Neural, Parallel, and Scientific Computations, **27**, No. 2 (2019), 131-139 ISSN: 1061-5369

ON SOME INTRINSIC PROPERTIES OF THE CHI AND WEIGHTED ERLANG CUMULATIVE DISTRIBUTION FUNCTIONS

ANNA MALINOVA[1], TODORKA TERZIEVA[2], ANTON ILIEV[3], ASEN RAHNEV[4], AND NIKOLAY KYURKCHIEV[5]

[1,2,3,4,5]Faculty of Mathematics and Informatics
University of Plovdiv Paisii Hilendarski
24, Tzar Asen Str., 4000 Plovdiv, BULGARIA

ABSTRACT: The aim of this note is to study "saturation" of the Chi and weighted Erlang cumulative distribution functions

$$M(t) = 1 - \frac{1}{\Gamma(\frac{k}{2})}\Gamma(\frac{k}{2}, \frac{t^2}{2})$$

$$M_1(t) = 1 - \frac{1}{\Gamma(\lambda+\theta)}\Gamma(\lambda+\theta, \frac{t}{\beta}),$$

respectively, to the horizontal asymptote with respect to Hausdorff distance.

We prove upper and lower estimates for the one–sided Hausdorff approximation of the Heaviside step–function $h_{t_1}(t)$ by means of these families.

Numerical examples on real datasets (1. "actual data to estimate the number of software residual faults" and 2. "2017 meningitis outbreak data were obtained from Nigeria Centre for Disease Control") using *CAS Mathematica*, illustrating our results are given.

AMS Subject Classification: 41A46

Key Words: Chi cumulative function, weighted Erlang cumulative function, Heaviside function, Hausdorff distance, Upper and lower bounds

Received: May 29, 2019 ; **Accepted:** August 18, 2019 ;
Published: August 31, 2019. **doi:** 10.12732/npsc.v27i2.7
Dynamic Publishers, Inc., Acad. Publishers, Ltd. https://acadsol.eu/npsc

1. INTRODUCTION AND PRELIMINARIES

In [1], the authors studied some properties of the Wilson–Hilferty cumulative distribution

$$WH(t) = 1 - \frac{1}{\Gamma(\alpha)}\Gamma(\alpha, \frac{\alpha}{\lambda}t^3),$$

and in particular "saturation" to the horizontal asymptote.

Similar studies can be applied to other known cumulative distributions.

Definition 1. The Chi cumulative distribution function (CCDF) is given by

$$M(t) = P(\frac{k}{2}, \frac{t^2}{2})$$

where $P(s, x)$ is regularized gamma function, or

$$M(t) = 1 - \frac{1}{\Gamma(\frac{k}{2})}\Gamma(\frac{k}{2}, \frac{t^2}{2}) \tag{1}$$

where

$$\Gamma(s, x) = \int_x^\infty t^{s-1}e^{-t}dt$$

is the upper incomplete gamma function.

Definition 2. The weighted Erlang cumulative distribution function (WECDF) is given by

$$M_1(t) = 1 - \frac{1}{\Gamma(\lambda + \theta)}\Gamma(\lambda + \theta, \frac{t}{\beta}). \tag{2}$$

Definition 3. The shifted Heaviside step function is defined by

$$h_{t_1}(t) = \begin{cases} 0, & \text{if} \quad t < t_1, \\ [0, 1], & \text{if} \quad t = t_1, \\ 1, & \text{if} \quad t > t_1. \end{cases} \tag{3}$$

Definition 4. [2] 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,

$$\rho(f, g) = \max\{\sup_{A\in F(f)} \inf_{B\in F(g)} ||A - B||, \sup_{B\in F(g)} \inf_{A\in F(f)} ||A - B||\}, \tag{4}$$

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.

When studying the intrinsic properties of these distributions, it is also appropriate to study the "saturation" to the horizontal asymptote.

In this Section we prove upper and lower estimates for the one–sided Hausdorff approximation of the Heaviside step–function $h_{t_1}(t)$ by means of families (1) and (2).

2.1. THE CHI CUMULATIVE FUNCTION

Let t_1 is the unique positive root of the nonlinear equation $M(t_1) - \frac{1}{2} = 0$.

The one–sided Hausdorff distance d satisfies the relation

$$M(t_1 + d) = 1 - d. \tag{5}$$

The following theorem gives upper and lower bounds for d

Theorem 1. Let

$$\begin{aligned} p &= -\frac{\Gamma\left(\frac{k}{2}, \frac{t_1^2}{2}\right)}{\Gamma(\frac{k}{2})}, \\ q &= 1 + \frac{2^{1-\frac{k}{2}} e^{-\frac{t_1^2}{2}} t_1^{k-1}}{\Gamma(\frac{k}{2})} \\ s &= 2.1q. \end{aligned} \tag{6}$$

Let $s > e^{1.05}$ and $k > 0.5$.

For the one–sided Hausdorff distance d between $h_{t_1}(t)$ and the cumulative sigmoid (1) the following inequalities hold:

$$d_l = \frac{1}{s} < d < \frac{\ln s}{s} = d_r. \tag{7}$$

Proof. Let us examine the function:

$$F(d) = M(t_1 + d) - 1 + d. \tag{8}$$

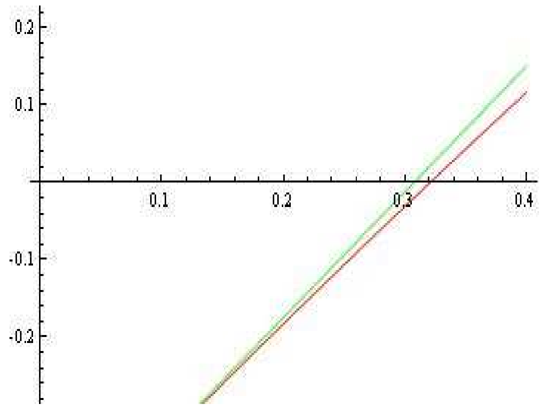

Figure 1: The functions $F(d)$ and $G(d)$ for $k = 1.1$.

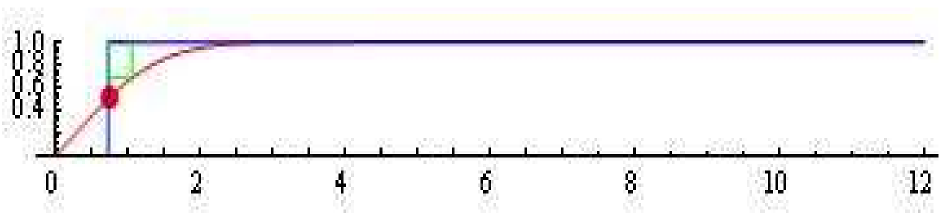

Figure 2: The cumulative function (1) for $k = 1.1$; $t_1 = 0.730256$; Hausdorff distance $d = 0.320908$; $d_l = 0.293014$; $d_r = 0.359685$.

$F'(d) > 0$ and the function F is increasing.

Consider the function

$$G(d) = p + qd. \tag{9}$$

From Taylor expansion we obtain $G(d) - F(d) = O(d^2)$.

Hence $G(d)$ approximates $F(d)$ with $d \to 0$ as $O(d^2)$ (see Fig. 1).

In addition $G'(d) > 0$ and the function G is also increasing.

Further, for $s > e^{1.05}$ and $k > 0.5$ we have

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

This completes the proof of the theorem.

The cumulative sigmoid (1) for $k = 1.1$ is visualized on Fig. 2.

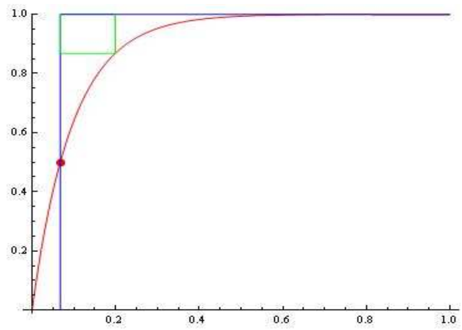

Figure 3: The cumulative function (2) for $\beta = 0.1; \lambda = 0.95; \theta = 0.03$; $t_2 = 0.0673801$; Hausdorff distance $d = 0.131901$; $d_l = 0.078354$; $d_r = 0.19953$.

2.2. THE WEIGHTED ERLANG CUMULATIVE FUNCTION

Let t_2 is the unique positive root of the nonlinear equation $M_1(t_2) - \frac{1}{2} = 0$.

The one–sided Hausdorff distance d satisfies the relation $M_1(t_2 + d) = 1 - d$.

The following theorem gives upper and lower bounds for d

Theorem 2. Let

$$p_1 = -\frac{\Gamma\left(\theta+\lambda, \frac{t_2}{\beta}\right)}{\Gamma(\theta+\lambda)},$$

$$q_1 = 1 + \frac{e^{-\frac{t_2}{\beta}} \left(\frac{t_2}{\beta}\right)^{\lambda+\theta-1}}{\beta\Gamma(\theta+\lambda)} \tag{10}$$

$$s_1 = 2.1 q_1.$$

Let $s > e^{1.05}$. With some constraints imposed on the parameters θ and λ which we will not explore here, for the one–sided Hausdorff distance d between $h_{t_2}(t)$ and the cumulative function (2) the following inequalities hold:

$$d_l = \frac{1}{s_1} < d < \frac{\ln s_1}{s_1} = d_r. \tag{11}$$

The proof follows the ideas given in this paper and will be omitted.

For some experiments, see Fig. 3–Fig. 4.

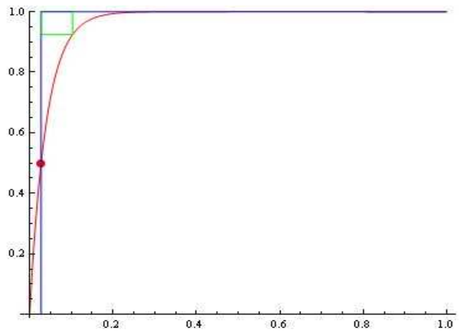

Figure 4: The cumulative function (2) for $\beta = 0.04; \lambda = 0.99; \theta = 0.01$; $t_2 = 0.0277259$; Hausdorff distance $d = 0.0755788$; $d_l = 0.0352734$; $d_r = 0.117976$.

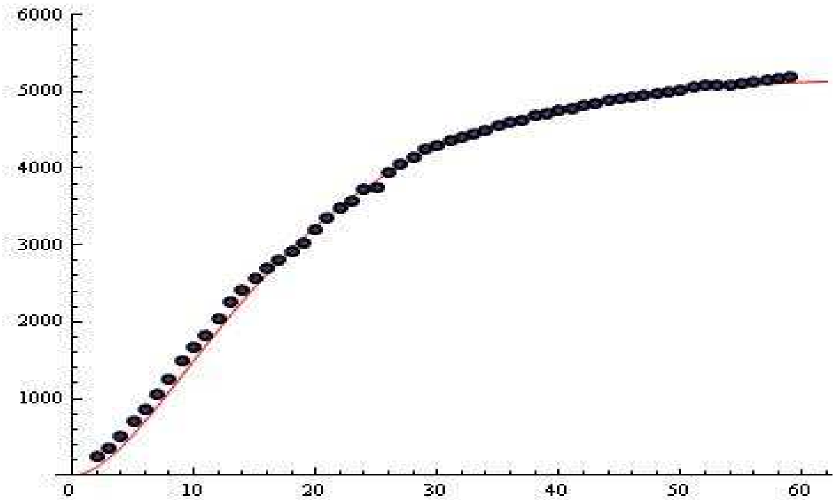

Figure 5: The fitted model $M_1^*(t)$.

3. NUMERICAL EXAMPLES

1. We analyze the following "actual data to estimate the number of software residual faults" [3]–[4].

After that using the model $M_1^*(t) = \omega M_1(t)$ for $\omega = 5186$, $\theta = 1.97$, $\lambda = 1 \times 10^{-9}$ and $\beta = 9.6$ we obtain the fitted model (see, Fig. 5).

2. We analyze the "2017 meningitis outbreak data were obtained from Nigeria Centre for Disease Control" [5]

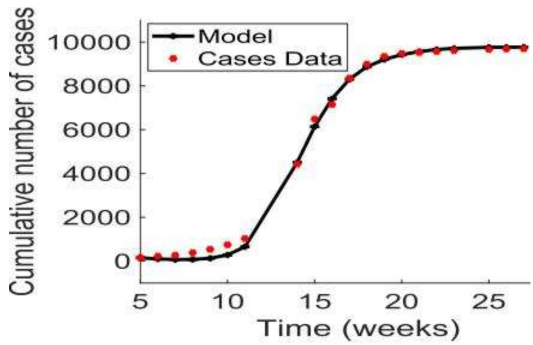

Figure 6: The model [5].

$data\_Meningitis$

$:= \{\{5, 158\}, \{6, 210\}, \{7, 263\}, \{8, 474\}, \{9, 579\}, \{10, 789\},$
$\{11, 1053\}, \{14, 4500\}, \{15, 6500\}, \{16, 7200\}, \{17, 8368\},$
$\{18, 9053\}, \{19, 9368\}, \{20, 9474\}, \{21, 9526\}, \{22, 9632\},$
$\{23, 9684\}, \{25, 9737\}, \{26, 9790\}, \{27, 9790\}\}.$

After that using the model $M_1^*(t) = \omega M_1(t)$ for $\omega = 9790$, $\theta = 26.95$, $\lambda = 0.01$ and $\beta = 0.531$ we obtain the fitted model (see, Fig. 7).

For some comparisons, see model [5] - Fig. 6.

4. CONCLUSION.

The aim of this note is to study "saturation" of the Chi and weighted Erlang cumulative functions.

For other results, see [6]–[11].

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.

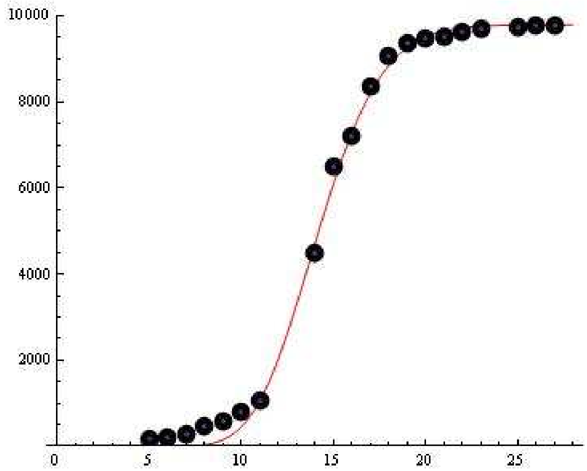

Figure 7: The fitted model $M_1^*(t)$.

REFERENCES

[1] N. Kyurkchiev, A. Iliev, A. Rahnev, A note on the Wilson–Hilferty distribution, *International Journal of Differential Equations and Applications*, **18**, No. 1 (2019), 31–36.

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

[3] T. Mitsuhashi, *A method of software quality evaluation*, JUSE Press, Tokyo (1981). (in Japanese)

[4] D. Satoh, A discrete Gompertz equation and a software reliability growth model, *IEICE Trans. Inf. and Syst.*, **E83–D**, No. 7 (2000), 1508–1513.

[5] F. Agusto, M. Leite, Optimal control and cost effective analysis of the 2017 meningitis outbreak in Nigeria, *Infectious Disease Modelling*, **4** (2019), 161–187.

[6] C. Kilber, S. Kotz, *Statistical Size Distributions in Economics and Actuarial Science*, Willey, New York (2003).

[7] E. Gomez–Deniz, E. Calderin–Ojeda, The Compound DGL/Erlang Distribution in the Collective Risk Model, *Revista de Metodos Cuantitativos Para la Economia y la Empresa*, **16** (2003), 121–142.

[8] R. Agarwal, S. Hristova, D. O'Regan, P. Kopanov, P–moment exponential stability of differential equations with random noninstantaneous impulses, *Int. J. of Pure and Appl. Math.*, **109**, No. 1 (2016), 9–28.

[9] J. Fang, Hyper–Erlang distribution model and its applications in wireless mobile networks, *Wireless Network*, **7** (2001), 211–219.

[10] K. Kim, N. Thomas, A fitting method with generalized Erlang distribution, *Simulation Modelling Practice and Theory*, **19** (2011), 1507–1517.

[11] 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.