International Journal of Differential Equations and Applications
Volume **19**, No. 1 (2020), pages: 1-14
ISSN (Print): 1311-2872; ISSN (Online): 1314-6084;
url: https://www.ijdea.eu
ijdea.eu

PROPERTIES OF A POWER TOPP–LEONE G–FAMILY WITH BASELINE GOMPERTZ CUMULATIVE DISTRIBUTION FUNCTION

Nikolay Kyurkchiev[1], Anton Iliev[2],
Asen Rahnev[3], Todorka Terzieva[4]

[1,2,3,4]Faculty of Mathematics and Informatics
University of Plovdiv Paisii Hilendarski
24, Tzar Asen Str., 4000 Plovdiv, BULGARIA

ABSTRACT: In [2] the authors introduced a new power Topp–Leone–G–Family (NTL–G) of distribution with c.d.f

$$F(t) = e^{\alpha\beta\left(1-\frac{1}{G(t)}\right)}\left(2 - e^{\beta\left(1-\frac{1}{G(t)}\right)}\right)^{\alpha}$$

where $\alpha, \beta \in R^+$ and $G(t)$ is a c.d.f. of a baseline continuous distribution.

Interesting particular case of offered new family of cdf with "correction of inverse Lomax-type cdf" is also considered.

During last 5 years are appeared in the literature modifications of classical and newer probability distributions and their generalized G-families obligatory are researched in the sense of other important characteristics (beside above mentioned "confidence bounds") - "supersaturation" of the cdf of these distributions to the horizontal asymptote about I-III quartile. This task is connected to approximation of shifted Heaviside function $h_{t_0}(t)$ by the fixed cdf about Hausdorff distance [12] where t_0 is the "median".

In this paper we study some properties of the new Topp–Leone–G–Family with baseline Gompertz–type cdf (NTLG–G).

The estimates of the value of the one–sided Hausdorff approximation obtained in this article can be used in practice as one possible additional criterion in "saturation" study.

We give example with real dataset: "dataModBlaster" worm, "data $\_Storm$", "data $\_Conficker$", "data$_Journal$", "data Level to literacy in Bulgaria – men (1887–1946)" and "data$_Witty_World$".

Numerical examples are presented using *CAS MATHEMATICA*.

AMS Subject Classification: 41A46

Key Words: new power Topp–Leone–G–Family (NTL–G) of distribution, "correction of inverse Lomax-type cdf", new Topp–Leone–G–Family with baseline Gompertz–type cdf (NTLG–G), Heaviside function, Hausdorff distance

Received: September 11, 2019 **Revised:** December 20, 2019
Published: February 21, 2020 **doi:** 10.12732/ijdea.v19i1.1
Academic Publications, Ltd. https://acadpubl.eu

1. INTRODUCTION

In [2] Bantan, Jamal, Chesneau and Elgarhy introduced a new power Topp–Leone–G–Family (NTL–G) of distribution.

The corresponding cumulative distribution function is written as [2]:

$$F(t) = e^{\alpha\beta\left(1-\frac{1}{G(t)}\right)}\left(2-e^{\beta\left(1-\frac{1}{G(t)}\right)}\right)^{\alpha} \tag{1}$$

where $\alpha, \beta \in R^+$ and $G(t)$ is a c.d.f. of a baseline continuous distribution.

The following result shows some inequalities involving $F(t)$ (see, Proposition 1 [2]):

$$e^{\alpha\beta\left(1-\frac{1}{G(t)}\right)}\left(2-G(t)^{\beta}\right)^{\alpha} \le F(t) \le 2^{\alpha} e^{\alpha\beta\left(1-\frac{1}{G(t)}\right)}. \tag{2}$$

In this paper we study some properties of the new Topp–Leone–G–Family with baseline Gompertz–type cdf (NTLG–G); $G(t) = 1 - e^{-a\left(e^{bt}-1\right)}$.

The corresponding cumulative distribution function is written as:

$$M(t) = e^{\alpha\beta\left(1-\frac{1}{1-e^{-a\left(e^{bt}-1\right)}}\right)}\left(2-e^{\beta\left(1-\frac{1}{1-e^{-a\left(e^{bt}-1\right)}}\right)}\right)^{\alpha} \tag{3}$$

where $\alpha, \beta, a, b \in R^+$.

For other results, see [1]–[11].

2. MAIN RESULTS

In this Section we study the Hausdorff approximation [12] of the Heaviside step function $h_{t_0}(t)$ where t_0 is the "median" by families of the new Topp–Leone–G–Family with

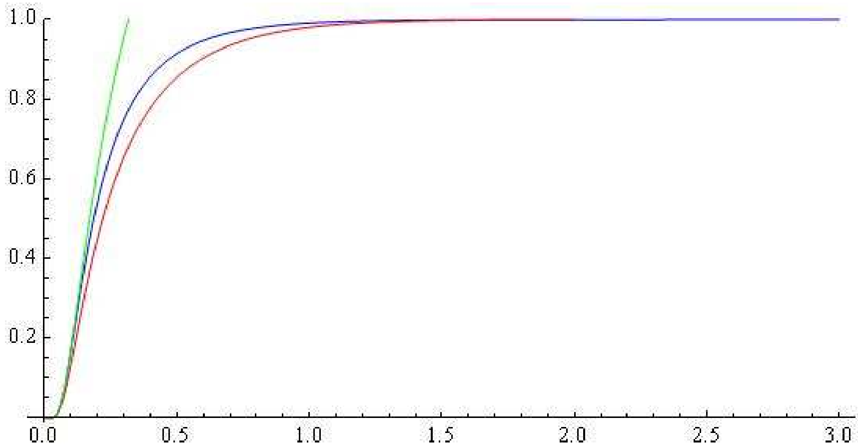

Figure 1: The two-sides estimations (4) for $\theta = 1.1$; $\alpha = 0.9$; $\beta = 0.3$; $a = 2.1$; $b = 0.5$.

baseline Gompertz–type cdf (NTLG–G)

$$h_{t_0}(t) = \begin{cases} 0, & \text{if} \quad t < t_0, \\ [0,1], & \text{if} \quad t = t_0, \\ 1, & \text{if} \quad t > t_0. \end{cases}$$

The obtained two-sides estimations (see Proposition 1. [2]) in particular case with usage of the baseline Gompertz–type cdf for $\theta = 1.1$; $\alpha = 0.9$; $\beta = 0.3$; $a = 2.1$; $b = 0.5$

$$e^{\alpha\beta\left(1-\frac{1}{1-e^{-a\left(e^{bt}-1\right)}}\right)}\left(2-\left(1-e^{-a\left(e^{bt}-1\right)}\right)^{\beta}\right)^{\alpha} \leq M(t) \leq 2^{\alpha} e^{\alpha\beta\left(1-\frac{1}{1-e^{-a\left(e^{bt}-1\right)}}\right)} \tag{4}$$

are given in Fig. 1.

From Fig.1 it can be seen that these estimations can be used as "confidence bounds", which are extremely useful for the specialists in the choice of model for cumulative data approximating in areas of Biostatistics, Population dynamics, Growth theory, Debugging and Test theory, Computer viruses propagation, Financial and Insurance mathematics.

Let t_0 is the value for which $M(t_0) = \frac{1}{2}$.

The Hausdorff distance d between the function $h_{t_0}(t)$ and $M(t)$ satisfies the relation

$$M(t_0 + d) = 1 - d. \tag{5}$$

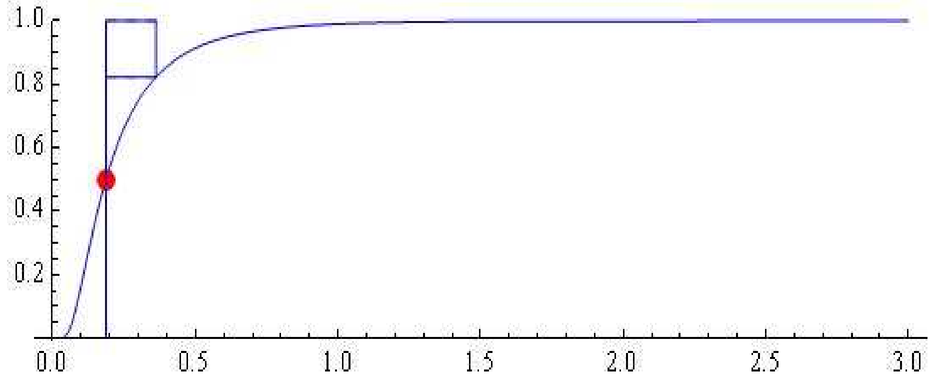

Figure 2: The model $M(t)$ for $\theta = 1.1$; $\alpha = 0.9$; $\beta = 0.3$; $a = 2.1$; $b = 0.5$, $t_0 = 0.1861174$; H–distance $d = 0.176102$.

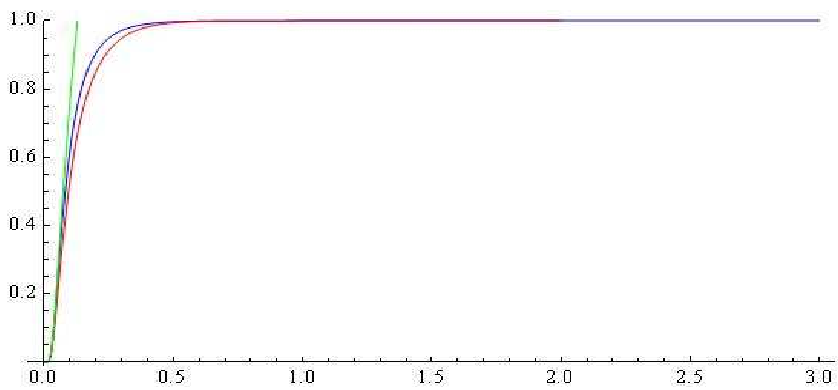

Figure 3: The two-sides estimations (4) for $\theta = 1.2$; $\alpha = 0.99$; $\beta = 0.4$; $a = 3.5$; $b = 0.95$.

For fixed $\theta = 1.1$; $\alpha = 0.9$; $\beta = 0.3$; $a = 2.1$; $b = 0.5$ we find $t_0 = 0.1861174$ and from the nonlinear equation (5) we have $d = 0.176102$ (see, Fig. 2).

The two-sides estimations (4) for $\theta = 1.2$; $\alpha = 0.99$; $\beta = 0.4$; $a = 3.5$; $b = 0.95$ are visualized on Fig. 3

For fixed $\theta = 1.2$; $\alpha = 0.99$; $\beta = 0.4$; $a = 3.5$; $b = 0.95$ we find $t_0 = 0.0810943$ and from the nonlinear equation (5) we have $d = 0.108504$ (see, Fig. 4).

From the graphics it can be seen that the ”saturation” is faster.

During last 5 years are appeared in the literature modifications of classical and newer probability distributions and their generalized G-families obligatory are researched in the sense of other important characteristics (beside above mentioned ”confidence

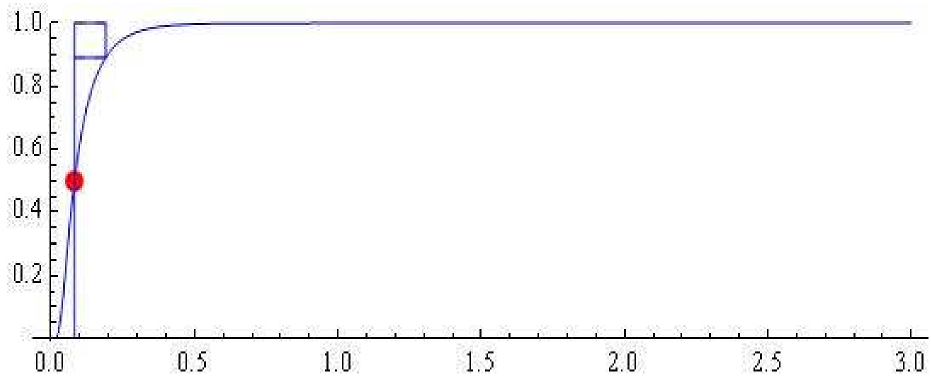

Figure 4: The model $M(t)$ for $\theta = 1.2$; $\alpha = 0.99$; $\beta = 0.4$; $a = 3.5$; $b = 0.95$, $t_0 = 0.0810943$; H–distance $d = 0.108504$.

bounds") - "supersaturation" of the cdf of these distributions to the horizontal asymptote about I-III quartile.

This task is connected to approximation of shifted Heaviside function $h_{t_0}(t)$ by the fixed cdf about Hausdorff distance [12] where t_0 is the "median".

Obviously, imposed combined research "confidence bounds" and "supersaturation" gives the opportunity to the researcher for choice of appropriate model when approximating cumulative specific data of pointed above mentioned domains.

3. APPLICATIONS

Example 1. We consider the following "dataModBlaster" worm:

$$dataModBlaster$$

$$:= \{\{1, 10\}, \{100, 410\}, \{200, 4103\}, \{300, 5517\}, \{400, 95345\}, \{500, 472414\},$$

$$\{600, 565517\}, \{700, 581034\}, \{800, 590345\}\};$$

The model $M^*(t) = \omega M(t)$ for

$$\alpha = 4722.33146377, \beta = 0.29842407, a = 3.71, b = 0.0014, \omega = 590345$$

is visualized on Fig. 5.

Example 2. Storm worm one of the most biggest cyber threats of 2008.

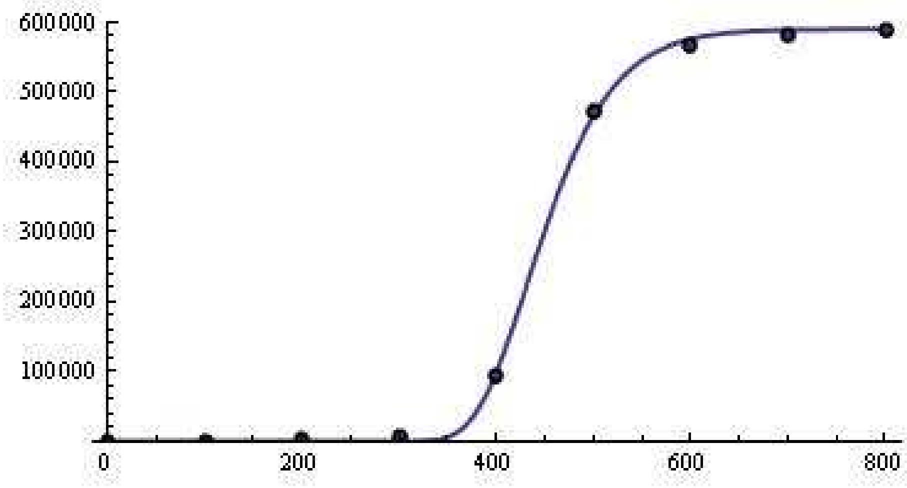

Figure 5: The model M^*.

We analyze the following data [27]

$$data\_Storm\_IDs := \{\{1, 0.843\}, \{4, 0.926\}, \{5, 0.954\}, \{6, 0.967\},$$
$$\{7, 0.976\}, \{8, 0.981\}, \{9, 0.985\}, \{10, 0.991\}, \{22, 0.995\},$$
$$\{38, 0.997\}, \{51, 0.998\}, \{64, 0.9985\}, \{74, 0.999\}, \{83, 1\}, \{100, 1\},$$
$$\{367, 1\}\}$$

The cdf $M(t)$ for $\alpha = 1.2458162728$, $\beta = 0.0035766929$, $a = 2.71$, $b = 0.0021$ is visualized on Fig. 6.

Example 3. Here we will present a new analysis of Conficker propagation in 2008 and we explore the Network Telescope project's daily dataset [28], [29] collected on November 21, 2008.

We analyze the following data

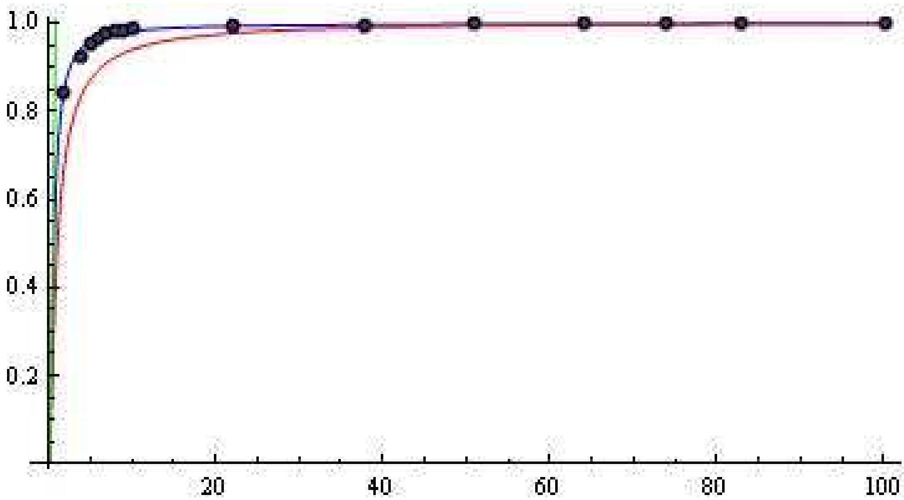

Figure 6: The fitted model $M(t)$ with "confidence bounds".

$$data\_Conficker :=$$

$$\{\{0.1, 10\}, \{1, 150\}, \{2, 300\}, \{3, 600\}, \{4, 2500\}, \{5, 5000\},$$

$$\{6, 7500\}, \{7, 13000\}, \{8, 19000\}, \{9, 25000\}, \{10, 31000\},$$

$$\{11, 37000\}, \{12, 44000\}, \{13, 52000\}, \{14, 58000\}, \{15, 66000\},$$

$$\{16, 74000\}, \{17, 81000\}, \{18, 86000\}, \{19, 89000\}, \{20, 92000\},$$

$$\{21, 92500\}\}$$

The $M^*(t) = \omega M(t)$ for $\omega = 92500$; $\alpha = 0.27398368$; $a = 11.75$; $b = 0.0041$ and $\beta = 4.09446898$ is visualized on Fig. 7.

Example 4. Analysis of data "growth of the cumulative number of TREZ publications" [30], [31]

$$data\_Journal$$

$$:= \{\{1.1, 5\}, \{2, 37\}, \{3, 107\}, \{4, 201\}, \{5, 298\}, \{6, 439\},$$

$$\{7, 617\}, \{8, 773\}, \{9, 936\}, \{10, 1121\}, \{11, 1316\},$$

$$\{12, 1451\}, \{13, 1563\}, \{14, 1629\}, \{15, 1722\}, \{16, 1788\}\};$$

After that using the model $M^*(t) = \omega M(t)$ for $a = 3.48$, $b = 0.0372$, $\alpha = 51.4456$,

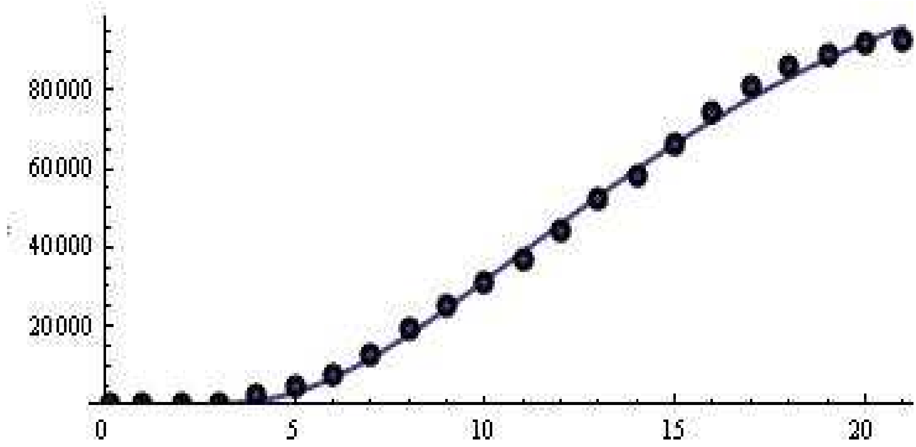

Figure 7: The fitted model $M^*(t)$.

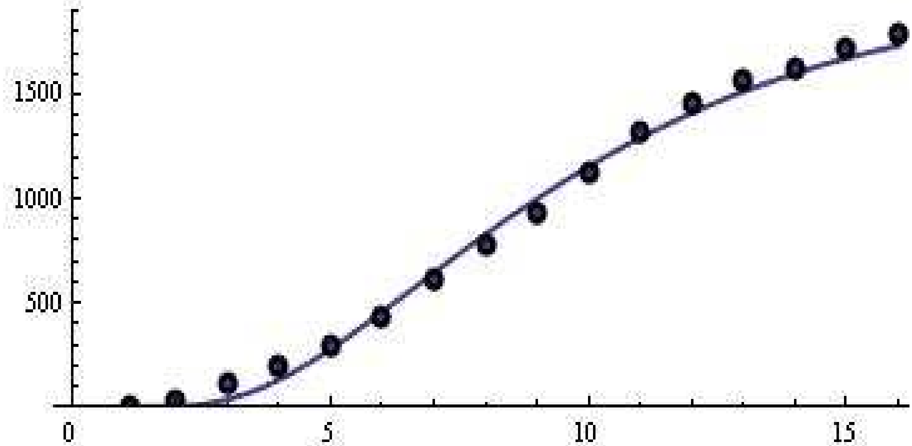

Figure 8: The fitted models $M^*(t)$.

$\beta = 0.040226$ and $\omega = 1970$ we obtain the fitted model (see, Fig. 8).

Example 5. Analysis of Witty worm infection behavior [33].

Here we will give an application of the model $M(t)$ when provide analysis of this real "data" [33].

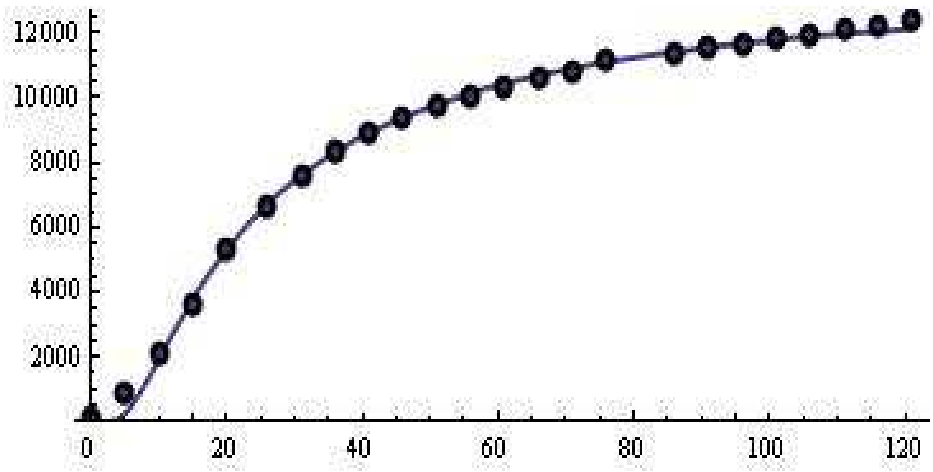

Figure 9: The fitted model $M^*(t)$

$$data\_Witty\_World =$$

$$\{\{0.1, 150\}, \{5, 869\}, \{10, 2141\}, \{15, 3637\}, \{20, 5312\},$$
$$\{26, 6602\}, \{31, 7562\}, \{36, 8340\}, \{41, 8941\}, \{46, 9389\}, \{51, 9734\},$$
$$\{56, 10060\}, \{61, 10349\}, \{66, 10586\}, \{71, 10800\}, \{76, 11169\},$$
$$\{86, 11362\}, \{91, 11532\}, \{96, 11684\}, \{101, 11823\}, \{106, 11972\},$$
$$\{111, 12118\}, \{116, 12256\}, \{121, 12372\}\}$$

For entire World spreading parameters are (see Fig. 9)

$$\alpha = 68.36217257; a = 0.7; b = 0.009; \beta = 0.00190972; \omega = 12650.$$

Example 6. Approximation of the data: "Growth of the number of ranchers adopting improved pasture technology in Uruguay (1961–1976)" [34]:

$$data\_URUGUAY :=$$

$$\{\{0.1, 141\}, \{1, 261\}, \{2, 397\}, \{3, 697\}, \{4, 944\}, \{5, 1455\},$$

$$\{6, 12060\}, \{7, 13247\}, \{8, 15284\}, \{9, 17999\}, \{10, 19554\},$$

$$\{11, 21465\}, \{12, 23767\}, \{13, 24678\}, \{14, 24998\}, \{15, 25473\}\}.$$

After that using the model $M^*(t) = \omega M(t)$ for $a = 38$, $b = 0.033$, $\alpha = 0.79019730$, $\beta = 2.41446995$ and $\omega = 25473$ we obtain the fitted model (see, Fig. 10).

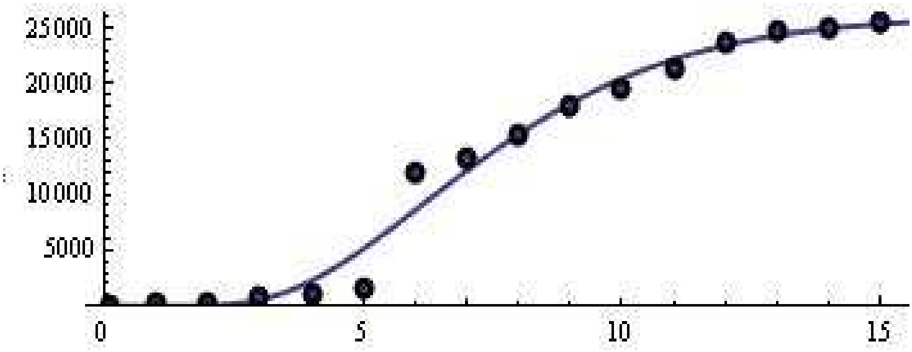

Figure 10: The fitted model $M^*(t)$ for approximation of the data: "Growth of the number of ranchers adopting improved pasture technology in Uruguay (1961–1976) [34].

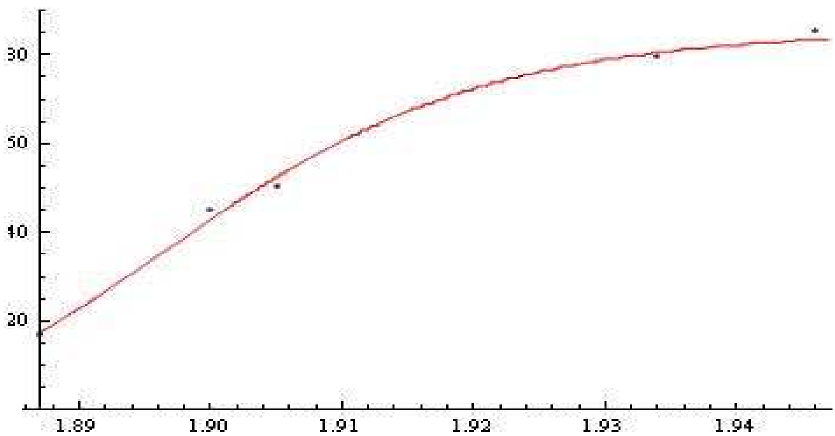

Figure 11: The fitted model $M^*(t)$ for approximation of the data: "Level to literacy in Bulgaria – men (1887–1946)" [35].

Example 7. Approximation of the data: "Level to literacy in Bulgaria – men (1887–1946)" [35]:

$$data\_literacy\_BULGARIA\_men\_(1887-1946) :=$$

$$\{\{1887, 17\}, \{1900, 44.96\}, \{1905, 50.57\}, \{1934, 79.60\}, \{1946, 85.51\}\}.$$

After that using the model $M^*(t) = \omega M(t)$ for $a = 2.18$, $b = 1.44$, $\alpha = 1.15511$, $\beta = 4.65 \times 10^{13}$ and $\omega = 85.51$ we obtain the fitted model (see, Fig. 11).

In addition, the experiments must be performed with very high accuracy.

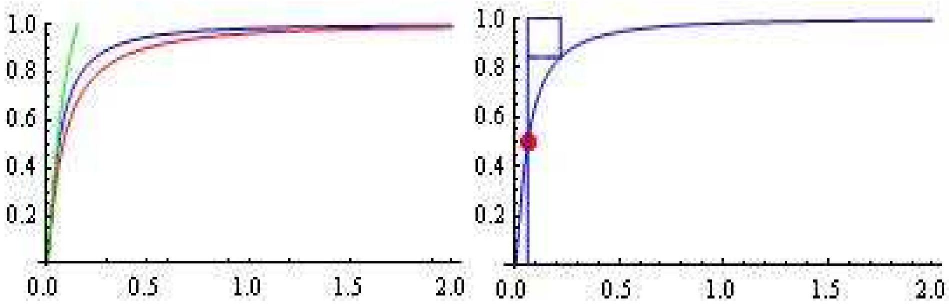

Figure 12: The combined research "confidence bounds" and "supersaturation" for the (NTLIL–G) family.

A excellent survey about monotonic decrease of upper limit estimated with Gompertz model for data described using logistic model can be found in Satoh and Matsumura [32].

Obviously, the model $M(t)$ can also be used for approximating of some "specific data".

4. CONCLUSIONS

Interesting particular case of offered new family of cdf with "correction of inverse Lomax-type cdf" is considered in [2], i.e. $G(t) = \left(1 + \frac{1}{t}\right)^{-\theta}$.

Research on this interesting family can be extended, in the light of the proposed methodology, which is the subject of this article.

For example, for $\theta = 0.6$, $\alpha = 0.9$, $\beta = 0.3$, $t_0 = 0.0640215$ for the one–sided Hausdorff distance we have $d = 0.156789$, see Fig. 12.

The results obtained in this paper can be used when controlling growth in areas of Biostatistics, Population dynamics, Debugging and Test theory, Computer viruses propagation, Financial and Insurance mathematics.

For some approximation, computational and modelling aspects, see [13]–[26].

ACKNOWLEDGMENTS

This paper is supported by the National Scientific Program "Information and Communication Technologies for a Single Digital Market in Science, Education and Security (ICTinSES)", financed by the Ministry of Education and Science.

REFERENCES

[1] G. W. Topp, F. C. Leone, A family of J–shaped frequency functions, *Journal of American Stat. Association*, **50** (1955), 209–219.

[2] Rashad A.R. Bantan, Farrukh Jamal, Christophe Chesneau, Mohammed Elgarhy, A New Power Topp-Leone Generated Family of Distributions with Applications, *Entropy* (2019).

[3] M. Elgarhy, M. Nazir, F. Jamal, G. Ozel, The type II Topp–Leone generated family of distributions: properties and applications, *J. of Statistics and Management Systems*, **21**, No. 8 (2018), 1525–155.

[4] M. Tahir, G. Cordeiro, Compounding of distributions: A survey and new generalized classes, *J. of Stat. Distr. and Appl.,* **3**, No. 13 (2016), 1–35.

[5] A. Hassan, M. Elgarhy, Z. Ahmad, Type II generalized Topp–Leone family of distributions: properties and applications, *J. of Data Sci.,* **17**, No 4 (2019), 638–659.

[6] A. Mahdavi, Generalized Topp–Leone family of distributions, *J. of Biost. and Epidemiology,* **3**, No 2 (2017), 65–75.

[7] H. Reyad, M. Alizadeh, F. Jamal, S. Othman, G. Hamedani, The exponentiated generalized Topp Leone–G family of distributions: properties and applications, *Pak. J. of Stat. and Oper. Res.,* **15**, No 1 (2019), 1–24.

[8] H. Reyad, M. Korkmaz, A. Afify, G. Hamedani, S. Othman, The Frechet Topp Leone–G family of distributions: properties, characterizations and applications, *Ann. of Data Sci.,* (2019).

[9] S. Rezaei, B. Sadr, M. Alizadeh, S. Nadarajah, Topp–Leone generated family of distributions: properties and applications, *Comm. in Stat.–Theory and Methods,* **46**, No 6 (2016), 2893–2909.

[10] H. Yousof, M. Alizadeh, S. Jahanshalu, T. Ramires, I. Ghosh, G. Hamedani, The transmuted Topp–Leone G family of distributions: theory, characterizations and applications, *J. of Data Sci.,* **15** (2017), 723–740.

[11] A. Iliev, A. Rahnev, N. Kyurkchiev, S. Markov, A Study on the Unit-logistic, Unit-Weibull and Topp-Leone Cumulative Sigmoids, *Biomath Communications*, **6**, No 1 (2019), 1–15.

[12] B. Sendov, *Hausdorff Approximations*, Kluwer, Boston (1990).

[13] N. Kyurkchiev, S. Markov, On the Hausdorff distance between the Heaviside step function and Verhulst logistic function, *J. Math. Chem.*, **54** No. 1 (2016), 109–119.

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

[15] N. Kyurkchiev, A. Iliev, S. Markov, *Some Techniques for Recurrence Generating of Activation Functions: Some Modeling and Approximation Aspects*, LAP LAMBERT Academic Publishing (2017), ISBN: 978-3-330-33143-3.

[16] S. Markov, A. Iliev, A. Rahnev, N. Kyurkchiev, A note on the Log-logistic and transmuted Log-logistic models. Some applications, *Dynamic Systems and Applications*, **27** No. 3 (2018), 593–607.

[17] S. Markov, N. Kyurkchiev, A. Iliev, A. Rahnev, On the approximation of the cut functions by hyper-log-logistic function, *Neural, Parallel and Scientific Computations*, **26** No. 2 (2018), 169–182.

[18] 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.

[19] N. Kyurkchiev, A. Iliev, A. Rahnev, *Some Families of Sigmoid Functions: Applications to Growth Theory*, LAP LAMBERT Academic Publishing, (2019), ISBN: 978-613-9-45608-6.

[20] A. Iliev, N. Kyurkchiev, S. Markov, On the Approximation of the step function by some sigmoid functions, *Mathematics and Computers in Simulation*, **133** (2017), 223–234.

[21] S. Markov, N. Kyurkchiev, A. Iliev, A. Rahnev, On the approximation of the generalized cut functions of degree $p+1$ by smooth hyper-log-logistic function, *Dynamic Systems and Applications*, **27** No. 4 (2018), 715–728.

[22] N. Pavlov, 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.

[23] 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.

[24] A. Iliev, N. Kyurkchiev, A. Rahnev, T. Terzieva, *Some models in the theory of computer viruses propagation*, LAP LAMBERT Academic Publishing (2019), ISBN:

978-620-0-00826-8.

[25] N. Kyurkchiev, Uniform Approximation of the Generalized Cut Function by Erlang Cumulative Distribution Function. Application in Applied Insurance Mathematics, *International Journal of Theoretical and Applied Mathematics*, **2**, No. 2 (2016), 40-44.

[26] N. Kyurkchiev, Mathematical Concepts in Insurance and Reinsurance: Some Moduli in Programming Environment MATHEMATICA, LAP LAMBERT Academic Publishing, 2016, ISBN:978-3-659-96906-5.

[27] S. Sarat, A. Terzis, HiNRG Technical Report: 01-10-2007 Measuring the Storm Worm Network, (2007).

[28] C. Zhang, S. Zhou, B. Chain, Hybrid Epidemics–A Case Study on Computer Worm Conficker, *PLoS ONE*, **10**, No. 5 (2015), e0127478.

[29] P. Porras, H. Saidi, V. Yegneswaran, An Analysis of Conficker's Logic and Rendezvous Points, SRI international technical report, March 19, (2009).

[30] D. Kucharavy, R. Guio, Application of logistic growth curve, *Proc. Eng.,* **131** (2015), 280–290.

[31] D. Satoh, Property of logistic data exposed with Gompertz model and resistance to noise in actual data, *Japan J. of Ind. and Appl. Math.,* (2019).

[32] D. Satoh, R. Matsumura, Monotonic decrease of upper limit estimated with Gompertz model for data described using logistic model, *Japan J. of Ind. and Appl. Math.,* (2018).

[33] C. Shannon, D. Moore, The Spread of the Witty Worm, *IEEE Security & Privacy*, **July/August**, (2004), 46–50.

[34] L. Jarvis, Predicting the diffusion of improved pastures in Uruguay, *American Journal of Agricultural Economics*, **63** (1981), 495–502.

[35] Bulgarska kniga. Enziklopediya. Sustavitel: A. Gergova, Pensoft, 2004, ISBN 954-642-210-XC.