*Neural, Parallel, and Scientific Computations*, **27**, No. 3&4 (2019), 165-175ISSN: 1061-5369

# INVESTIGATIONS ON THE SATURATION OF C.D.F. OF THE MODIFIED "ODD GENERALIZED EXPONENTIAL POWER FUNCTION DISTRIBUTION". SOME APPLICATIONS

TODORKA TERZIEVA[1], ANTON ILIEV[2], ASEN RAHNEV[3], AND NIKOLAY KYURKCHIEV[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 proposed a new family of continuous distributions called the odd generalized exponential family.

Their reviews have brought to life a number of modifications that have appeared in the literature over the las few years, with applications to approximate specific "cumulative data" from different branches of science.

In this article we will look at the one suggested in [2].

The authors' assertion that probability distribution produces very good results encouraged us to conduct further studies on "saturation" in Hausdorff sense of the corresponding commutative function to the horizontal asymptote hoping to partially contribute to uncovering some of the "intrinsic properties" of this apparently good model.

We also analyze some experimental data: the cumulative number of Welchia attackers; "growth of the cumulative number of TREZ publications"; data set "US mobile"; Software Failure Data – Release #1; Defects (AT & T).

Numerical examples, illustrating our results are presented using programming environment *CAS Mathematica*.

**AMS Subject Classification:** 41A46

**Key Words:** c.d.f. of the ODD–Generalized–Exponential–Power–Function–Distribution (cdfOGEPF), Heaviside step–function $h_{t_0}(t)$, Hausdorff distance

**Received:** June 1, 2019 ; **Accepted:** October 4, 2019 ;
**Published:** October 18, 2019. **doi:** 10.12732/npsc.v27i3&4.3

Dynamic Publishers, Inc., Acad. Publishers, Ltd. https://acadsol.eu/npsc

## 1. INTRODUCTION AND PRELIMINARIES

In this note we study the Hausdorff approximation of the Heaviside function $h_{t_0}(t)$ by the cdf of the ODD–Generalized–Exponential–Power–Function–Distribution (cdfOGEPF), suggested in [2].

The model have been tested with real-world data.

**Definition 1.** The cdf of the ODD–Generalized–Exponential–Power–Function–Distribution (cdfOGEPF) is defined by [2]:

$$M(t) = \left(1 - e^{-\lambda\left(\frac{t^\alpha}{\beta^\alpha - t^\alpha}\right)}\right)^\theta \tag{1}$$

where $0 \leq t < \beta$, $\theta > 0$, $\alpha > 0$.

**Definition 2.** The shifted Heaviside step function is defined by

$$h_{t_0}(t) = \begin{cases} 0, & \text{if} \quad t < t_0, \\ [0,1], & \text{if} \quad t = t_0, \\ 1, & \text{if} \quad t > t_0 \end{cases} \tag{2}$$

**Definition 3.** [3] The Hausdorff distance (the H–distance) $\rho(f,g)$ 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,

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

wherein $||.||$ 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 $||A - B|| = \max(|t_A - t_B|, |x_A - x_B|)$.

## 2. MAIN RESULTS

### 2.1. A NOTE ON THE NEW (CDFOGEPF)

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

The quantile function is defined by:

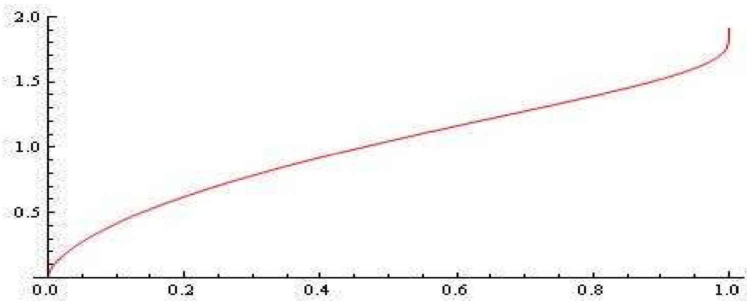


Figure 1: The function $Q(u)$ for $\alpha = 1.9$, $\beta = 2$, $\lambda = 1.5$ and $\theta = 0.9$.

$$Q(u) = \beta \left( -\frac{\ln(1 - u^{\frac{1}{\theta}})}{\lambda - \ln(1 - u^{\frac{1}{\theta}})} \right)^{\frac{1}{\alpha}} \tag{3}$$

(see, Fig. 1 for $\alpha = 1.9$, $\beta = 2$, $\lambda = 1.5$ and $\theta = 0.9$).

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

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

The one–sided Hausdorff distance $d$ between the function $h_{t_0}(t)$ and the cdf (1) satisfies the relation

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

For given $\beta$, $\theta$, $\alpha$, $\lambda$, the nonlinear equation $M(t_0 + d) - 1 + d = 0$ has unique positive root $-$ $d$.

The cdf (1) for $\lambda = 1.5$, $\beta = 2$, $\theta = 0.9$, $\alpha = 1.9$ and $t_0 = 1.04818$ is visualized on Fig. 2.

From the nonlinear equation (4) we have: $d = 0.266933$.

The cdf (1) for $\lambda = 25.6$, $\beta = 1.8$, $\theta = 0.999$, $\alpha = 1.4$ and $t_0 = 0.133999$ is visualized on Fig. 3.

From the nonlinear equation (4) we have: $d = 0.139343$.

From the above examples, it can be seen that the "supersaturation" by the (cdf) $M(t)$ is faster.

Obviously, this "advantage" can actually be used to approximate some specific "cumulative data" from different branches of science.

In the next Section, we will support what is said by analyzing real datasets.

## 2.2. APPLICATIONS

Welchia worm ransomware has a long growing phase in contrast to many other threats.

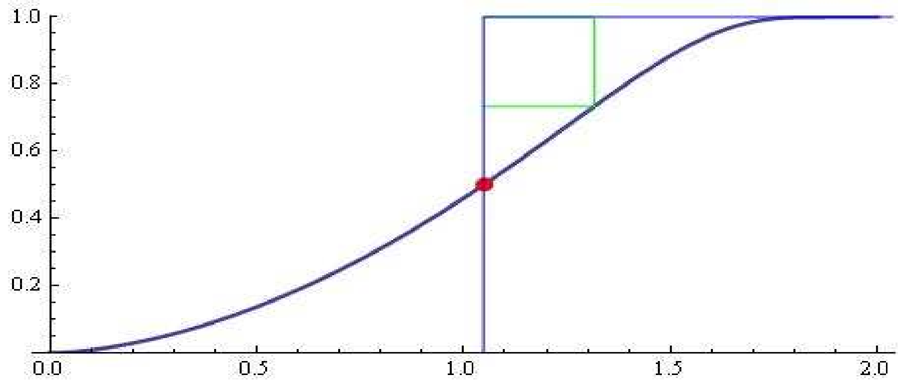


Figure 2: The cdf (1) for $\lambda = 1.5$, $\beta = 2$, $\theta = 0.9$, $\alpha = 1.9$ and $t_0 = 1.04818$; H–distance $d = 0.266933$.

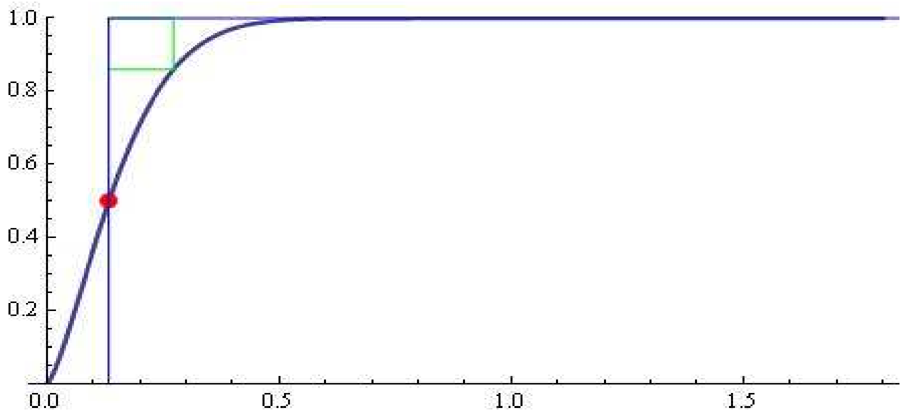


Figure 3: The cdf (1) for $\lambda = 25.6$, $\beta = 1.8$, $\theta = 0.999$, $\alpha = 1.4$ and $t_0 = 0.133999$; H–distance $d = 0.139343$.

Welchia worm uses a vulnerability in the Microsoft remote procedure call service. Welchia firstly checks for Blaster worm and if it is exists continues with Blaster deletion as well as takes care for computer to be immunised for Blaster worm.

**Example 1.** Analysis of Welchia worm infection behavior

For epidemic as Welchia worm it is appropriately to use a model

$$M^*(t) = \omega \left(1 - e^{-\lambda\left(\frac{t^\alpha}{\beta^\alpha - t^\alpha}\right)}\right)^\theta \tag{5}$$

for approximating data from the statistics collected on an individual Welchia [4] honeypot administered by Frederic Perriot between August 24th, 2003 and February 24th,

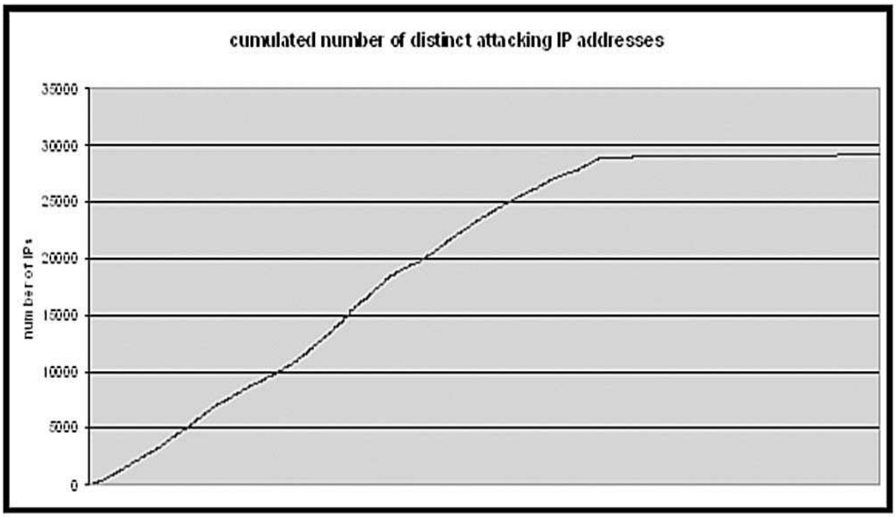


Figure 4: The cumulative number of Welchia attackers [4].

2004, shown in Fig. 4.

We will explore this example by photographing the data from Fig. 4.

$$data\_Welchia :=$$

$$\{\{1.1, 1000\}, \{2, 2333\}, \{3, 3500\}, \{4, 5000\}, \{5, 6833\}, \{6, 8000\},$$
$$\{7, 9333\}, \{8, 10500\}, \{9, 12000\}, \{10, 14000\}, \{11, 16333\},$$
$$\{12, 18167\}, \{13, 19667\}, \{14, 21000\}, \{15, 22667\}, \{16, 23667\},$$
$$\{17, 25000\}, \{18, 26333\}, \{19, 27500\}, \{20, 28333\}, \{21, 29333\},$$
$$\{22, 29500\}, \{23, 29500\}, \{24, 29500\}, \{25, 29500\}, \{26, 29500\},$$
$$\{27, 29500\}, \{28, 29500\}, \{29, 29500\}, \{30, 29500\}, \{31, 29667\},$$
$$\{32, 29667\}\}$$

The fitted model is given by

$$\lambda = 2.51375;\ \ \beta = 33;\ \ \theta = 0.999;\ \ \alpha = 1.31223;\ \ \omega = 30000.$$

We receive an impressive result when approximating these data, see Fig. 5.

Comparison between $M^*(t)$ and the model $M_1^*(t)$ [5] with $\lambda = -0.1$; $\beta = 0.0385232$; $\theta = 0.602142$; $\omega = 30000$ (see, Fig. 6):

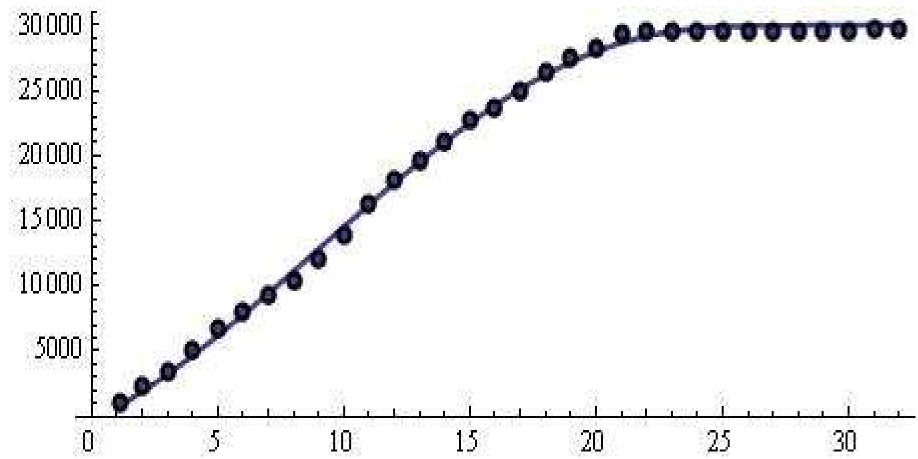


Figure 5: The fitted model $M^*(t)$.

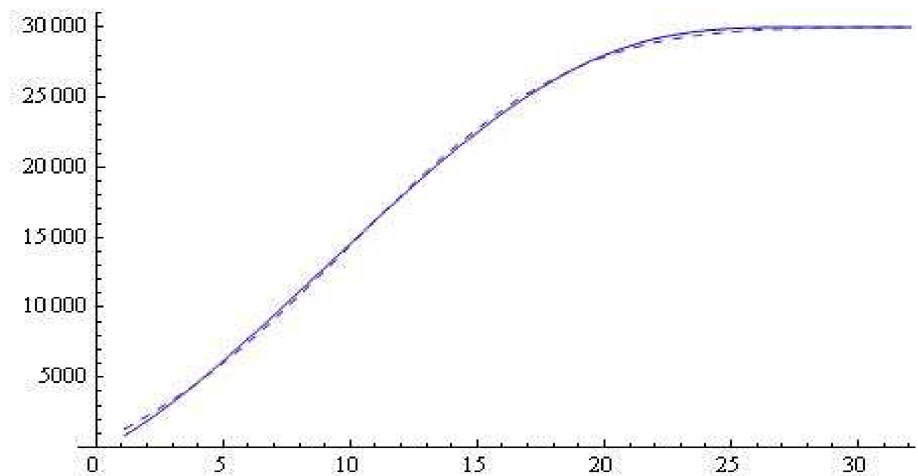


Figure 6: The fitted models $M^*(t)$–blue and $M_1^*(t)$–dashed.

$$M_1^*(t) = \omega \left( 1 - \left( 1 + \frac{e^{\theta((1-\lambda t^{\theta})^{-\frac{1}{\lambda}}-1)} - 1}{e^{\theta} - 1} \right)^{-\beta} \right) \tag{6}$$

**Example 2.** Analysis of data "growth of the cumulative number of TREZ publications" [6], [7]

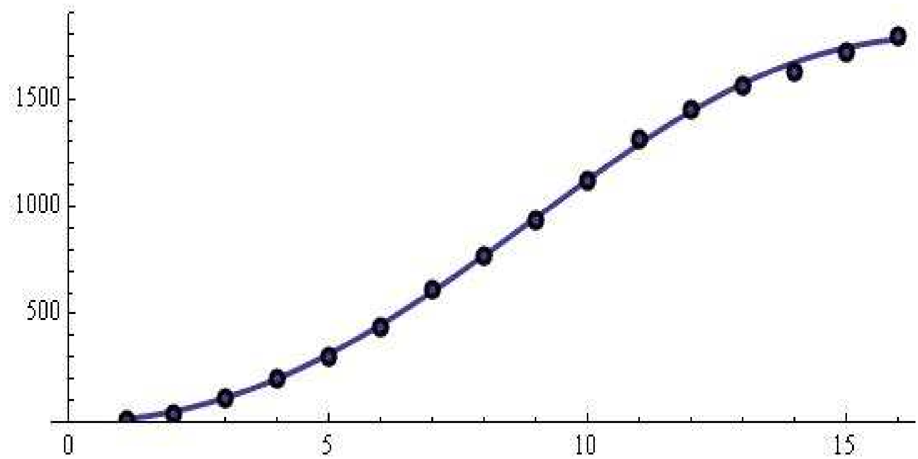


Figure 7: The fitted model $M^*(t)$.

$$
\begin{aligned}
&data\_Journal\\
&:= \{\{1.1, 5\}, \{2, 37\}, \{3, 107\}, \{4, 201\}, \{5, 298\}, \{6, 439\},\\
&\{7, 617\}, \{8, 773\}, \{9, 936\}, \{10, 1121\}, \{11, 1316\},\\
&\{12, 1451\}, \{13, 1563\}, \{14, 1629\}, \{15, 1722\}, \{16, 1788\}\};
\end{aligned}
$$

After that using the model $M^*(t) = \omega M(t)$ for $\lambda = 4.80604$, $\beta = 22.1$, $\theta = 0.9$, $\alpha = 2.33693$ and $\omega = 1800$ we obtain the fitted model (see, Fig. 7).

**Example 3.** Analysis of data set "US mobile", [7]

$$
\begin{aligned}
&dataMobile\\
&:= \{\{3, 20.29\}, \{4, 25.08\}, \{5, 30.81\}, \{6, 38.75\}, \{7, 45\}, \{8, 49.16\},\\
&\{9, 55.15\}, \{10, 62.852\}, \{11, 68.63\}, \{12, 76.64\}, \{13, 82.47\},\\
&\{14, 85.68\}, \{15, 89.14\}, \{16, 91.864\}, \{17, 95.28\}, \{18, 98.17\}\};
\end{aligned}
$$

After that using the model $M^*(t) = \omega M(t)$ for $\lambda = 3.10703$, $\beta = 20.1$, $\theta = 0.183849$, $\alpha = 5.26477$ and $\omega = 100$ we obtain the fitted model (see, Fig. 8).

**Example 4.** Software Failure Data – Release #1. These data derive from one major release of software products at Tandem Computers [8], [9]

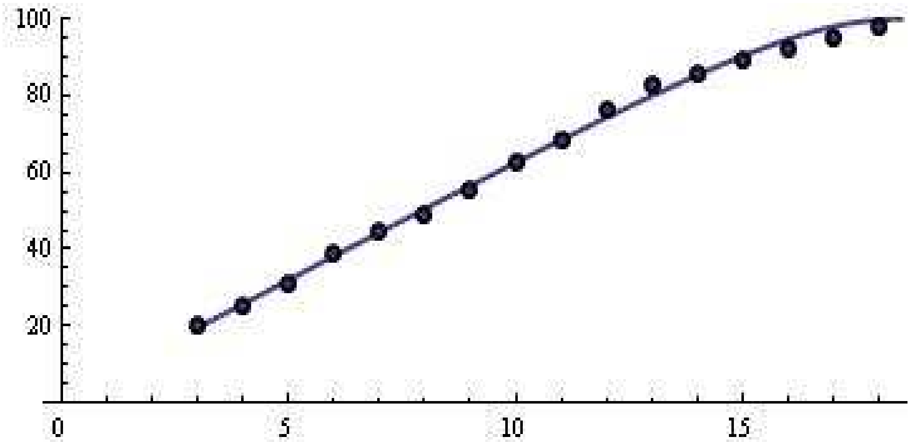


Figure 8: The fitted model $M^*(t)$.

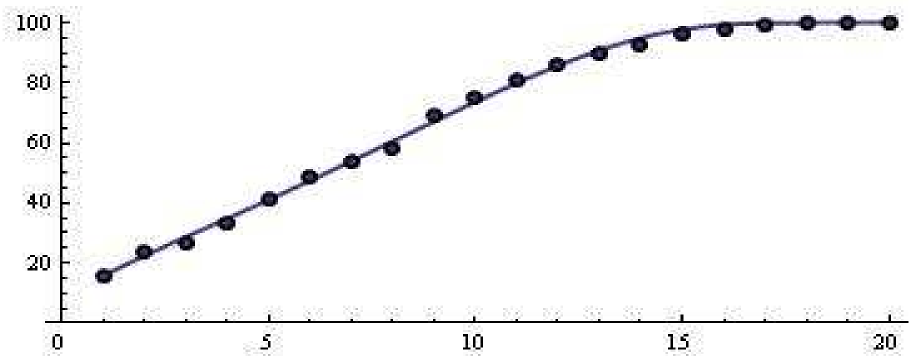


Figure 9: The fitted model $M^*(t)$.

$$data\_PhamNordmanZhang := \{\{1,16\},\{2,24\},\{3,27\},\{4,33\},$$

$$\{5,41\},\{6,49\},\{7,54\},\{8,58\},\{9,69\},\{10,75\},\{11,81\},\{12,86\},$$

$$\{13,90\},\{14,93\},\{15,96\},\{16,98\},\{17,99\},\{18,100\},\{19,100\},$$

$$\{20,100\}\};$$

For this data the fitted model for estimated parameters: $\omega = 100$; $\alpha = 0.0265009$; $\beta = 20.1$; $\lambda = 0.0336371$; $\theta = 1.69507$ is plotted on Fig. 9.

## 3. CONCLUDING REMARKS

Finally, we note that the studied model produces extremely good results, generally when approximating specific "cumulative data" from Computer Viruses Propagation, Debugging and Test theory.

For other approximation and modelling results, see [10]–[23].

We hope that the results will be useful for specialists in this scientific area.

## ACKNOWLEDGMENTS

This work has been accomplished with the financial support by the Grant No BG05M2 OP001-1.001-0003, financed by the Science and Education for Smart Growth Operational Program (2014-2020) and co-financed by the European Union through the European structural and Investment funds.

## REFERENCES

[1] M. Tahir, G. Cordeiro, M. Alizadeh, M. Mansoor, M. Zubair, G. Hamedani, The odd generalized exponential family of distributions with applications, *J. of Stat. Distr. and Appl.,* **2**, No. 1 (2015), 1–28.

[2] A. Hassan, E. Elshrpieny, R. Mohamed, Odd generalized exponential power function distribution: properties and applications, *J. of Sci., Gazi Univ.,* **32**, No. 1 (2019), 351–370.

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

[4] P. Szor, *The Art of Computer Virus Research and Defense*, Addison Wesley Professional (2005), ISBN: 0-321-30454-3.

[5] T. Hussain, H. Bakouch, Ch. Chesneau, A new probability model with application to heavy–tailed hydrological data, *Environmental and Ecological Statistics*, **26**, No. 2 (2019), 127–151.

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

[7] 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).

[8] A. Wood, Predicting software reliability, *IEEE Computer*, **11** (1996), 69–77.

[9] H. Pham, L. Nordmann, X. Zhang, A General Imperfect-Software-Debugging Model with S-Shaped Fault-Detection Rate, *IEEE Trans. Rel.*, **48**, No. 2 (1999), 169–175.

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

[11] A. Malinova, O. Rahneva, A. Golev, V. Kyurkchiev, Investigations on the Odd-Burr-III-Weibull cumulative sigmoid. Some applications, *Neural, Parallel, and Scientific Computations*, **27**, No. 1 (2019), 35–44.

[12] O. Rahneva, H. Kiskinov, I. Dimitrov, V. Matanski, Application of a Weibull Cumulative Distribution Function Based on m Existing Ones to Population Dynamics, *International Electronic Journal of Pure and Applied Mathematics*, **12**, No. 1 (2018), 111–121.

[13] E. Angelova, A. Malinova, V. Kyurkchiev, O. Rahneva, Investigations on a logistic model with weighted exponential Gompertz type correction. Some applications, *Neural, Parallel, and Scientific Computations*, **27**, No. 2 (2019), 105–114.

[14] N. Kyurkchiev, S. Markov, On the Hausdorff distance between the Heaviside step function and Verhulst logistic function, *J. Math. Chem.*, **54** (2016), 109–119.

[15] N. Kyurkchiev, On a Sigmoidal Growth Function Generated by Reaction Networks. Some Extensions and Applications, *Communications in Applied Analysis*, **23**, No. 3 (2019), 383–400.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] V. Kyurkchiev, A. Iliev, A. Rahnev, N. Kyurkchiev, *Some New Logistic Differential Models: Properties and Applications*, LAP LAMBERT Academic Publishing (2019), ISBN: 978-620-0-43442-5.

[21] 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.

[22] 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.

[23] 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.

[24] L. Pham, H. Pham, Software Reliability Models with Time-Dependent Hazard Function Based on Bayesian Approach, *IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans*, **30**, No. 1 (2000), 25–35.

[25] A. Gupta, R. Garg, Software Reliabilty for Distributed systems: Industry Perspective, *International Journal of Advanced Research in Computer Engineering & Technology*, **3**, No. 7 (2014), 2522–2526.