A novel technique for generating families of continuous distributions

Boikanyo Makubate*, Regent Retrospect Musekwa

Department of Mathematics, Botswana International University of Science and Technology, Palapye, Botswana

Abstract In this paper, we present the generalized flexible-G family for creating several continuous distributions. Our new technique features are that it adds only two extra shape parameters to any chosen continuous distribution and is not derived from any parent distribution that currently exists. Several special cases of this family are provided. The generalized flexible-G family offers significant improvements in flexibility, fit, and applicability across a wide range of fields. The family's model parameters are estimated using the maximum likelihood estimation method. A simulation study is conducted to assess the consistency of the maximum likelihood estimates. The generalized flexible log-logistic, a specific case of our novel family, is applied to both patient's analgesia and reliability data in order to illustrate the significance of the family. The generalized flexible log-logistic outperforms several competitive models provided in this paper. Furthermore, the generalized flexible log-logistic performs better than traditional distributions such as the BurrXII, Gumbel, and Weibull models.

Keywords New flexible-G; Generalized flexible-G; Family of distributions; Continuous distributions; Maximum likelihood estimation.

AMS 2010 subject classifications 62E10, 62F30

DOI: 10.19139/soic-2310-5070-2068

1. Introduction

In our real-life applications, statistics is essential through the use of statistical methods that heavily rely on the standard probability distributions. Nonetheless, a number of statistical issues defy the assumptions of these traditional models. Many real-world phenomena exhibit data patterns that are not adequately captured by these models. Different fields have unique requirements. For instance, hydrologists might need distributions that accurately model extreme rainfall, while reliability engineers might require models for lifetimes of complex systems. Also, existing distributions might fail to capture certain empirical regularities observed in specific domains, such as the heavy tails in financial returns or the asymmetry in biological measurements, flexibility and a closer fit to empirical data, capturing features such as skewness, kurtosis, and multimodality. As a result of this, there is a growing need to create flexible distributions in order to draw trustworthy conclusions and make informed judgments [\[1,](#page-16-0) [2,](#page-16-1) [6,](#page-16-2) [11,](#page-16-3) [13,](#page-16-4) [14,](#page-16-5) [15,](#page-16-6) [22\]](#page-17-0). By expanding the toolbox of available distributions, statisticians and researchers can achieve more accurate and insightful analyses. More work on the development of new models was done (see [\[8,](#page-16-7) [9,](#page-16-8) [10,](#page-16-9) [12,](#page-16-10) [19,](#page-16-11) [24\]](#page-17-1))

Recently, [\[25\]](#page-17-2) introduced a new technique of creating families of distributions called the new flexible generalized family (NFGF). This family was primarily designed to take account of the non-symmetrical behavior of the parent distribution. For any arbitrary baseline cdf distribution $G(t)$, the cdf and pdf of the NFGF are given by

$$
F_{NFGF}(t) = 1 - \bar{G}(t)^{G(t)},\tag{1}
$$

ISSN 2310-5070 (online) ISSN 2311-004X (print) Copyright © 2024 International Academic Press

[∗]Correspondence to: Boikanyo Makubate (Email: makubateb@biust.ac.bw). Department of Mathematics, Botswana International University of Science and Technology, Palapye, Botswana.

1232 A NOVEL TECHNIQUE FOR GENERATING FAMILIES OF CONTINUOUS DISTRIBUTIONS

and

$$
f_{NFGF}(t) = g(t; \varphi)\bar{G}(t)^{G(t)} \left(G(t)(G(t)^{-1} - \log(\bar{G}(t))) \right), \tag{2}
$$

respectively. In a similar spirit, we extend the family of distributions given in [\(1\)](#page-0-1) with the incorporation of two additional shape parameters. These parameters control the skewness and the tail weight. The new family of distributions termed the generalized flexible -G (GFG) is defined as follows. Consider a continuous distribution function $G(t; \varphi)$ with density function $g(t; \varphi)$. Then, the GFG has its cdf as

$$
F(t; \phi, \psi, \varphi) = \left[1 - \bar{G}(t; \varphi)^{\phi G(t; \varphi)}\right]^{\psi}, \tag{3}
$$

where ϕ and ψ are positive non zero shape parameters. Thus, this family of distributions has its pdf as

$$
f(t; \phi, \psi, \varphi) = \psi \phi g(t; \varphi) \bar{G}(t; \varphi)^{\phi G(t; \varphi)} \left[\frac{G(t; \varphi)}{\bar{G}(t; \varphi)} - \log(\bar{G}(t; \varphi)) \right] \left[1 - \bar{G}(t; \varphi)^{\phi G(t; \varphi)} \right]^{\psi - 1}.
$$
 (4)

The GFG has its survival function and hazard rate function (hrf) as

$$
s(t) = 1 - \left[1 - \bar{G}(t; \varphi)^{\phi G(t; \varphi)}\right]^{\psi},\tag{5}
$$

and

$$
hrf(t) = \frac{\psi \phi g(t; \varphi) \bar{G}(t; \varphi)^{\phi G(t; \varphi)} \left[\frac{G(t; \varphi)}{\bar{G}(t; \varphi)} - \log(\bar{G}(t; \varphi)) \right] \left[1 - \bar{G}(t; \varphi)^{\phi G(t; \varphi)} \right]^{\psi - 1}}{1 - \left[1 - \bar{G}(t; \varphi)^{\phi G(t; \varphi)} \right]^{\psi}},
$$
(6)

respectively. For $u \in (0, 1)$, the GFG has its quantile function as

$$
\phi G(t; \varphi) \log(\bar{G}(t; \varphi) = \log(1 - u^{1/\psi}). \tag{7}
$$

Equation [\(7\)](#page-1-0) can be solved numerically via some softwares such as R, MATHEMATICA, MAPLE and Ox. The strength of this research is solely based on the fact that the family of distributions defined in [\(3\)](#page-1-1) is not developed from any well-known parent model similar to the T-X family [\[1\]](#page-16-0), cubic rank transmuted-G [\[2\]](#page-16-1), generalized exponentiated-G family [\[6\]](#page-16-2), exponentiated-G family [\[11\]](#page-16-3), alpha power transformation family [\[13\]](#page-16-4), Marshall Olkin family [\[14\]](#page-16-5), alpha log power transformed-G family [\[15\]](#page-16-6), transmuted-G family [\[22\]](#page-17-0) and the new flexible-G family [\[25\]](#page-17-2).

This research was primarily driven by a combination of practical needs, theoretical advancements, the desire to improve statistical modeling and the need to develop a novel technique for generating families of continuous distributions. This novel technique can provide better fits to empirical data, unify existing models, offer more flexible parameterizations, and address specific challenges in various applications.

2. Special Cases

Here we consider examples of the family in [\(3\)](#page-1-1) for different standard distributions, namely for uniform (U), exponential (E), log-logistic (LLoG), Topp-Leone (TL), Pareto (P) and half-logistic (HL) distributions. Figures [1,](#page-2-0) [2,](#page-3-0) [3,](#page-5-0) [4,](#page-6-0) [5,](#page-7-0) and [6](#page-9-0) demonstrate how the hrf of the GFG distribution can take on a variety of very flexible shapes, including bathtub, bathtub followed by upside-down bathtub, upside-down bathtub, constant, increasing and decreasing. We also provide moments, standard deviation (SD), variance $(Var(T))$, skewness (S), kurtosis (K) and quantiles for selected values of $\Omega = (\phi, \psi, \delta)$.

2.1. Generalized flexible uniform distribution

If the parent distribution is uniform, such that $G(t; \delta) = t/\delta$ and $q(t; \delta) = 1/\delta$ where $0 < t < \delta$, then the cdf and pdf of the generalized flexible uniform (GFU) model are respectively given by

$$
F_{GFU}(t; \phi, \psi, \delta) = \left[1 - (1 - t/\delta)^{\phi t/\delta}\right]^{\psi},\tag{8}
$$

and

$$
f_{GFU}(t; \phi, \psi, \delta) = \psi \phi \delta^{-1} \left(\frac{\delta - t}{\delta} \right)^{\phi t/\delta} \left[\frac{t/\delta}{1 - t/\delta} - \log(1 - t/\delta) \right] \left[1 - (1 - t/\delta)^{\phi t/\delta} \right]^{\psi - 1}.
$$
 (9)

The GFU has its i^{th} moment and quantile function as

Figure 1. GFU distribution pdf and hrf plots

$$
E(T^{i}) = \int_{0}^{\infty} t^{i} f_{GFU}(t; \phi, \psi, \delta) dt,
$$
\n(10)

and

$$
\phi t/\delta \log(1 - t/\delta) = \log(1 - u^{1/\psi}),\tag{11}
$$

respectively. Moments associated with equation [\(10\)](#page-2-1) for selected values of Ω are given in Table [1,](#page-3-1) whereas Table [2](#page-3-2) shows some quantiles for selected values of Ω .

2.2. Generalized flexible exponential distribution

Let the parent distribution be exponential with cdf and pdf given by $G(t; \delta) = 1 - e^{-\delta t}$ and $g(t; \delta) = \delta e^{-\delta t}$ for non negative δ , then the cdf and pdf of the generalized flexible exponential (GFE) model are respectively given by

$$
F_{GFE}(t; \phi, \psi, \delta) = \left[1 - e^{-\delta t \phi (1 - e^{-\delta t})}\right]^{\psi}, \qquad (12)
$$

and

$$
f_{GFE}(t; \phi, \psi, \delta) = \frac{\psi \phi \delta e^{-\delta t}}{e^{\delta t \phi (1 - e^{-\delta t})}} \left[\frac{1 - e^{-\delta t}}{e^{-\delta t}} - \log \left(e^{-\delta t} \right) \right] \left[1 - e^{-\delta t \phi (1 - e^{-\delta t})} \right]^{\psi - 1}.
$$
 (13)

	(0.5, 0.5, 0.5)	(0.5, 1.3, 0.5)	(0.5, 1.8, 0.5)	(1.5, 1.9, 0.5)	(2.2, 2.5, 0.5)
E(T)	0.2962	0.5111	0.5868	0.2978	0.2502
$E(T^2)$	0.1654	0.3344	0.4075	0.1182	0.0809
$E(T^3)$	0.1132	0.2463	0.3100	0.0564	0.0314
$E(T^4)$	0.0856	0.1943	0.2494	0.0305	0.0140
$E(T^5)$	0.0688	0.1602	0.2084	0.0181	0.0069
SD	0.2787	0.2705	0.2514	0.1719	0.1354
Var(T)	0.0777	0.0732	0.0632	0.0296	0.0183
S	6.5438	12.9317	17.6919	25.2097	38.6252
K	2.5739	1.8980	2.0198	3.0740	3.5263

Table 1. GFU distribution table of moments

Table 2. GFU distribution table of quantiles

\boldsymbol{u}	(0.5, 0.5, 0.5)	(1.5, 0.3, 0.1)	(2.5, 0.8, 0.5)	(3.5, 0.9, 0.5)	(3.2, 0.5, 1.5)
0.1	0.0071	0.0000	0.0513	0.0633	0.0084
0.2	0.0279	0.0007	0.0876	0.0973	0.0334
0.3	0.0617	0.0001	0.1202	0.1261	0.0747
0.4	0.1073	0.0005	0.1514	0.1527	0.1320
0.5	0.1633	0.0017	0.1823	0.1787	0.2054
0.6	0.2280	0.0047	0.2141	0.2053	0.2960
0.7	0.2999	0.0011	0.2481	0.2339	0.4068
0.8	0.3766	0.0023	0.2869	0.2667	0.5459
0.9	0.4527	0.0433	0.3368	0.3102	0.7383

Figure 2. GFE distribution pdf and hrf plots

The GFE has its i^{th} moment and quantile function as

$$
E(T^{i}) = \int_{0}^{\infty} t^{i} f_{GFE}(t; \phi, \psi, \delta) dt,
$$
\n(14)

and

$$
-\phi \delta t (1 - e^{-\delta t}) = \log(1 - u^{1/\psi}),
$$
\n(15)

respectively. For some selected values of Ω , moments associated with equation [\(14\)](#page-3-3) and quantiles associated with Equation [\(15\)](#page-3-4) are given in Table [3](#page-4-0) and Table [4,](#page-4-1) respectively.

	(0.5, 0.5, 0.5)	(0.5, 1.3, 0.5)	(0.5, 1.8, 1.5)	(1.5, 1.9, 2.5)	(2.2, 2.5, 3.5)
E(T)	0.1450	0.0321	0.1590	0.4416	0.3122
$E(T^2)$	0.0940	0.0245	0.1199	0.2480	0.1173
$E(T^3)$	0.0693	0.0198	0.0956	0.1582	0.0518
$E(T^4)$	0.0548	0.0166	0.0792	0.1107	0.0264
$E(T^5)$	0.0453	0.0143	0.0675	0.0829	0.0152
SD	0.2701	0.1532	0.3076	0.2300	0.1409
Var(T)	0.0729	0.0235	0.0946	0.0529	0.0198
S	3.5752	5.4201	3.0557	18.0660	48.9788
K	4.7213	25.7681	3.8723	2.5986	4.4810

Table 3. GFE table of moments

Table 4. GFE table of quantiles

\boldsymbol{u}	(0.5, 0.5, 0.5)	(1.5, 0.3, 0.1)	(2.5, 0.8, 0.5)	(3.5, 0.9, 0.5)	(1.2, 1.5, 0.5)
0.1	0.2940	0.1767	0.3164	0.3155	1.0156
0.2	0.6160	0.5671	0.5102	0.4854	1.3915
0.3	0.9765	1.1342	0.6891	0.6374	1.7185
0.4	1.3913	1.8795	0.8691	0.7873	2.0403
0.5	1.8852	2.8279	1.0604	0.9441	2.3809
0.6	2.5013	4.0368	1.2746	1.1172	2.7650
0.7	3.3241	5.6268	1.5299	1.3208	3.2306
0.8	4.5538	7.8833	1.8644	1.5837	3.8586
0.9	6.8647	11.7415	2.3969	1.9946	4.9066

2.3. Generalized flexible log-logistic distribution

For a log-logistic parent distribution with cdf and pdf given by $G(t; \delta) = 1 - (1 + t^{\delta})^{-1}$ and $g(t; \delta) = \delta t^{\delta - 1} (1 + t^{\delta})^{-1}$ $(t^{\delta})^{-2}$ for non negative δ , then the cdf and pdf of the generalized flexible log-logistic (GFLLoG) model are respectively given by

$$
F_{GFLLoG}(t; \phi, \psi, \delta) = \left[1 - (1 + t^{\delta})^{-\phi(1 - (1 + t^{\delta})^{-1})}\right]^{\psi}, \qquad (16)
$$

and

$$
f_{GFLLoG}(t; \phi, \psi, \delta) = \psi \phi \delta t^{\delta - 1} (1 + t^{\delta})^{-2} (1 + t^{\delta})^{-\phi (1 - (1 + t^{\delta})^{-1})}
$$

$$
\times \left[\frac{1 - (1 + t^{\delta})^{-1}}{(1 + t^{\delta})^{-1}} - \log((1 + t^{\delta})^{-1}) \right]
$$

$$
\times \left[1 - (1 + t^{\delta})^{-\phi (1 - (1 + t^{\delta})^{-1})} \right]^{\psi - 1} .
$$
 (17)

Consequently, the i^{th} moment and the quantile funtion of the GFLLoG distribution are given by

$$
E(T^{i}) = \int_{0}^{\infty} t^{i} f_{GFLLoG}(t; \phi, \psi, \delta) dt,
$$
\n(18)

Figure 3. GFLLoG distribution pdf and hrf plots

and

$$
\phi(1 - (1 + t^{\delta})^{-1}) \log((1 + t^{\delta})^{-1}) = \log(1 - u^{1/\psi}),\tag{19}
$$

respectively. For some selected values of Ω , moments associated with equation [\(18\)](#page-4-2) and quantiles associated with [\(19\)](#page-5-1) are given in Tables [5](#page-5-2) and [6,](#page-5-3) respectively.

	(0.5, 0.5, 1.5)	(1.5, 0.3, 1.1)	(2.5, 0.8, 0.5)	(3.5, 0.9, 1.5)	(1.2, 1.5, 1.5)
E(T)	0.2042	0.2293	0.1848	0.4218	0.1445
$E(T^2)$	0.1327	0.1247	0.1000	0.2824	0.1122
$E(T^3)$	0.0971	0.0838	0.0675	0.2069	0.0911
$E(T^4)$	0.0761	0.0625	0.0507	0.1610	0.0764
$E(T^5)$	0.0624	0.0497	0.0405	0.1308	0.0656
SD	0.3017	0.2686	0.2566	0.3232	0.3022
Var(T)	0.0910	0.0722	0.0658	0.1044	0.0913
S	3.6111	5.3177	4.7576	6.0865	3.0514
K	2.9937	3.2199	4.1033	1.6905	4.3738

Table 5. GFLLoG distribution table of moments

Table 6. GFLLoG distribution table of quantiles

\boldsymbol{u}	(0.5, 0.5, 1.5)	(1.5, 0.3, 1.1)	(2.5, 0.8, 0.5)	(3.5, 0.9, 1.5)	(1.2, 1.5, 1.5)
0.1	0.2927	0.0257	0.0294	0.3079	0.7593
0.2	0.5067	0.0755	0.0844	0.4225	1.0035
0.3	0.7345	0.1456	0.1692	0.5203	1.2284
0.4	1.0034	0.2386	0.2962	0.6150	1.4653
0.5	1.3489	0.3618	0.4890	0.7139	1.7366
0.6	1.8384	0.5299	0.7946	0.8242	2.0730
0.7	2.6321	0.7749	1.3199	0.9566	2.5321
0.8	4.2451	1.1800	2.3719	1.1339	3.2595
0.9	9.6441	2.0778	5.3593	1.4305	4.8335

2.4. Generalized flexible Topp-Leone distribution

Let the parent distribution be the Topp-Leone distribution with $G(t; \delta) = t^{\delta}(2-t)^{\delta}$ and $g(t; \delta) = 2\delta t^{\delta-1}(1-t)^{\delta}$ $t(2-t)^{\delta-1}$ where δ is non negative, then the cdf and pdf of the generalized flexible Topp-Leone (GFTL) model are respectively given by

$$
F_{GFTL}(t; \phi, \psi, \delta) = \left[1 - (1 - t^{\delta}(2 - t)^{\delta})^{\phi t^{\delta}(2 - t)^{\delta}}\right]^{\psi},\tag{20}
$$

and

$$
f_{GFTL}(t; \phi, \psi, \delta) = \psi \phi 2 \delta t^{\delta - 1} (1 - t) (2 - t)^{\delta - 1} (1 - t^{\delta} (2 - t)^{\delta})^{\phi t^{\delta} (2 - t)^{\delta}}
$$

$$
\times \left[\frac{t^{\delta} (2 - t)^{\delta}}{1 - t^{\delta} (2 - t)^{\delta}} - \log(1 - t^{\delta} (2 - t)^{\delta}) \right]
$$

$$
\times \left[1 - (1 - t^{\delta} (2 - t)^{\delta})^{\phi t^{\delta} (2 - t)^{\delta}} \right]^{\psi - 1} .
$$
 (21)

Figure 4. GFTL distribution pdf and hrf plots

The GFTL distribution has its i^{th} moment and quantile function as

$$
E(T^{i}) = \int_{0}^{\infty} t^{i} f_{GFTL}(t; \phi, \psi, \delta) dt,
$$
\n(22)

and

$$
\phi t^{\delta} (2-t)^{\delta} \log(1-t^{\delta} (2-t)^{\delta}) = \log(1-u^{1/\psi}), \tag{23}
$$

respectively. Tables [7](#page-7-1) and [8](#page-7-2) show, respectively, moments associated with equation [\(22\)](#page-6-1) and quantiles associated with Equation [\(23\)](#page-6-2) for some selected values of Ω .

2.5. Generalized flexible Pareto distribution

Let the parent distribution be Pareto with cdf and pdf given by $G(t; \delta) = 1 - 1/t^{\delta}$ and $g(t; \delta) = \delta/x^{\delta+1}$ for non negative δ , then the cdf and pdf of the generalized flexible Pareto (GFP) model are respectively given by

$$
F_{GFP}(t; \phi, \psi, \delta) = \left[1 - \left(1/t^{\delta}\right)^{\phi(1-1/t^{\delta})}\right]^{\psi},\tag{24}
$$

	(0.5, 0.5, 0.5)	(0.5, 1.3, 0.5)	(0.5, 1.8, 1.5)	(1.5, 1.9, 2.5)	(2.2, 2.5, 3.5)
E(T)	0.2962	0.5111	0.7399	0.6237	0.6514
$E(T^2)$	0.1654	0.3344	0.5778	0.4044	0.4328
$E(T^3)$	0.1132	0.2463	0.4695	0.2712	0.2929
$E(T^4)$	0.0856	0.1943	0.3931	0.1874	0.2017
$E(T^5)$	0.0688	0.1602	0.3369	0.1330	0.1412
SD	0.2787	0.2705	0.1743	0.1239	0.0921
Var(T)	0.0777	0.0732	0.0304	0.0154	0.0085
S	6.5438	12.9317	53.2120	153.6846	378.6294
K	2.5739	1.8980	2.6386	2.7430	2.8940

Table 7. GFTL distribution table of moments

Table 8. GFTL distribution table of quantiles

\boldsymbol{u}	(0.5, 0.5, 0.5)	(0.5, 1.3, 0.5)	(0.5, 1.8, 1.5)	(1.5, 1.9, 2.5)	(2.2, 2.5, 3.5)
0.1	0.0094	0.1429	0.4922	0.4618	0.5323
0.2	0.0358	0.2408	0.5846	0.5173	0.5735
0.3	0.0776	0.3305	0.6528	0.5578	0.6032
0.4	0.1347	0.4178	0.7106	0.5925	0.6285
0.5	0.2079	0.5054	0.7630	0.6250	0.6521
0.6	0.2995	0.5953	0.8124	0.6574	0.6757
0.7	0.4139	0.6890	0.8602	0.6918	0.7008
0.8	0.5588	0.7876	0.9073	0.7315	0.7300
0.9	0.7479	0.8917	0.9540	0.7849	0.7699

and

$$
f_{GFP}(t; \phi, \psi, \delta) = \psi \phi \delta t^{-(\delta+1)} (1/t^{\delta})^{\phi(1-1/t^{\delta})} \left[\frac{1-1/t^{\delta}}{1/t^{\delta}} - \log(1/t^{\delta}) \right]
$$

$$
\times \left[1 - (1/t^{\delta})^{\phi(1/t^{\delta})} \right]^{\psi-1} . \tag{25}
$$

Figure 5. GFP distribution pdf and hrf plots

The i^{th} moment and the quantile function of the GFP distribution are given by

$$
E(T^{i}) = \int_{0}^{\infty} t^{i} f_{GFP}(t; \phi, \psi, \delta) dt,
$$
\n(26)

and

$$
\phi(1 - 1/t^{\delta})\log(1/t^{\delta}) = \log(1 - u^{1/\psi}),\tag{27}
$$

respectively. Moments associated with Equation [\(26\)](#page-8-0) and quantiles associated with Equation [\(27\)](#page-8-1) for some selected values of Ω are given in Tables [9](#page-8-2) and [10,](#page-8-3) respectively.

Table 9. GFP distribution table of moments

Table 10. GFP distribution table of quantiles

2.6. Generalized flexible half-logistic distribution

For a half-logistic parent distribution with cdf and pdf given by $G(t; \delta) = \frac{1-e^{-\delta t}}{1+e^{-\delta t}}$ and $g(t; \delta) = \frac{2\delta e^{-\delta t}}{(1+e^{-\delta t})^2}$ for non negative δ , then the cdf and pdf of the generalized flexible half-logistic (GFHL) model are respectively given by

$$
F_{GFHL}(t; \phi, \psi, \delta) = \left[1 - \left(1 - \frac{1 - e^{-\delta t}}{1 + e^{-\delta t}}\right)^{\phi} \right]^{\psi}, \tag{28}
$$

and

$$
f_{GFHL}(t; \phi, \psi, \delta) = \frac{2\psi\phi\delta e^{-\delta t}}{(1 + e^{-\delta t})^2} \left[\frac{\frac{1 - e^{-\delta t}}{1 + e^{-\delta t}}}{1 - \frac{1 - e^{-\delta t}}{1 + e^{-\delta t}}} - \log\left(1 - \frac{1 - e^{-\delta t}}{1 + e^{-\delta t}}\right) \right]
$$

$$
\times \left(1 - \frac{1 - e^{-\delta t}}{1 + e^{-\delta t}}\right)^{\phi} \frac{\frac{1 - e^{-\delta t}}{1 + e^{-\delta t}}}{1 - \left(1 - \frac{1 - e^{-\delta t}}{1 + e^{-\delta t}}\right)^{\phi} \frac{1 - e^{-\delta t}}{1 + e^{-\delta t}}}
$$
 (29)

Figure 6. GFHL distribution pdf and hrf plots

The GFHL distribution has its i^{th} moment and quantile function as

$$
E(T^{i}) = \int_{0}^{\infty} t^{i} f_{GFHL}(t; \phi, \psi, \delta) dt,
$$
\n(30)

and

$$
\phi \frac{1 - e^{-\delta t}}{1 + e^{-\delta t}} \log \left(1 - \frac{1 - e^{-\delta t}}{1 + e^{-\delta t}} \right) = \log(1 - u^{1/\psi}),\tag{31}
$$

respectively. For some selected values of Ω , moments and quantiles of the GFHL model are given in Tables [11](#page-10-0) and [12,](#page-10-1) respectively.

3. Estimation

Let $t_1, t_2, t_3, ..., t_i$ be values of a random sample of size i from the GFG family. The log-likelihood function $log(L(\mathbf{\Omega})) = \ell(\mathbf{\Omega})$, for $\mathbf{\Omega} = (\phi, \psi, \varphi)$ of the GFG family is given by

$$
\ell(\Omega) = i \log(\psi \phi) + \sum_{j=1}^{i} \log(g(t; \varphi)) + \phi \sum_{j=1}^{i} G(t; \varphi) \log(\bar{G}(t; \varphi)) \n+ \sum_{j=1}^{i} \log \left[\frac{G(t; \varphi)}{\bar{G}(t; \varphi)} - \log(\bar{G}(t; \varphi)) \right] + (\psi - 1) \sum_{j=1}^{i} \log \left[1 - \bar{G}(t; \varphi)^{\phi G(t; \varphi)} \right].
$$
\n(32)

	(0.5, 0.5, 0.5)	(0.5, 1.3, 0.5)	(0.5, 1.8, 1.5)	(1.5, 1.9, 2.5)	(2.2, 2.5, 3.5)
E(T)	0.0928	0.0089	0.0741	0.5110	0.4440
$E(T^2)$	0.0621	0.0070	0.0596	0.3465	0.2240
$E(T^3)$	0.0466	0.0057	0.0497	0.2528	0.1254
$E(T^4)$	0.0373	0.0049	0.0425	0.1947	0.0767
$E(T^5)$	0.0311	0.0042	0.0372	0.1561	0.0506
SD	0.2312	0.0831	0.2325	0.2922	0.1639
Var(T)	0.0535	0.0069	0.0541	0.0854	0.0269
S	3.7673	9.9732	3.7714	9.7774	50.2795
K	8.0509	98.2922	10.1590	2.2314	3.3319

Table 11. GFHL distribution table of moments

Table 12. GFHL distribution table of quantiles

\boldsymbol{u}	(0.5, 0.5, 0.5)	(1.5, 0.3, 0.1)	(2.5, 0.8, 0.5)	(3.5, 0.9, 0.5)	(1.2, 1.5, 0.5)
0.1	0.5503	0.3504	0.5896	0.5880	1.6859
0.2	1.0863	1.1037	0.9163	0.8756	2.2042
0.3	1.6298	2.1526	1.2007	1.1200	2.6290
0.4	2.2039	3.4609	1.4729	1.3507	3.0290
0.5	2.8382	5.0299	1.7498	1.5828	3.4374
0.6°	3.5786	6.9046	2.0469	1.8298	3.8831
0.7	4.5111	9.2058	2.3865	2.1094	4.4074
0.8	5.8348	12.2361	2.8123	2.4563	5.0941
0.9	8.2185	16.9942	3.4562	2.9731	6.2050

The log-likelihood in [\(32\)](#page-9-1) has its score functions as

$$
\begin{array}{rcl}\n\frac{\partial \ell(\mathbf{\Omega})}{\partial \phi} & = & \frac{i}{\phi} + \sum_{j=1}^{i} G(t; \varphi) \log(\bar{G}(t; \varphi)) - (\psi - 1) \sum_{j=1}^{i} \frac{\phi g(t; \varphi) \left[\frac{G(t; \varphi)}{\bar{G}(t; \varphi)} - \log(\bar{G}(t; \varphi)) \right]}{\bar{G}(t; \varphi)^{-G(t; \varphi)} \left[1 - \bar{G}(t; \varphi)^{\phi G(t; \varphi)} \right]}, \\
\frac{\partial \ell(\mathbf{\Omega})}{\partial \psi} & = & \frac{i}{\psi} + \sum_{j=1}^{i} \log \left[1 - \bar{G}(t; \varphi)^{\phi G(t; \varphi)} \right],\n\end{array}
$$

and

$$
\frac{\partial \ell(\mathbf{\Omega})}{\partial \varphi} = \sum_{j=1}^{i} \partial \left(\log(g(t; \varphi)) \right) / \partial \varphi_k + (\psi - 1) \sum_{j=1}^{i} \partial \left(\log \left[1 - \bar{G}(t; \varphi)^{\phi G(t; \varphi)} \right] \right) / \partial \varphi_k + \sum_{j=1}^{i} \partial \left(\log \left[\frac{G(t; \varphi)}{\bar{G}(t; \varphi)} - \log(\bar{G}(t; \varphi)) \right] \right) / \partial \varphi_k + \phi \sum_{j=1}^{i} \partial \left(G(t; \varphi) \log(\bar{G}(t; \varphi)) \right) / \partial \varphi_k.
$$

The score functions are not linear in the parameters. Hence iterative methods are required to solve them [\[16,](#page-16-12) [18\]](#page-16-13).

4. Simulation Study

In this section, we evaluate the effectiveness of the maximum likelihood estimators (MLEs) in the proposed model using the GFLLoG distribution. The algorithm of the simulation process is as follows:

1242 A NOVEL TECHNIQUE FOR GENERATING FAMILIES OF CONTINUOUS DISTRIBUTIONS

- i) Select initial parameter values for the GFLLoG distribution;
- ii) Generate *n* random values from a uniform distribution with pdf $f(x) = 1$; $0 < x < 1$;
- iii) Use Equation [19](#page-5-1) to compute n values of the GFLLoG distribution defined in step i);
- iv) Repeat the above steps 2000 times;

The performance of the estimators is evaluated using the Average Bias (ABIAS) and Root Mean Square Error (RMSE). These metrics are calculated based on 2000 samples of each selected sample size. The simulation results, presented in Tables [13,](#page-11-0) [14](#page-12-0) and [15,](#page-12-1) showcase the performance under selected initial parameter values. It is observed that both ABIAS and RMSE decrease as the sample size n increases, indicating the consistency of the MLEs. These findings demonstrate that the MLEs provide reliable results when estimating model parameters in the GFG distribution.

			(1.0, 1.0, 1.0)			(0.5, 0.5, 0.5)	
Parameter	Sample Size	MLE	RMSE	ABias	MLE	RMSE	ABias
ϕ	75	2.3003	3.0880	1.3003	1.1171	2.0709	0.6171
	100	2.1599	2.8389	1.1599	0.9690	1.6822	0.4690
	200	1.6116	1.8416	0.6116	0.6189	0.6375	0.1189
	400	1.3074	1.1639	0.3074	0.5478	0.4085	0.0478
	500	1.1861	0.8389	0.1861	0.5264	0.1390	0.0264
	800	1.0986	0.5287	0.0986	0.5244	0.1115	0.0244
ψ	75	9.9784	95.6833	8.9784	2.8951	33.4374	2.3951
	100	6.9022	47.2207	5.9022	1.6000	10.8882	1.1000
	200	2.5199	9.2566	1.5199	0.6229	0.9254	0.1229
	400	1.5297	4.1140	0.5297	0.5823	1.6032	0.0823
	500	1.2724	1.9318	0.2724	0.5219	0.0834	0.0219
	800	1.1134	0.8077	0.1134	0.5224	0.0700	0.0224
δ	75	0.9447	0.4882	-0.0553	0.5024	0.2032	0.0024
	100	0.9324	0.4482	-0.0676	0.4938	0.1759	-0.0062
	200	0.9519	0.3257	-0.0481	0.5056	0.1116	0.0056
	400	0.9696	0.2399	-0.0304	0.5072	0.0755	0.0072
	500	0.9821	0.2016	-0.0179	0.5104	0.0682	0.0104
	800	0.9901	0.1582	-0.0099	0.5036	0.0555	0.0036

Table 13. Parameter estimation from the GFLLoG distribution Results 1

5. Data Analysis

We consider patients receiving an analgesic data [\[3\]](#page-16-14), and reliability data [\[4\]](#page-16-15). For each dataset, the MLEs for the distributions are obtained. The GFLLoG is compared to non-nested models of different parameters namely the exponentiated generalized XLindley (EGXL) [\[17\]](#page-16-16), exponentiated generalized logarithmic (EGEL) [\[20\]](#page-16-17), exponentiated power lindley poisson (EPLP) [\[21\]](#page-17-3), generalized Gompertz-poisson (GGP) [\[26\]](#page-17-4) and the exponentiated Burr-XII poisson (EBXIIP) [\[7\]](#page-16-18). We also compared the GFLLoG with some parent distributions namely the Burr-XII [\[27\]](#page-17-5), Gambel [\[5\]](#page-16-19), Weibull distribution [\[23\]](#page-17-6). Model selection criteria such as the $-2 \log L$, AIC, AICC, BIC, W^* , A^* , K-S statistic and its associated p-value (see [\[17\]](#page-16-16) for more details on the GoF) were employed. For the selected model, plots representing the fitted density, probability plot, Kaplain Meier (KM), estimated cdf (ECDF), estimated hrf, and total time on test (TTT) are displayed.

5.1. Patients receiving an analgesic

The data collection contains information on patients' lifetime alleviation times (measured in minutes) after taking an analgesic as reported [\[3\]](#page-16-14).

			(1.0, 1.5, 1.1)			(2.5, 0.5, 2.5)	
Parameter	Sample Size	Mean	RMSE	Bias	Mean	RMSE	A.Bias
ϕ	75	2.3204	3.3551	1.3204	3.4363	2.1962	0.9363
	100	2.0541	2.8295	1.0541	3.1777	1.8011	0.6777
	200	1.5286	1.8124	0.5286	2.8190	1.1928	0.3190
	400	1.2216	1.0461	0.2216	2.6455	0.8632	0.1455
	500	1.1365	0.7890	0.1365	2.6196	0.7726	0.1196
	800	1.0664	0.4428	0.0664	2.5703	0.6189	0.0703
ψ	75	17.8273	107.6070	16.3273	1.0344	2.2969	0.5344
	100	10.6293	69.0287	9.1293	0.8160	1.1024	0.3160
	200	4.4779	37.2887	2.9779	0.6188	0.3679	0.1188
	400	2.1849	5.4492	0.6849	0.5536	0.2208	0.0536
	500	2.0384	10.9684	0.5384	0.5447	0.1945	0.0447
	800	1.6337	1.5989	0.1337	0.5266	0.1475	0.0266
δ	75	1.0616	0.5446	-0.0384	2.4731	1.1849	-0.0269
	100	1.0517	0.4755	-0.0483	2.5084	1.0941	0.0084
	200	1.0646	0.3362	-0.0354	2.5579	0.8923	0.0579
	400	1.0835	0.2376	-0.0165	2.5656	0.6859	0.0656
	500	1.0912	0.2009	-0.0088	2.5523	0.6183	0.0523
	800	1.0952	0.1551	-0.0048	2.5409	0.4905	0.0409

Table 14. Parameter estimation from the GFLLoG distribution Results 2

Table 15. Parameter estimation from the GFLLoG distribution Results 3

			(1.5, 1.5, 1.5)			(2.5, 0.5, 1.5)	
Parameter	Sample Size	Mean	RMSE	Bias	Mean	RMSE	A.Bias
φ	75	3.1544	3.6789	1.6544	3.4104	2.1566	0.9104
	100	3.1081	3.4807	1.6081	3.2119	1.8399	0.7119
	200	2.5387	2.4622	1.0387	2.8446	1.2080	0.3446
	400	2.1628	1.7562	0.6628	2.6578	0.8685	0.1578
	500	2.0655	1.5949	0.5655	2.6193	0.7793	0.1193
	800	1.8827	1.1965	0.3827	2.5686	0.6267	0.0686
ψ	75	22.3013	184.7170	20.8013	0.9987	1.8218	0.4987
	100	14.7299	104.0332	13.2299	0.8373	1.1373	0.3373
	200	5.0081	21.4077	3.5081	0.6273	0.3799	0.1273
	400	2.8911	5.0725	1.3911	0.5570	0.2258	0.0570
	500	2.6303	4.0689	1.1303	0.5451	0.1964	0.0451
	800	2.1260	2.2063	0.6260	0.5265	0.1495	0.0265
δ	75	1.4287	0.8087	-0.0713	1.4891	0.7188	-0.0109
	100	1.3788	0.7435	-0.1212	1.5002	0.6790	0.0002
	200	1.3942	0.5841	-0.1058	1.5252	0.5319	0.0252
	400	1.4165	0.4700	-0.0835	1.5346	0.4107	0.0346
	500	1.4250	0.4315	-0.0750	1.5333	0.3784	0.0333
	800	1.4429	0.3619	-0.0571	1.5275	0.3017	0.0275

Tables [16](#page-13-0) and [17](#page-13-1) show maximum likelihood estimates (standard errors in parenthesis) and GoF statistics of the fitted models. In table [17,](#page-13-1) it is clear that the GFLLoG consistently has the lowest GoF statistics values and a corresponding high value of the K-S statistic as compared to comparative models of different number of parameters. It is evident that the GFLLoG distribution is the best fit for data set on patients receiving an analgesic.

Estimates									
Model	Φ	ψ	δ	ϵ					
GFLLoG	1.0377	4.5151	3.7132						
	(1.1912)	(2.9350)	(3.2739)						
EGXL	0.1937	36.6798	11.6052						
	(1.0717)	(25.3081)	(63.5798)						
EGEL	2.5195×10^{-8}	4.4828	44.2770	0.5239					
	(0.0381)	(37.1390)	(34.4620)	(4.3409)					
EPLP	1.1407×10^{-8}	1.3382	1.3909	6.3251					
	(0.0442)	(0.3369)	(0.6274)	(5.1764)					
GGP	8.0847	5.8056	2.3723	3.0970×10^{-8}					
	(5.0881)	(4.9606)	(0.4817)	(0.0584)					
EBXIIP	4.0755	0.6325	40.7090	5.9194×10^{-6}					
	(0.7475)	(0.0960)	(0.0013)	(0.5487)					
BXII		157.1300	0.0108						
		(1.6598×10^{-7})	(0.0024)						
Gambel		0.4386	1.6189						
		(0.0811)	(0.1023)						
Weibull		0.1215	2.7870						
		(0.0562)	(0.4273)						

Table 16. Model's parameter estimates

Table 17. Model's GoF statistics

GoF Statistics								
Model	$-2\log L$	AIC	AICC	BIC	W^*	A^*	$K-S$	p-value
GFLLoG	30.8524	36.8524	38.3524	39.8396	0.0259	0.1487	0.0956	0.9931
EGXL	32.5239	38.5239	40.0239	41.5111	0.0542	0.3185	0.1343	0.8631
EGEL	32.5951	40.5951	43.2617	44.5780	0.0515	0.3020	0.1282	0.8974
EPLP	34.7641	42.7641	45.4307	46.7470	0.0912	0.5411	0.1625	0.6657
GGP	33.4128	41.4127	44.0793	45.3956	0.0643	0.3822	0.1199	0.9358
EBXIIP	30.8148	38.8148	41.4814	42.7977	0.0281	0.1597	0.0997	0.9886
BXII	42.4142	46.4142	47.1201	48.4057	0.0381	0.2193	0.2850	0.0775
Gambel	32.6661	36.6661	37.3720	38.6575	0.0554	0.3261	0.1340	0.8647
Weibull	41.1728	45.1728	45.8786	47.1642	0.1857	1.0928	0.1849	0.5005

5.2. Reliability data

The reliability data considered here consisting of 20 mechanical components failure times as reported [\[4\]](#page-16-15).

Table [18](#page-15-0) shows some parameter estimates and standard errors in parenthesis of the fitted models. In table [19,](#page-15-1) it is evident that the GFLLoG distribution is the best fit for reliability data since the GFLLoG has the lowest GoF statistics values and the highest K-S statistic value as compared to comparative models presented in the table.

Figure 7. Fitted density, probability plot and KM survival plot of the GFLLoG distribution for patients receiving an analgesic data

Figure 8. ECDF, estimated hrf plot and TTT of the GFLLoG distribution for patients receiving an analgesic data

Figure 9. Fitted density, probability plot and KM survival plot of the GFLLoG distribution for reliability data

6. Concluding Remarks

We introduce an original flexible generalized family for univariate distributions called the generalized flexible-G family, which has the flexible-G family as its sub-model and was not created using any well-known parent model. The hazard rate function of the generalized flexible-G family can take on a variety of very flexible shapes, such as bathtub, bathtub followed by upside-down bathtub, upside-down bathtub, constant, increasing

Table 18. Model's parameter estimates

Table 19. Model's GoF statistics

GoF Statistics								
Model	$-2\log L$	AIC	AICC	BIC	W^*	A^\ast	$K-S$	p-value
GFLLoG	-76.0595	-70.0595	-68.5595	-67.0723	0.0563	0.4387	0.1245	0.9159
EGXL	-65.9524	-59.9524	-58.4524	-56.9652	0.1758	1.2511	0.1603	0.6826
EGEL	-65.9528	-57.9528	-55.2861	-53.9699	0.1756	1.2504	0.1601	0.6840
EPLP	-34.8670	-26.8668	-24.2001	-22.8839	0.2350	1.5937	0.4585	0.0004
GGP	-58.1901	-50.1899	-47.5232	-46.2070	0.2490	1.6718	0.2988	0.0561
EBXIIP	-78.0870	-70.0870	-67.4203	-66.1041	0.0421	0.2901	0.1261	0.9080
BXII	-53.8717	-49.8717	-49.1658	-47.8802	0.3723	2.3262	0.2641	0.1227
Gambel	-65.1988	-61.1988	-60.4930	-59.2074	0.1798	1.2751	0.1602	0.6832
Weibull	-52.8456	-48.8456	-48.1397	-46.8541	0.3970	2.4519	0.2641	0.1227

and decreasing. Because of its appealing flexibility, the generalized flexible-G family's hazard rate function can be used to non-monotonic empirical hazard behaviors, which are more likely to occur in or be seen in real-world scenarios. We used the log-logistic as our baseline model for the presented simulation study and data analysis. Our generalized flexible log-logistic fit the two real-life datasets better than compared models presented in this article. This technique was limited to the univariate case, in the same spirit, more work can be done including the multivariate extension, truncation, censoring schemes and regression.

Figure 10. ECDF, estimated hrf plot and TTT of the GFLLoG distribution for reliability data

Standard model selection criteria like AIC, BIC, or cross-validation might be used in choosing between the various special cases within the generalized flexible-G family.

REFERENCES

- 1. Ayman Alzaatreh, Carl Lee, and Felix Famoye. A new method for generating families of continuous distributions. *Metron*, 71(1):63– 79, 2013.
- 2. Muhammad Aslam, Zawar Hussain, and Zahid Asghar. Cubic transmuted-g family of distributions and its properties. *Stochastics and Quality Control*, 33(2):103–112, 2018.
- 3. Rana Ali Bakoban and Ashwaq Mohammad Al-Shehri. A new generalization of the generalized inverse rayleigh distribution with applications. *Symmetry*, 13(4):711, 2021.
- 4. Marcelo Bourguignon, Helton Saulo, and Rodrigo Nobre Fernandez. A new pareto-type distribution with applications in reliability and income data. *Physica A: Statistical Mechanics and its Applications*, 457:166–175, 2016.
- 5. Kahadawala Cooray. Generalized gumbel distribution. *Journal of Applied Statistics*, 37(1):171–179, 2010.
- 6. Gauss M Cordeiro, Edwin MM Ortega, and Daniel CC da Cunha. The exponentiated generalized class of distributions. *Journal of data science*, 11(1):1–27, 2013.
- 7. Ronaldo V da Silva, Frank Gomes-Silva, Manoel Wallace A Ramos, and Gauss M Cordeiro. The exponentiated burr xii poisson distribution with application to lifetime data. *International Journal of Statistics and Probability*, 4(4):112, 2015.
- 8. Fatehi Yahya Eissa and Chhaya Dhanraj Sonar. Generalization of power lindley distribution: Properties and applications. *Statistics, Optimization & Information Computing*, 2024.
- 9. Mahmoud Eltehiwy. The odd log-logistic power inverse lindley distribution: Model, properties and applications. *Statistics, Optimization & Information Computing*, 12(1):91–108, 2024.
- 10. Masoud Ganji, Hossein Bevrani, and Nasrin Hami Golzar. A new method for generating continuous bivariate distribution families. *Journal of The Iranian Statistical Society*, 17(1):109–129, 2022.
- 11. Ramesh C Gupta, Pushpa L Gupta, and Rameshwar D Gupta. Modeling failure time data by lehman alternatives. *Communications in Statistics-Theory and methods*, 27(4):887–904, 1998.
- 12. MA Lone, IH Dar, and TR Jan. A new method for generating distributions with an application to weibull distribution. *Reliability: Theory & Applications*, 17(1 (67)):223–239, 2022.
- 13. Abbas Mahdavi and Debasis Kundu. A new method for generating distributions with an application to exponential distribution. *Communications in Statistics-Theory and Methods*, 46(13):6543–6557, 2017.
- 14. Albert W Marshall and Ingram Olkin. A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. *Biometrika*, 84(3):641–652, 1997.
- 15. Regent Retrospect Musekwa, Lesego Gabaitiri, and Boikanyo Makubate. A new technique to generate families of continuous distributions. *Revista Colombiana de Estad´ıstica*, 47(2):329–354, 2024.
- 16. Regent Retrospect Musekwa and Boikanyo Makubate. Statistical analysis of saudi arabia and uk covid-19 data using a new generalized distribution. *Scientific African*, 22:e01958, 2023.
- 17. Regent Retrospect Musekwa and Boikanyo Makubate. A flexible generalized xlindley distribution with application to engineering. *Scientific African*, page e02192, 2024.
- 18. Violet Zivai Nyamajiwa, Regent Retrospect Musekwa, and Boikanyo Makubate. Application of the new extended topp-leone distribution to complete and censored data. *Revista Colombiana de Estad´ıstica*, 47(1):37–65, 2024.
- 19. Broderick Oluyede, Bakang Percy Tlhaloganyang, and Whatmore Sengweni. The topp-leone odd burr xg family of distributions: Properties and applications. *Statistics, Optimization & Information Computing*, 12(1):109–132, 2024.
- 20. Broderick O Oluyede, Baitshephi Mashabe, A Fagbamigbe, Boikanyo Makubate, and Divine Wanduku. The exponentiated generalized power series: Family of distributions: theory, properties and applications. *Heliyon*, 6(8), 2020.

1248 A NOVEL TECHNIQUE FOR GENERATING FAMILIES OF CONTINUOUS DISTRIBUTIONS

- 21. Mavis Pararai, Gayan Warahena-Liyanage, and Broderick O Oluyede. Exponentiated power lindley–poisson distribution: Properties and applications. *Communications in Statistics-Theory and Methods*, 46(10):4726–4755, 2017.
- 22. Md Mahabubur Rahman, Bander Al-Zahrani, and Muhammad Qaiser Shahbaz. A general transmuted family of distributions. *Pakistan Journal of Statistics and Operation Research*, pages 451–469, 2018.
- 23. Horst Rinne. *The Weibull distribution: a handbook*. CRC press, 2008.
- 24. Subhradev Sen, Morad Alizadeh, Mohamed Aboraya, M Masoom Ali, Haitham M Yousof, and Mohamed Ibrahim. On truncated versions of xgamma distribution: Various estimation methods and statistical modelling. *Statistics, Optimization & Information Computing*, 12(4):943–961, 2024.
- 25. MH Tahir, M Adnan Hussain, and Gauss M Cordeiro. A new flexible generalized family for constructing many families of distributions. *Journal of Applied Statistics*, 49(7):1615–1635, 2022.
- 26. Saeed Tahmasebi and Ali Akbar Jafari. Generalized gompertz-power series distributions. *Hacettepe Journal of Mathematics and Statistics*, 45(5):1579–1604, 2015.
- 27. William J Zimmer, J Bert Keats, and FK Wang. The burr xii distribution in reliability analysis. *Journal of quality technology*, 30(4):386–394, 1998.