Neural, Parallel, and Scientific Computations, **28**, No. 1 (2020), 13-25 ISSN: 1061-5369

PROPERTIES OF SOME TRUNCATED FAMILIES OF CUMULATIVE DISTRIBUTION FUNCTION

TODORKA TERZIEVA[1], NIKOLAY PAVLOV[2],
ANNA MALINOVA[3], EVGENIYA ANGELOVA[4]
[1,2,3,4]Faculty of Mathematics and Informatics
University of Plovdiv Paisii Hilendarski
24, Tzar Asen Str., 4000 Plovdiv, BULGARIA

ABSTRACT: In [1] the authors introduced a truncated inverted Kumuraswamy –G family (TIK–G) of distribution with c.d.f.

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

where $a, b \in R^{+}$ 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 Exponential–type cdf" (TIK–Ex), i.e. $G(t) = 1 - e^{-\theta t}$ is also considered

$$F_1(t) = \frac{1}{\left(1-2^{-a}\right)^{b}} \left(1-\left(2-e^{-\theta t}\right)^{-a}\right)^{b}$$

where $\theta > 0$.

In [2] the authors considered the following doubly truncated inverse Lomax (DTIL) cdf

$$F_2(t) = \frac{\left(1+\frac{\beta}{t}\right)^{-\alpha} - \left(1+\frac{\beta}{a}\right)^{-\alpha}}{\left(1+\frac{\beta}{b}\right)^{-\alpha} - \left(1+\frac{\beta}{a}\right)^{-\alpha}}$$

where $\alpha, \beta > 0$ and $a \le t \le b$.

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 "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 [3] where t_0 is the "median".

In this paper we study some properties of the new (TIK–G) and (DTIL) families of c.d.f.

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: truncated inverted Kumuraswamy –G family (TIK–G), doubly truncated inverse Lomax (DTIL) cdf, Heaviside function, Hausdorff distance

Received: January 23, 2020 ; **Accepted:** April 12, 2020 ;
Published: May 28, 2020. **doi:** 10.12732/npsc.v28i1.2
Dynamic Publishers, Inc., Acad. Publishers, Ltd. https://acadsol.eu/npsc

1. INTRODUCTION

In [1] Bantan, Jamal, Chesneau and Elgarhy introduced a truncated inverted Kumuraswamy –G family (TIK–G) of distribution with "correction of Exponential–type cdf" (TIK–Ex)

$$F_1(t) = \frac{1}{(1-2^{-a})^b}\left(1-\left(2-e^{-\theta t}\right)^{-a}\right)^b \tag{1}$$

where $a, b, theta \in R^+$.

In [2] Yadav, Shukla and Kumari considered the following doubly truncated inverse Lomax (DTIL) cdf

$$F_2(t) = \frac{\left(1+\frac{\beta}{t}\right)^{-\alpha} - \left(1+\frac{\beta}{a}\right)^{-\alpha}}{\left(1+\frac{\beta}{b}\right)^{-\alpha} - \left(1+\frac{\beta}{a}\right)^{-\alpha}} \tag{2}$$

where $\alpha, \beta > 0$ and $a \leq t \leq b$.

In this paper we study some properties of the families (1) and (2).

For other results, see [9]–[29].

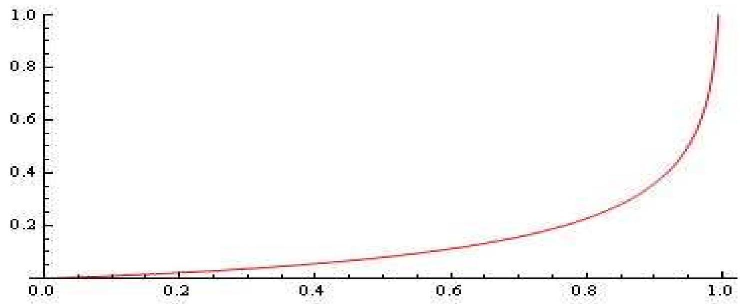

Figure 1: The function $Q(u)$ for $a = 0.99$, $b = 0.9$, $\theta = 4.5$.

2. MAIN RESULTS

2.1. "SATURATION" BY THE MODEL (1)

In this Section we study the Hausdorff approximation [3] of the Heaviside step function $h_{t_0}(t)$

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

where t_0 is the "median" by families of the type (1)

The investigation of the characteristic "supersaturation" of the cdf (1) to the horizontal asymptote is important.

The quantile function is defined by [1]:

$$Q(u) = -\frac{1}{\theta} \ln \left(2 - \left(1 - u^{\frac{1}{b}} \left(1 - 2^{-a} \right) \right)^{\frac{-1}{a}} \right) \tag{3}$$

(see, Fig. 1 for $a = 0.99$, $b = 0.9$, $\theta = 4.5$).

The median is obtained by substituting $u = 0.5$ in (3).

Let t_0 is the value for which $F_1(t_0) = \frac{1}{2}$, i.e. $t_0 = Q(0.5)$.

The Hausdorff distance d between the function $h_{t_0}(t)$ and $F_1(t)$ satisfies the relation

$$F_1(t_0 + d) = 1 - d. \tag{4}$$

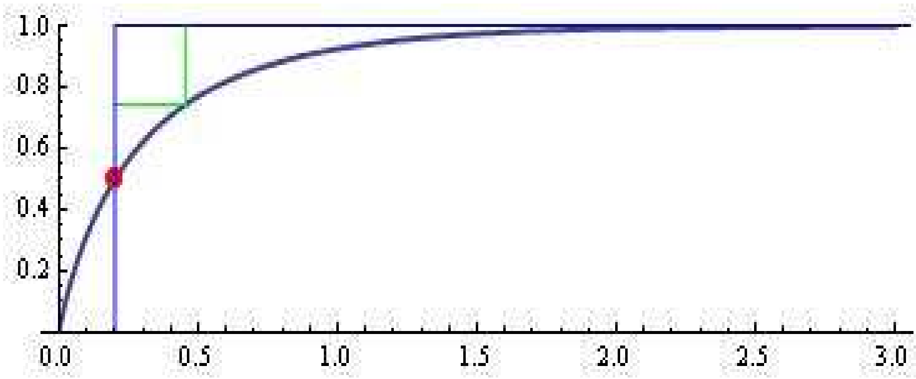

Figure 2: The model $F_1(t)$ for $a = 0.6$; $b = 0.9$; $\theta = 2$, $t_0 = 0.200718$; H–distance $d = 0.255252$.

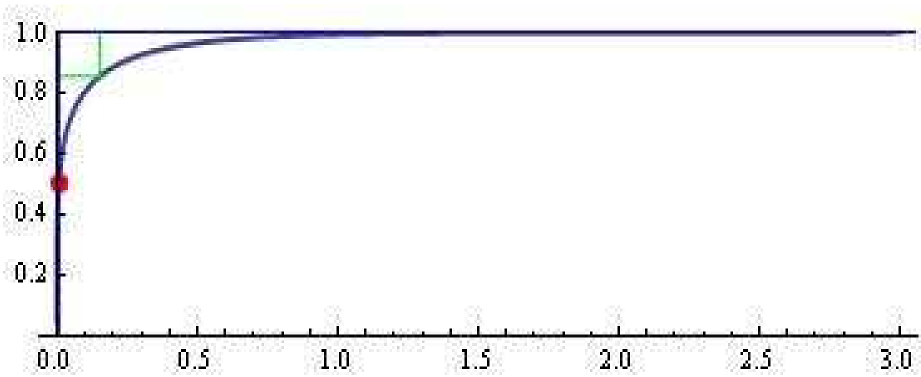

Figure 3: The model $F_1(t)$ for $a = 0.1$; $b = 0.2$; $\theta = 3$, $t_0 = 0.00713246$; H–distance $d = 0.144553$.

For fixed $a = 0.6$; $b = 0.9$; $\theta = 2$ we find $t_0 = 0.200718$ and from the nonlinear equation (4) we have $d = 0.255252$ (see, Fig. 2).

For fixed $a = 0.1$; $b = 0.2$; $\theta = 3$ we find $t_0 = 0.00713246$ and from the nonlinear equation (4) we have $d = 0.144553$ (see, Fig. 3).

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

We note that, for some conditions the following is valid

Let

$$p = -\frac{1}{2},$$

$$q = 1 + \frac{ab\theta\left(0.5^{\frac{1}{b}}(1-2^{-a})\right)^{b-1}\left(1-0.5^{\frac{1}{b}}(1-2^{-a})\right)^{\frac{1+a}{a}}\left(2-\left(1-0.5^{\frac{1}{b}}(1-2^{-a})\right)^{-\frac{1}{a}}\right)}{(1-2^{-a})^{b}}.$$

For the one–sided Hausdorff distance d between $h_{t_0}(t)$ and the function (1) the following inequalities hold for:

$$2.1q > e^{1.05}$$

$$d_l = \frac{1}{2.1q} < d < \frac{\ln(2.1q)}{2.1q} = d_r. \tag{5}$$

Evidently, function $G(d) = p + qd$ approximates $F(d) = F_1(t_0 + d) - 1 + d$ with $d \to 0$ as $O(d^2)$ (see, Fig. 4).

The proof follows the ideas given in [4] and will be omitted.

For example, for $a = 0.6$; $b = 0.9$; $\theta = 2$ we have

$$d_l = 0.194133 < d = 0.255252 < d_r = 0.318225.$$

2.2. APPLICATION

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

We analyze the following data [5]

$$data\_Storm\_IDs := \{\{1, 0.843\}, \{4, 0.926\}, \{5, 0.954\}, \{6, 0.967\},$$

$$\{7, 0.976\}, \{8, 0.981\}, \{9, 0.985\}, \{10, 0.991\}, \{22, 0.995\},$$

$$\{38, 0.997\}, \{51, 0.998\}, \{64, 0.9985\}, \{74, 0.999\}, \{83, 1\}, \{100, 1\},$$

$$\{367, 1\}\}$$

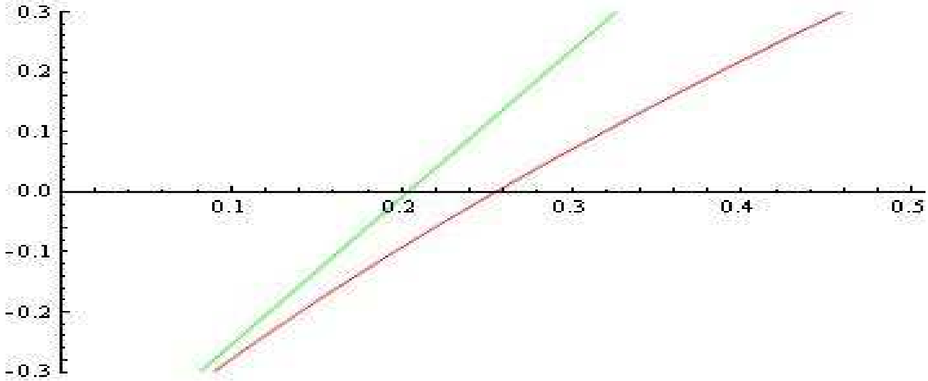

Figure 4: The functions G and F for $a = 0.6$; $b = 0.9$; $\theta = 2$.

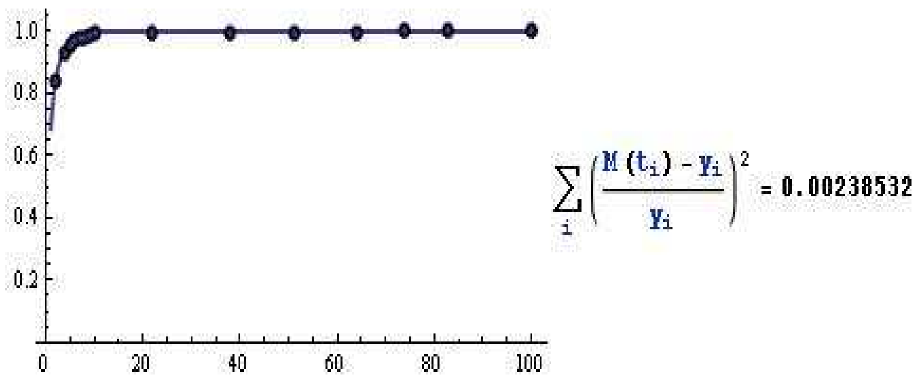

Figure 5: The fitted model $F_1(t)$.

The cdf $F_1(t)$ for $a = 0.001$; $b = 0.5$; $\theta = 0.492$ is visualized on Fig. 5.

For the error we have

$$error = \sum_i \left(\frac{F_1(t_i) - y_i}{y_i} \right)^2 = 0.00238532.$$

2.3. "SATURATION" BY THE MODEL (2)

Let t_0 is the value for which $F_2(t_0) = \frac{1}{2}$.

The Hausdorff distance d between the function $h_{t_0}(t)$ and $F_2(t)$ satisfies the relation

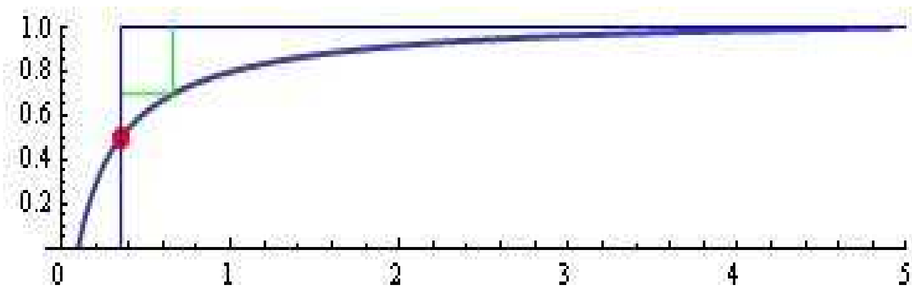

Figure 6: The model $F_2(t)$ for $a = 0.1$; $b = 4.9$; $\alpha = 0.3$; $\beta = 0.4$, $t_0 = 0.353516$; H–distance $d = 0.304494$.

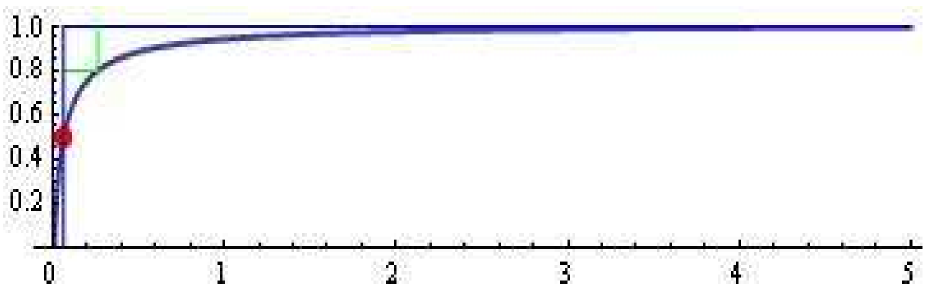

Figure 7: The model $F_2(t)$ for $a = 0.01$; $b = 5$; $\alpha = 0.1$; $\beta = 0.2$, $t_0 = 0.0629688$; H–distance $d = 0.198471$.

$$F_2(t_0 + d) = 1 - d. \tag{6}$$

For fixed $a = 0.1$; $b = 4.9$; $\alpha = 0.3$; $\beta = 0.4$ we find $t_0 = 0.353516$ and from the nonlinear equation (6) we have $d = 0.304494$ (see, Fig. 6).

For fixed $a = 0.01$; $b = 5$; $\alpha = 0.1$; $\beta = 0.2$ we find $t_0 = 0.0629688$ and from the nonlinear equation (6) we have $d = 0.198471$ (see, Fig. 7).

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

2.4. APPLICATION

Example. Analysis of Witty worm infection behavior [6].

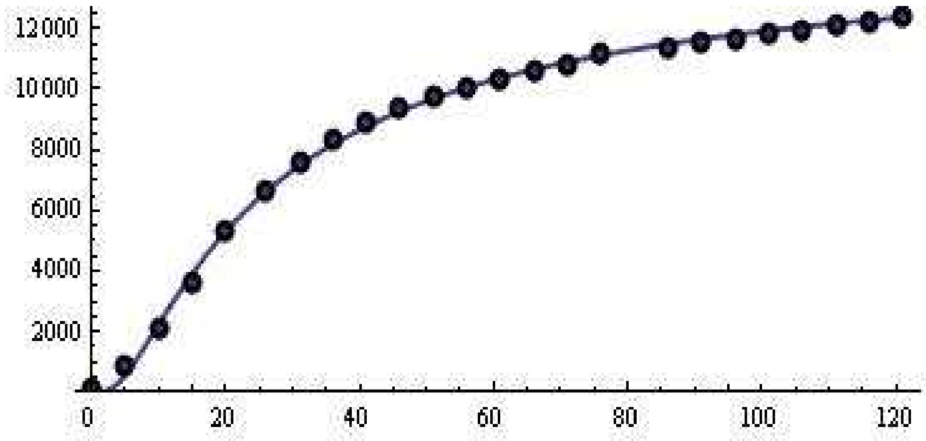

Figure 8: The fitted model $F_2^*(t) = \omega F_2(t)$

Here we will give an application of the model (2) when provide analysis of this real "data" [6].

$$data\_Witty\_World =$$

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

For entire World spreading parameters are (see Fig. 8)

$$a = 0.1; b = 21; \alpha = 65.24324; \beta = 4.34943; \omega = 12372.$$

Example. We will illustrate the advances of the new family for approximation and modelling of "cancer data" (for some details see, [7]–[8]).

days	4	7	10	12	14	17	19	21
$R(t)$	0.415	0.794	1.001	1.102	1.192	1.22	1.241	1.3

Table 1: The "cancer data" [7]–[8]

The model $F_2^*(t) = \omega F_2(t)$ based on the data from Table 1 for the estimated parameters:

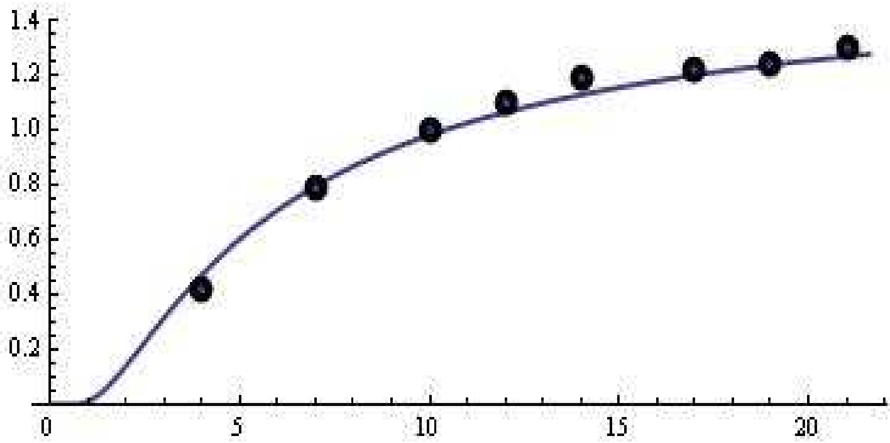

Figure 9: The model $F_2^*(t)$ based on the "cancer data".

$$\omega = 1.3045;\ a = 0.1;\ b = 24;\ \alpha = 552.301;\ \beta = 0.00884363$$

is plotted on Fig. 9.

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 "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 [3] 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.

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

Obviously, the models $F_1(t)$ and $F_2(t)$ can also be used for approximating of some "specific data".

We will note that the parameters a and b can be used as "limiters of specifically located data".

3. CONCLUSIONS

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 [30]–[43].

Acknowledgment

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] R. Bantan, F. Jamal, Ch. Chesneau, M. Elgarhy, Truncated inverted Kumaraswamy generated family of distributions with application, *Entropy* (2019), 22 pp.

[2] A. Yadav, S. Shukla, A. Kumari, Statistical inference for truncated inverse Lomax distribution and its application to survival data, *Ann. of Data Sci.* (2019), 17 pp.

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

[4] 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.

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

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

[7] M. Vinci, S. Gowan, F. Boxall, L. Patterson, M. Zimmermann, W. Court, C. Lomas, M. Mendila, D. Hardisson, S. Eccles, Advances in establishment and analysis of three-dimensional tumor spheroid-based functional assays for target validation and drug evaluation, *BMC Biology*, **10** (2012).

[8] A. Antonov, S. Nenov, T. Tsvetkov, Impulsive controllability of tumor growth, *Dynamic Systems and Appl.*, **28**, No. 1 (2019), 93–109.

[9] Abd Al-Fattah, A.M., El-Helbawy, A.A., Al-Dayian, G.R., Inverted Kumaraswamy distribution: Properties and estimation, *Pak. J. Stat.*, **33** (2017), 37-61.

[10] P. Kumaraswamy, A generalized probability density function for double-bounded random processes, *J. Hydrol.*, **46** (1980), 79-88.

[11] Abid, S.H., Abdulrazak, R.K., [0,1] truncated Frechet-G generator of distributions, *Appl. Math.*, **7** (2017), 51-66.

[12] H. Najarzadegan, Alamatsaz, M.H., Hayati, S., Truncated Weibull-G more flexible and more reliable than geta-G distribution, *Int. J. Stat. Probab.*, **6** (2017), 1-17.

[13] Jamal, F., Bakouch, H.S., Nasir, M.A., A truncated general-G class of distributions with application to truncated Burr-G family, *Rev. Stat.* (2019).

[14] 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.

[15] M. Mandel, Censoring and truncation-highlighting the differences, *Am. Stat.*, **61**, No4 (2007), 321-324.

[16] Aban IB, Meerschaert M., Parameter estimation for the truncated Pareto distribution, *J. Am. Stat. Ass.*, **101**, No 473 (2006), 270-277.

[17] Dey S., Raheem E., Mukherjee S., Statistical properties and different methods of estimation of transmuted Rayleigh distribution, *Rev Colomb Estad.*, **40**, No 1 (2017), 165-203.

[18] Dixit UJ, Phal KD, On estimating scale parameter of a truncated gamma distribution, *Soochow J Math*, **31**, No 4 (2011), 515-523.

[19] Hassan AS, Elgarhy M, Nassr SG, Ahmad Z, Alrajhi S, Truncated Weibull Frechet distribution: statistical inference and applications, *J Comput Theoret Nanosci.*, **16**, No 1 (2019), 52-60.

[20] Kumar D, Yadav AS, Singh SK, Singh U, Chauhan PK, Transmuted inverse Lomax distribution, *J Adv Comput.* (2017).

[21] Mitra D, Balakrishnan N., Statistical inference based on left truncated and interval censored data from log-location-scale family of distribution, *Commun Stat Simul Comput.* (2019).

[22] Nadarajah S, Nassiri V, Mohammadpour A., Truncated-exponential skew-symmetric distributions, *Statistics*, **48**, No 4 (2014), 872-895.

[23] Okasha MK, Alqanoo IM, Inference on the doubly truncated gamma distribution for lifetime data, *Int J Math Stat Invent.*, **2**, No 11 (2014), 1-17.

[24] Singh SK, Singh U, Sharma VK, The truncated Lindley distribution: inference and application, *J Stat Appl Probab.*, **3**, No 2 (2014).

[25] Yadav AS, Singh SK, Singh U., Reliability estimation for inverse Lomax distribution under type-II censored data using Markov chain Monte Carlo method, *Int J Math Stat.*, **17**, No 1 (2016), 128-146.

[26] Yadav AS, Singh SK, Singh U., Bayesian estimation of R = P[Y¡X] for inverse Lomax distribution under progressive type-II censoring scheme, *Int J Syst Assur Eng Manag.*, **10**, No 5 (2019), 905-917.

[27] Zaninetti L., A right and left truncated gamma distribution with application to the stars, *Adv Stud Theoret Phys.,* **23** (2013), 1139-1147.

[28] Zhang T, Xie M., On the upper truncated Weibull distribution and its reliability implications, *Reliab Eng Syst Saf.*, **96**, No 1 (2011), 194-200.

[29] 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.

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

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] N. Kyurkchiev, V. Kyurkchiev, A. Iliev, A. Rahnev, Some nonstandard differential models with applications to the population dynamics and computer viruses propagation, *Dynamic Systems and Applications*, **28**, No. 3 (2019), 757–788.

[41] 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.

[42] 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.

[43] 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.