A NOTE ON THE MODIFIED INVERSE RAYLEIGH CUMULATIVE SIGMOID. SOME APPLICATIONS

EVGENIA ANGELOVA\textsuperscript{1}, ANNA MALINOVA\textsuperscript{2}, TODORKA TERZIEVA\textsuperscript{3}, AND OLGA RAHNEVA\textsuperscript{4}

\textsuperscript{1,2,3}Faculty of Mathematics and Informatics
University of Plovdiv Paisii Hilendarski
24, Tzar Asen Str., 4000 Plovdiv, BULGARIA
\textsuperscript{4}Faculty of Economy and Social Sciences
University of Plovdiv Paisii Hilendarski
24, Tzar Asen Str., 4000 Plovdiv, BULGARIA

ABSTRACT: In this paper we prove upper and lower estimates for the one–sided Hausdorff approximation of the Heaviside step–function $h_{t_0}(t)$ by means of a modified inverse Rayleigh cumulative sigmoid (MIRCS).

Some applications of the presented cumulative sigmoid for analysis of the "cancer data" \cite{3}-\cite{4} and for approximating cdf of the number of Bitcoin received per address \cite{5} are considered. We analyze the "data on the development of the \textit{Drosophila melanogaster} population", published by biologist Raymond Pearl in 1920 (see, also \cite{6}).

Numerical examples, illustrating our results are given.

AMS Subject Classification: 41A46

Key Words: modified inverse Rayleigh cumulative sigmoid (MIRCS), supersaturation, heaviside function, Hausdorff distance, Upper and lower bounds

Received: December 28, 2018; Accepted: April 20, 2019; Published: April 24, 2019; doi: 10.12732/npsc.v27i2.1

1. INTRODUCTION

In a series of papers, we have explored the interesting task of approximating the Heaviside function \( h(t) \) with some logistic functions [7]–[26].

The task is important in the treatment of questions related to the study of the "supersaturation" - the object of the research in various fields.

In this paper we prove upper and lower estimates for the one–sided Hausdorff approximation of the Heaviside step–function \( h_{t_0}(t) \) by means of a new modified inverse Rayleigh cumulative sigmoid (MIRCS).

The proposed model can be successfully used to approximating data from Population Dynamics and Computer Viruses Propagation Theory.

2. PRELIMINARIES

**Definition 1.** The shifted Heaviside step function is defined by

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

(1)

**Definition 2.** [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,

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

(2)

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

**Definition 3.** The modified inverse Rayleigh cumulative sigmoid (MIRCS) is defined by [1]:

\[
    M(t) = e^{-\frac{t}{\alpha} - \theta \left( \frac{1}{t} \right)^2}
\]

(3)

where \( \alpha > 0 \) and \( \theta > 0 \).
3. MAIN RESULTS

3.1. APPROXIMATION RESULTS

Let

\[ t_0 = \frac{2\theta}{-\alpha + \sqrt{\alpha^2 + 4\theta \ln 2}}. \]  

(4)

Evidently, for the median we have \( M(t_0) = \frac{1}{2} \).

It is important to study the characteristic - ”supersaturation” of the sigmoid ((3)–(4)) to the horizontal asymptote.

In this Section we prove upper and lower estimates for the one–sided Hausdorff approximation of the Heaviside step–function \( h_{t_0}(t) \) by means of families (3)–(4).

The one–sided Hausdorff distance \( d \) satisfies the relation

\[ M(t_0 + d) = 1 - d. \]  

(5)

The following theorem gives upper and lower bounds for \( d \)

**Theorem 1.** Let

\[ p = -\frac{1}{2}, \]
\[ q = 1 + \frac{1}{2t_0^2} \left( \alpha + \frac{2\theta}{t_0} \right), \]
\[ \gamma = 2.1q. \]  

(6)

Let \( \gamma > e^{1.05} \). For the one–sided Hausdorff distance \( d \) between \( h_{t_0}(t) \) and the sigmoid (3)–(4) the following inequalities hold:

\[ d_l = \frac{1}{\gamma} < d < \frac{\ln \gamma}{\gamma} = d_r. \]  

(7)

**Proof.** Let us examine the function:

\[ F(d) = M(t_0 + d) - 1 + d. \]  

(8)

From \( F'(d) > 0 \) we conclude that function \( F \) is increasing.

Consider the function

\[ G(d) = p + qd. \]  

(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.
Figure 1: The functions $F(d)$ and $G(d)$ for $\alpha = 0.005; \theta = 0.001$.

Figure 2: The sigmoid (3)–(4) for $\alpha = 0.05; \theta = 0.01; t_0 = 0.161478$; Hausdorff distance $d = 0.195984; d_l = 0.109888; d_r = 0.242655$.

Further, for $\gamma > e^{1.05}$ we have

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

This completes the proof of the theorem.

Approximations of the $h_{t^*}(t)$ by model (3)–(4) for various $k$ and $p$ are visualized on Fig. 2–Fig. 3.

From the graphic it can be seen that the "saturation" is faster.
Figure 3: The sigmoid (3)–(4) for $\alpha = 0.005; \theta = 0.001; t_0 = 0.0417604$; Hausdorff distance $d = 0.090549; d_l = 0.0294588; d_r = 0.103835$.

4. SOME APPLICATIONS

The proposed model can be successfully used to approximating data from Population Dynamic, Biostatistics, Debugging Theory and Computer Viruses Propagation Theory.

4.1. APPROXIMATING CDF OF THE NUMBER OF BITCOIN RECEIVED PER ADDRESS

We consider the following data (see, [5]):

$$data_{CDF\ of\ Bitcoin\ received\ (in\ ransoms)\ per\ address\ in\ CL} := \{(1, 0.0857), (2, 0.1238), (3, 0.6571), (4, 0.6854), (5, 0.8381),
{6, 0.8476}, {7, 0.8810}, {8, 0.9095}, {9, 0.9143}, {10, 0.9333},
{12, 0.9429}, {14, 0.9571}, {18, 0.9667}, {20, 0.9762}, {23, 0.9810},
{27, 0.9857}, {40, 0.9905}, {46, 0.9952}, {59, 0.9981}\}.$$

Fig. 4 show cdf of the number of Bitcoin received per address respectively [5]. After that using the model

$$M(t) = \omega e^{-\frac{\alpha}{t_0}} - \theta t^2$$

for $\omega = 0.972975$, $\alpha = -0.387164$ and $\theta = 6.5794$ we obtain the fitted model (see, Fig. 5).
4.2. APPLICATION OF THE NEW CUMULATIVE SIGMOID FOR ANALYSIS OF THE "CANCER DATA"

We will illustrate the advances of the modified inverse Rayleigh cumulative sigmoid for approximation and modelling of "cancer data" (for some details see, [3]–[4]).

<table>
<thead>
<tr>
<th>days</th>
<th>4</th>
<th>7</th>
<th>10</th>
<th>12</th>
<th>14</th>
<th>17</th>
<th>19</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R(t)$</td>
<td>0.415</td>
<td>0.794</td>
<td>1.001</td>
<td>1.102</td>
<td>1.192</td>
<td>1.22</td>
<td>1.241</td>
<td>1.3</td>
</tr>
</tbody>
</table>
Figure 6: The model $M(t)$ based on the "cancer data".

<table>
<thead>
<tr>
<th>$t$</th>
<th>$M(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>39</td>
</tr>
<tr>
<td>12</td>
<td>105</td>
</tr>
<tr>
<td>15</td>
<td>152</td>
</tr>
<tr>
<td>18</td>
<td>225</td>
</tr>
<tr>
<td>21</td>
<td>390</td>
</tr>
<tr>
<td>25</td>
<td>547</td>
</tr>
<tr>
<td>29</td>
<td>618</td>
</tr>
<tr>
<td>33</td>
<td>791</td>
</tr>
<tr>
<td>36</td>
<td>877</td>
</tr>
<tr>
<td>39</td>
<td>938</td>
</tr>
</tbody>
</table>

Table 1: The "cancer data" [3]–[4]

The model $M(t)$ based on the data from Table 1 for the estimated parameters:

$$\omega = 1.555; \alpha = 3.60844; \theta = 6.93094$$

is plotted on Fig. 6.

4.3. APPROXIMATING THE "DATA ON THE DEVELOPMENT OF THE DROSOPHILA MELANOGASTER POPULATION"

We analyze the "data on the development of the Drosophila melanogaster population", published by biologist Raymond Pearl in 1920 (see, also [6]).

We consider the following data:

```
data_Pearl := {{9, 39}, {12, 105}, {15, 152}, {18, 225}, {21, 390}, {25, 547},
{29, 618}, {33, 791}, {36, 877}, {39, 938}}.
```

After that using the model $M(t)$ for $\omega = 3479.18$, $\alpha = 53.2571$ and $\theta = -111.646$ we obtain the fitted model (see, Fig. 7).
5. CONCLUSION

In this paper we prove upper and lower estimates for the one–sided Hausdorff approximation of the Heaviside step–function \( h_{t_0}(t) \) by means of a modified inverse Rayleigh cumulative sigmoid (MIRCS).

Some applications of the presented cumulative sigmoid for analysis of the: "cancer data" [3]–[4], "data for approximating cdf of the number of Bitcoin received per address" [5] and "data on the development of the Drosophila melanogaster population" [6] are considered.

We propose a software module within the programming environment CAS Mathematica for the analysis of the considered family of functions.

The module offers the following possibilities:

- calculation of the H-distance between the \( h_{t_0} \) and the function \( M(t) \);
- software tools for animation and visualization.

Finally, we will point out that when the data accepted in the literature as basis are approximated when compare existing and new models the results obtained with the model in this article are satisfactory.
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


