuarial sciences. Wiley Series in Probability and Statistics. Hoboken, NJ, Wiley-Interscience,2003.

[8] F. Domma, P. Perri. Some developments on the log–Dagum distribution. *Stat. Methods Appl.* **18**, *2* (2009), 205–220.

[9] S. Huang, B. Oluyede. Exponential Kumaraswamy–Dagum distribution with applications to income and lifetime data. *J. Stat. Distrib. Appl.* **1** (2014), Paper No 8, 20 pp.

[10] B. Oluyede, Sh. Huang, M. Pararai. A new class of generalized Dagum distribution with applications to income and lifetime data. *J. Statistical and Econometric Methods* **30**, *2* (2014), 125–151.

[11] F. Hausdorff Set Theory, 2nd ed. New York, Chelsea Publishing Co., 1962

[12] B. Sendov. Hausdorff Approximations. Mathematics and its Applications (East European Series) vol. **50**. Dordrecht, Kluwer Academic Publishers Group, 1990.

[13] N. Kyurkchiev. A note on the Volmer's activation (VA) function. *C. R. Acad. Bulgare Sci.* **70**, *6* (2017), 769–776.

[14] N. Kyurkchiev. A family of recurrence generated parametric functions based on Volmer-Weber-Kaishew activation function. *Pliska Stud. Math.* **29** (2018), 69–80.

[15] N. Kyurkchiev. A family of recurrence generated sigmoidal functions based on the Verhulst logistic function. Some approximation and modeling aspects. *Biomath Communications* **3**, *2* (2016) DOI: http://dx.doi.org/10.11145/bmc.2016.12.171.

[16] N. Kyurkchiev, S. Markov. Hausdorff approximation of the sign function by a class of parametric activation functions. *Biomath Communications* **3**, *2* (2016) DOI: http://dx.doi.org/10.11145/bmc.2016.12.217.

[17] A. Iliev, N. Kyurkchiev, S. Markov. A family of recurrence generated parametric activation functions with application to neural networks. *Int. J. Research Innovations in Engineering Science and Technology (IJRIEST)* **2**, *1* (2017), 60–69.

[18] N. Kyurkchiev, A. Iliev, S. Markov. Families of recurrence generated three and four parametric activation functions. *Int. J. Sci. Res. and Development (IJSRD)* **4**, *12* (2017), 746–750.

[19] V. Kyurkchiev, N. Kyurkchiev. A family of recurrence generated functions based on Half-hyperbolic tangent activation functions. *Biomedical Statistics and Informatics* **2**, *3* (2017), 87–94.

[20] V. Kyurkchiev, A. Iliev, N. Kyurkchiev. On some families of recurrence generated activation functions. *Int. J. Sci. Eng. and Appl. Sci. (IJSEAS)* **3**, *3* (2017), 243–248.

[21] A. Golev, T. Djamiykov, N. Kyurkchiev. Sigmoidal functions in antenna-feeder technique. *Int. J. Pure Appl. Math.* **116**, *4* (2017), 1081–1092.

[22] N. Kyurkchiev. A new transmuted cumulative distribution function based on Verhulst logistic function with application in population dynamics. *Biomath Communications* **4**, *1* (2017), 15 pp. `http://dx.doi.org/10.11145/bmc.2017.05.132`

[23] N. Kyurkchiev, S. Markov. On the Hausdorff distance between the Heaviside step function and Verhulst logistic function. *J. Math. Chem.* **54**, *1* (2016), 109–119. DOI: 10.1007/s10910-015-0552-0

[24] N. Kyurkchiev, A. Iliev, S. Markov. Some techniques for recurrence generating of activation functions. LAP LAMBERT Academic Publishing, 2017, ISBN 978-3-330-33143-3.

[25] N. Kyurkchiev, A. Iliev. Extension of Gompertz-type Equation in Modern Science: 240 Anniversary of the birth of B. Gompertz, LAP LAMBERT Academic Publishing, 2018, ISBN 978-613-9-90569-0.

[26] N. Kyurkchiev, S. Markov. Sigmoid functions: Some Approximation and Modelling Aspects. LAP LAMBERT Academic Publishing, Saarbrucken, 2015, ISBN 978-3-659-76045-7.

[27] Pavlov N., A. Iliev, A. Rahnev, N. Kyurkchiev. Some software reliability models: Approximation and modeling aspects, LAP LAMBERT Academic Publishing, 2018, ISBN 978-613-9-82805-0.

[28] N. Pavlov, A. Iliev, A. Rahnev, N. Kyurkchiev. Nontrivial Models in Debugging Theory (Part 2), LAP LAMBERT Academic Publishing, 2018, ISBN 978-613-9-87794-2.

[29] N. Kyurkchiev, A. Iliev. A note on Dagum sigmoid function with applications to income and lifetime data. *International Journal of Latest Research in Engineering and Technology*, **4**, *7* (2018), 1–5.

Nikolay Kyurkchiev
Faculty of Mathematics and Informatics
University of Plovdiv Paisii Hilendarski
24, Tzar Asen Str.
4000 Plovdiv, Bulgaria
e-mail: `nkyurk@uni-plovdiv.bg`
and
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Acad. G. Bonchev Str., Bl. 8
1113 Sofia, Bulgaria

Anton Iliev
Faculty of Mathematics and Informatics
University of Plovdiv Paisii Hilendarski
24, Tzar Asen Str.
4000 Plovdiv, Bulgaria
e-mail: `aii@uni-plovdiv.bg`