COMMENTS ON A ZUBAIR–G FAMILY OF CUMULATIVE LIFETIME DISTRIBUTIONS. SOME EXTENSIONS

NIKOLAY KYURKCHIEV¹, ANTON ILIEV², AND ASEN RAHNEV³
¹,²,³Faculty of Mathematics and Informatics
University of Plovdiv “Paisii Hilendarski”
24 Tzar Asen Str., 4000 Plovdiv, BULGARIA

ABSTRACT: In this paper we study the one–sided Hausdorff approximation of the shifted Heaviside step function by a class of the Zubair–G family of cumulative lifetime distribution. The estimates of the value of the best Hausdorff approximation obtained in this article can be used in practice as one possible additional criterion in ”saturation” study.

As an illustrative example we consider the modelling of the growth of red abalone (Haliotis Rufescens) in Northern California.

We also look at a possible extension, which we call α–Zubair–G Family. For the analysis of said dataset with the new (cdf), some comparisons are made. Finally, the potentiality of the software reliability models analyzed by means of real dataset. Numerical examples, illustrating our results are presented using programming environment CAS Mathematica.

AMS Subject Classification: 68N30, 41A46
Key Words: Zubair–G family of cumulative lifetime distribution, α–Zubair–G Family of (cdf), Heaviside function, Hausdorff approximation, upper and lower bounds

Received: March 7, 2018; Accepted: October 20, 2018; Published: November 16, 2018; doi: 10.12732/caa.v23i1.1
1. INTRODUCTION

In [1], a new family of lifetime distributions, called the Zubair–G family of distributions is introduced.

The new family is defined by the following cumulative distribution function (cdf)
\[ F(t; \lambda) = \frac{e^{\lambda G^2(t)} - 1}{e^\lambda - 1}, \]  
(1)

where \( \lambda > 0 \).

If, \( G \) is the (cdf) of the baseline model, then the distribution function (1) will be the (cdf) of the Zubair–G family.

For example, if \( G \) be (cdf) of the Weibull distribution given by
\[ G(t) = 1 - e^{-\beta t^m} \]  
(2)

then the (cdf) of the Z–Weibull distribution has the form
\[ F(t) = \frac{e^{\lambda (1 - e^{-\beta t^m})^2} - 1}{e^\lambda - 1} \]  
(3)

where \( m > 0 \) and \( \gamma > 0 \).

We consider the following class of this family with application to the population dynamics and debugging theory:
\[ M(t) = \frac{e^{\lambda (1 - e^{-\beta t^m})^2} - 1}{e^\lambda - 1}, \]  
(4)

with
\[ t_0 = \left( -\frac{1}{\beta} \ln \left( 1 - \frac{1}{\sqrt{\lambda}} \sqrt{\ln \left( \frac{e^\lambda + 1}{2} \right)} \right) \right)^{-\frac{1}{m}}; \quad M(t_0) = \frac{1}{2}. \]  
(5)

Some extensions of the well-known Poisson, Poisson–exponential, Chen, Exponentiated Chen, modified Weibull and Burr distributions can be found in: [2]–[13].

For some approximation, computational and modelling aspects, see [14]–[30].

Some software reliability models, can be found in [31]–[50].
In this note we study the Hausdorff approximation of the *shifted Heaviside step function*

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

by this family.

**Definition 1.** [51] 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,

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

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

We recall that completed graph of \( f \) is the closure of the graph of \( f \) as a subset of \( \Omega \times \mathbb{R} \). If the graph of an interval function \( f \) equals \( F(f) \), then the \( f \) is called S-continuous.

The Hausdorff distance \( \rho(f, g) = \max\{\vec{\rho}(f, g), \vec{\rho}(g, f)\} \) defines a metric in the set of the S-continuous interval functions [52]–[55].

We propose a software modules (intellectual properties) within the programming environment CAS Mathematica for the analysis.

We also look at a possible extension, which we call \( \alpha \)–Zubair–G Family. Some comparisons are made.

The models have been tested with real data (the growth of red abalone (*Haliotis Rufescens*) in Northern California).

**2. MAIN RESULTS**

The one–sided Hausdorff distance \( d \) between the function \( h_{t_0}(t) \) and the sigmoidal function - ((4)–(5)) satisfies the relation

\[ M(t_0 + d) = 1 - d. \]
The following theorem gives upper and lower bounds for $d$

**Theorem 1.** Let

\[
p = -\frac{1}{2},
\]

\[
q = 1 + \frac{\beta \lambda m (1 + e^\lambda) \sqrt{\ln \frac{e^\lambda + 1}{2}}}{(e^\lambda - 1) \sqrt{\lambda}} \left( 1 - \frac{1}{\sqrt{\lambda}} \sqrt{\ln \frac{e^\lambda + 1}{2}} \right) \times \left( -\frac{1}{\beta} \ln \left( 1 - \frac{1}{\sqrt{\lambda}} \sqrt{\ln \frac{e^\lambda + 1}{2}} \right) \right)^{m-1},
\]

\[
r = 2.1q.
\]

For the one-sided Hausdorff distance $d$ between $h_{t_0}(t)$ and the sigmoid ((4)–(5)) the following inequalities hold for: $q > e^{1.05}/2.1$

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

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

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

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

Consider the function

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

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 Figure 1).

In addition $G'(d) > 0$.

Further, for $q > e^{1.05}/2.1$ we have $G(d_l) < 0$ and $G(d_r) > 0$.

This completes the proof of the theorem.

3. NUMERICAL EXAMPLES

The model ((4)–(5)) for $\beta = 15; \lambda = 0.1; m = 2; t_0 = 0.289642$ is visualized on Figure 2. From the nonlinear equation (6) and inequalities (8) we have: $d = 0.134996, d_l = 0.104943, d_r = 0.236577.$
Figure 1: The functions $F(d)$ and $G(d)$ for $\beta = 15; \lambda = 0.1; m = 2$.

The model (4–5) for $\beta = 20, \lambda = 0.15, m = 10, t_0 = 0.759288$ is visualized on Figure 3. From the nonlinear equation (6) and inequalities (8) we have: $d = 0.0762637, d_l = 0.0612591, d_r = 0.171075$.

The model (4–5) for $\beta = 30, \lambda = 0.5, m = 30, t_0 = 0.902554$ is visualized on Figure 4. From the nonlinear equation (6) and inequalities (8) we have: $d = 0.0346169, d_l = 0.0255515, d_r = 0.093699$.

From the above examples, it can be seen that the proven estimates (see Theorem 1) for the value of the Hausdorff approximation is reliable when assessing the important characteristic - "saturation".

This characteristic (as we have already shown in our previous publications) has its equal participation together with the other two characteristics - "confidence intervals" and "confidence bounds" in the area of the Software Reliability Theory.

We propose a software module (intellectual properties) within the programming environment CAS Mathematica for the analysis of the considered family $M(t)$.

The module offers the following possibilities:

— generation of the function under user defined values of the parameters $m, \lambda, \theta$;

— calculation of the H-distance $d$ between the function $h_{t_0}(t)$ and the
Figure 2: The model (4)-(5) for $\beta = 15; \lambda = 0.1; m = 2, t_0 = 0.289642$; H-distance $d = 0.134996$, $dl = 0.104943$, $dr = 0.236577$.

Figure 3: The model (4)-(5) for $\beta = 20, \lambda = 0.15, m = 10, t_0 = 0.759288$; H-distance $d = 0.0762637$, $dl = 0.0612591$, $dr = 0.171075$.

sigmoid $M(t)$;
— software tools for animation and visualization.

We examine the following data. (The small data for modeling the growth of red abalone is shown in Table 1. For more details, see [56]).
Figure 4: The model \((4)-(5)\) for \(\beta = 30, \lambda = 0.5, m = 30, t_0 = 0.902554; H\)-distance \(d = 0.0346169, d_l = 0.0255515, d_r = 0.093699\).

The model
\[
M(t) = A \left( \frac{e^{\lambda(1-e^{-\beta m t})^2} - 1}{e^\lambda - 1} \right)
\]

based on the data of Table 1 for the estimated parameters:
\[
A = 179.6; \quad \beta = 0.40179; \quad \lambda = 1.48583; \quad m = 0.962607
\]
is plotted on Figure 5.

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

For the predictive power (PP) criterion:
\[
PP = \sum_{i=1}^{n} \left( \frac{M(t_i) - y_i}{y_i} \right)^2
\]
measures the distance of model actual data from the estimates against the actual data, we find \(PP = 0.207925\).

4. APPENDIX

Following the ideas for extending of the well-known Poisson, Poisson–exponential, Chen, Exponentiated Chen, modified Weibull and Burr distributions
we consider the new modified $\alpha$–Zubair–$G$ Family of Cumulative Lifetime Distributions:

$$M_{\alpha}(t) = \left( \frac{e^{\lambda (1 - e^{-\beta t m})^2} - 1}{e^{\lambda} - 1} \right)^{\alpha}, \quad \alpha > 0$$

(11)

with

$$t_0 = \left( -\frac{1}{\beta} \ln \left( 1 - \frac{1}{\sqrt{\lambda}} \sqrt{\ln \left( 1 + (e^{\lambda} - 1) \left( \frac{\frac{1}{2}}{\alpha} \right)^{\frac{1}{\alpha}} \right)} \right) \right)^{\frac{1}{m}},$$

(12)

$$M_{\alpha}(t_0) = \frac{1}{2}.$$

The one–sided Hausdorff distance $d$ between the function $h_{t_0}(t)$ and the sigmoidal function - ((11)–(12)) satisfies the relation

$$M_{\alpha}(t_0 + d) = 1 - d.$$

(13)

The following theorem gives upper and lower bounds for $d$. 

<table>
<thead>
<tr>
<th>Age</th>
<th>Length (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16.1</td>
</tr>
<tr>
<td>2</td>
<td>33.9</td>
</tr>
<tr>
<td>3</td>
<td>54.3</td>
</tr>
<tr>
<td>4</td>
<td>76.2</td>
</tr>
<tr>
<td>5</td>
<td>97.8</td>
</tr>
<tr>
<td>6</td>
<td>117.1</td>
</tr>
<tr>
<td>7</td>
<td>133.3</td>
</tr>
<tr>
<td>8</td>
<td>146.5</td>
</tr>
<tr>
<td>9</td>
<td>157.2</td>
</tr>
<tr>
<td>10</td>
<td>166</td>
</tr>
<tr>
<td>11</td>
<td>173.3</td>
</tr>
<tr>
<td>12</td>
<td>179.6</td>
</tr>
</tbody>
</table>

Table 1: Data for modeling the growth of red abalone *Haliotis Rufescens* in Northern California [56]
Theorem 2. Let
\[ p = -\frac{1}{2}, \]
\[ q = 1 + \frac{2\beta\lambda m}{\sqrt{\lambda(\lambda^2 - 1)}} (1 + (\lambda^2 - 1)0.5^{\frac{1}{\alpha}})0.5^{\frac{\alpha - 1}{\alpha}} \times \sqrt{\ln(1 + (\lambda^2 - 1)0.5^{\frac{1}{\alpha}})} \left(1 - \sqrt{\frac{\ln(1 + (\lambda^2 - 1)0.5^{\frac{1}{\alpha}})}{\lambda}}\right) \times \left(-\frac{1}{\beta} \ln \left(1 - \sqrt{\frac{\ln(1 + (\lambda^2 - 1)0.5^{\frac{1}{\alpha}})}{\lambda}}\right)\right)^{\frac{m-1}{m}}, \]
\[ r = 2.1q. \]
Figure 6: The functions $F_\alpha(d)$ and $G_\alpha(d)$ for $\beta = 20; \lambda = 0.15; m = 10; \alpha = 2$.

For the one–sided Hausdorff distance $d$ between $ht_0(t)$ and the sigmoid ((11)–(12)) the following inequalities hold for: $q > \frac{e^{1.05}}{2.1}$

$$d_l = \frac{1}{r} < d < \frac{\ln r}{r} = d_r.$$  (15)

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

$$F_\alpha(d) = M_\alpha(t_0 + d) - 1 + d.$$  (16)

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

Consider the function

$$G_\alpha(d) = p + qd.$$  (17)

From Taylor expansion we obtain $G_\alpha(d) - F_\alpha(d) = O(d^2)$.

Hence $G_\alpha(d)$ approximates $F_\alpha(d)$ with $d \to 0$ as $O(d^2)$ (see Figure 6).

In addition $G'_\alpha(d) > 0$.

Further, for $q > \frac{e^{1.05}}{2.1}$ we have $G_\alpha(d_l) < 0$ and $G_\alpha(d_r) > 0$.

This completes the proof of the theorem.

For example, the model ((11)–(12)) for $\beta = 20, \lambda = 0.15, m = 10, \alpha = 2, t_0 = 0.790114$ is visualized on Figure 7.
Figure 7: The model ((11)–(12)) for $\beta = 20$, $\lambda = 0.15$, $m = 10$, $\alpha = 2$, $t_0 = 0.790114$; H–distance $d = 0.0647847$, $d_l = 0.0478774$ and $d_r = 0.145505$.

Figure 8: The model $M_\alpha(t)$.

From the nonlinear equation (13) and inequalities (15) we have: $d = 0.0647847$, $d_l = 0.0478774$ and $d_r = 0.145505$. 
4.1. NUMERICAL EXAMPLE

The model

\[ M_{\alpha}(t) = A \left( \frac{e^{\lambda(1-e^{-\beta t})^2} - 1}{e^\lambda - 1} \right)^\alpha \]

based on the data of Table 1 for the estimated parameters: \( A = 179.6; \ \beta = 0.0034666666; \ \lambda = 0.289174; \ m = 2.75779; \ \alpha = 0.214726 \) is plotted on Figure 8.

For the predictive power (PP) criterion:

\[ PP = \sum_{i=1}^{n} \left( \frac{M_{\alpha}(t_i) - y_i}{y_i} \right)^2 \]

we find \( PP = 0.0041167 \).

Comparison between the models \( M(t) \) and \( M_{\alpha}(t) \) show a good fit by the presented new model \( M_{\alpha}(t) \) (see Figure 9).
4.2. APPLICATION TO THE SOFTWARE RELIABILITY GROWTH THEORY

We examine the following data. (The small on-line data entry software package test data, available since 1980 in Japan [58], is shown in Table 2. For more details, see [57]).

<table>
<thead>
<tr>
<th>Testing time (day)</th>
<th>Failures</th>
<th>Cumulative failures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>19</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>26</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>31</td>
</tr>
<tr>
<td>16</td>
<td>6</td>
<td>37</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>38</td>
</tr>
<tr>
<td>18</td>
<td>3</td>
<td>41</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>42</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>45</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>46</td>
</tr>
</tbody>
</table>

Table 2: On-line IBM entry software package [58]

The fitted model $M_\alpha(t)$ based on the data of Table 2 for the estimated parameters:

$A = 46; \beta = 2.05825 \times 10^{-6}; \lambda = 2.72793; m = 4.75613; \alpha = 0.145456$
is plotted on Figure 10.

We will explicitly note that in some cases the presented software reliability model provides better results than other much more sophisticated models.

ACKNOWLEDGMENTS

This work has been supported by the project BG05M2OP001-1.001-0003-01 “CoE on Informatics and ICT” supported by the Operational Programme “Science and Education for Smart Growth”.

REFERENCES


