Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 28 (2025), No. 3, pp. 425–443
DOI : 10.47974/JSMS-1233

Another extension of the Gompertz distribution : Properties, applications and regression

Selasi Kwaku Ocloo *
Department of Engineering
Faculty of Engineering and Computer Science
Ashesi University
Berekuso
Ghana

Suleman Nasiru †
Department of Statistics and Actuarial Science
School of Mathematical Sciences
C. K. Tedam University of Technology and Applied Sciences
Navrongo
Ghana

Benjamin Odoi §
Lewis Brew ‡
Department of Mathematical Sciences
Faculty of Computing and Mathematical Sciences
University of Mines and Technology
Tarkwa
Ghana

* *ORCID-ID:* 0000-0001-6137-4717
† *ORCID-ID:* 0000-0001-6652-4251
§ *ORCID-ID:* 0000-0001-6759-0744
‡ *ORCID-ID:* 0000-0002-4820-7241

* *E-mail:* selasi.ocloo@ashesi.edu.gh (Corresponding Author)
† *E-mail:* snasiru@cktutas.edu.gh; sulemanstat@gmail.com
§ *E-mail:* bodoi@umat.edu.gh
‡ *E-mail:* lbrew@umat.edu.gh

Abstract

Modelling heterogeneous data sets is an emerging research area that is growing steadily. This paper introduces the Harmonic Mixture Gompertz distribution, a new mixture model derived from the weighted harmonic mean of the survival functions of two Gompertz distributions. Some statistical properties are obtained. A simulation study is conducted to evaluate the performance of some estimation techniques used. Two-lifetime data sets were applied to the proposed distribution to ascertain its versatility while comparing it with other modifications of the Gompertz distribution. A regression model from the proposed model is developed using logarithmic link functions.

Subject Classification: *60E05, 62F10, 62PO5.*

Keywords: *Harmonic mixture-G family, Gompertz distribution, Heavy-tailed distribution, Parameter estimation, Monte Carlo simulation, Applications.*

1. Introduction

The goal of developing new types of distributions over time has been to offer greater flexibility in fitting actual data sets than the well-known classical distributions. In recent times, modifying classical distributions require the using what are called families of distributions. These families of distributions serve as the engine block that transforms the given baseline distributions into modified distributions. Using mixture families of distributions or generators is one of the contemporary methods for enhancing the flexibility of classical distributions because of their ability to handle heterogeneous data sets. The Harmonic Mixture-G (HMG) family [1] is used as generator and the Gompertz distribution as baseline distribution for this study. This new family introduces two extra parameters, a scale and a shape parameter which would improve the skewness and kurtosis of the Gompertz distribution. Several modifications of the Gompertz distribution exist in literature. They include: the Zubair Gompertz distribution [2], the new odd Lindley-Gompertz distribution [3], the Kumaraswamy Gompertz distribution [4],the Rayleigh Gompertz [5], the exponentiated generalisedWeibull-Gompertz distribution [6], the Rayleigh Gamma Gompertz distribution [7], the inverse power Gompertz [8], the odd generalised exponential Gompertz distribution [9], the inverted shifted Gompertz distribution [10], the Nadarajah Haghighi Gompertz distribution [11] and the transmuted power Gompertz distribution [12]

In this paper, a new distribution is proposed to address the deficiency in skewness in the Gompertz distribution and improve kurtosis. A distribution that is heavy-tailed and provides better fit than some modifications of the Gompertz distribution. The study adopted some of the estimation methods used by [13]and [14].

The structure of the paper is organized as follows: Section 2 introduces the probability density function (pdf), the cumulative distribution function (cdf), and the hazard rate function (hrf) of the Harmonic Mixture Gompertz (HMGOM) distribution. Sections 3 and 4 outline key statistical properties and estimators for the HMGOM distribution respectively. Section 5 illustrates simulations performed to assess the performance of these estimators. Sections 6 and 7 illustrate the practical application of the model using two lifetime data and the development of a regression model, respectively. Finally, Section 8 concludes the study by summarizing the findings.

2. Development of Harmonic Mixture Gompertz Distribution

The random variable (rv) Y belongs to the HMG family if it's pdf, cdf and hrf are defined respectively as

$$f_{Hm}(y) = g(y)\bar{G}^{\Psi-1}(y)\frac{\Psi(1-\Phi)+\Phi\bar{G}^{\Psi-1}(y)}{[1-\Phi(1-\bar{G}^{\Psi-1}(y))]^2}, \quad y \in R, \quad \Psi \geq 0, \quad 0 \leq \Phi \leq 1. \tag{1}$$

$$F_{Hm}(y) = 1 - \frac{\bar{G}^{\Psi}(y)}{1-\Phi(1-\bar{G}^{\Psi-1}(y))}, \quad y \in R, \quad \Psi \geq 0, \quad 0 \leq \Phi \leq 1. \tag{2}$$

and

$$h_{Hm}(y) = \frac{g(y)}{\bar{G}(y)}\frac{\Psi(1-\Phi)+\Phi\bar{G}^{\Psi-1}(y)}{1-\Phi(1-\bar{G}^{\Psi-1}(y))}, \tag{3}$$

$\bar{G}(y)$ is the SF or complement of the cdf.

The rv Y also follows the Gompertz distribution if it's PDF and survival function (SF) are given respectively as

$$f_G(y;\delta,\Omega) = \Omega e^{\delta y}\exp\left[-\frac{\Omega}{\delta}(e^{\delta y}-1)\right], \quad y > 0, \quad \delta > 0, \quad \Omega > 0. \tag{4}$$

and

$$S_G(y;\delta,\Omega) = \exp\left[-\frac{\Omega}{\delta}(e^{\delta y}-1)\right], \quad y > 0, \quad \delta > 0, \quad \Omega > 0. \tag{5}$$

where $\delta > 0$ and $\Omega > 0$.

Hence, the rv Y follows the HMGOM distribution if it has pdf given by

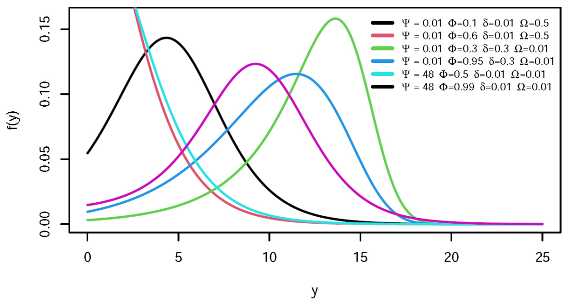

Figure 1

The shapes of the pdf

$$f(y)=\frac{\Omega\Psi(1-\Phi)e^{\delta y}e^{-\frac{\Omega\Psi}{\delta}(e^{\delta y}-1)}+\Omega\Phi e^{\delta y}e^{-\frac{\Omega(2\Psi-1)}{\delta}(e^{\delta y}-1)}}{\left[1-\Phi\left(1-e^{-\frac{\Omega(\Psi-1)}{\delta}(e^{\delta y}-1)}\right)\right]^2}, \tag{6}$$

where $\delta>0,\ \ \Psi>0,\ \ \Omega>0,\ \ y>0$ and $0<\Phi<1$.

Adjusting the parameter values, the densities displayed patterns that were decreasing, left-skewed, or right-skewed as shown in Figure 1.

The SF of the rv $Y\sim HMGOM(\delta,\Omega,\Psi,\Phi)$ is given by

$$S(y)=\frac{e^{-\frac{\Omega\Psi}{\delta}(e^{\delta y}-1)}}{\left[1-\Phi\left(1-e^{-\frac{\Omega(\Psi-1)}{\delta}(e^{\delta y}-1)}\right)\right]}, \quad y>0. \tag{7}$$

The hrf of the rv $Y\sim HMGOM(\delta,\Omega,\Psi,\Phi)$ is given as

$$h(y)=\frac{\Omega\Psi(1-\Phi)e^{\delta y}e^{-\frac{\Omega\Psi}{\delta}(e^{\delta y}-1)}+\Omega\Phi e^{\delta y}e^{-\frac{\Omega(2\Psi-1)}{\delta}(e^{\delta y}-1)}}{e^{-\frac{\Omega\Psi}{\delta}(e^{\delta y}-1)}\left[1-\Phi\left(1-e^{-\frac{\Omega(\Psi-1)}{\delta}(e^{\delta y}-1)}\right)\right]}, \quad y>0. \tag{8}$$

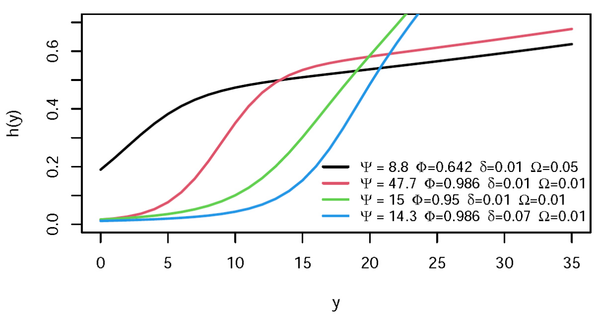

Figure 2

The hrf plots of the HMGOM

Adjusting some parameters, the hazard rate plots revealed increasing and concave patterns as shown in Figure 2.

Lemma 1: *The series expansion of the HMGOM density provided* $\Psi > 1$ *is given as*

$$f(y) = \sum_{a=0}^{\infty}\sum_{b=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} \varpi_{abklm} y^{m} e^{-fly}, \tag{9}$$

where

$$\varpi_{ijklm} = [\Omega\Psi(1-\Phi)\omega_{ijklm} + \Omega\Phi\omega^{*}_{ijklm}],$$

$$\omega_{abklm} = \frac{(-1)^{b+k+l}}{k!m!}(a+1)\binom{a}{b}\binom{l}{k}\Phi^{a}\left(\frac{\Omega}{\delta}(\Psi(b+1)-b)\right)^{k}(\delta(k+1))^{m},$$

$$\omega^{*}_{abklm} = \frac{(-1)^{b+k+l}}{k!m!}(a+1)\binom{a}{b}\binom{l}{k}\Phi^{a}\left(\frac{\Omega}{\delta}(\Psi(b+2)-(b+1))\right)^{k}(\delta(k+1))^{m},$$

$$y > 0, \quad \delta > 0, \quad \Omega > 0, \quad \Psi > 1, \quad 0 < \Phi < 1.$$

Proof: For $\Omega > 0,$ the Taylor expansions for $(1-w)^{-\Omega}$, $(1-w)^{\eta}$ for $|w|<1$ and e^{-t} are $(1-w)^{-\Omega} = \sum_{a=0}^{\infty} \binom{\Omega+a-1}{a} (w)^{a}$, $(1-y)^{\eta} = \sum_{b=0}^{\infty} (-1)^{b} \binom{\eta}{b} (y)^{b}$ and $e^{-t} = \sum_{k=0}^{\infty} \frac{(-1)^{k}}{k!}(t)^{k}$.

Since $0 < e^{-\frac{\Omega(\Psi-1)}{\delta}(e^{\delta y}-1)} < 1$ when $\Psi > 1,$ implies the above lemma could be derived.

3. Statistical Properties of the HMGOM

The quantile function (qf) of the rv $Y \sim HMGOM(\delta,\Omega,\Psi,\Phi)$ can be expressed as

$$(1-u)\left[1-\Phi\left(1-e^{-\frac{\Omega(\Psi-1)}{\delta}(e^{\delta y_u}-1)}\right)\right] - e^{-\frac{\Omega\Psi}{\delta}(e^{\delta y_u}-1)} = 0, \tag{10}$$

where $u \in (0,1)$ and $Q(u) = y_u$. We use numerical approximations to estimate the values of the gf since it does not have a closed-form solution.

The j^{th} non-central moment of the rv $Y \sim HMGOM(\delta,\Omega,\Psi,\Phi)$ is given as

$$\mu'_j = \sum_{a=0}^{\infty}\sum_{b=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\varpi_{abklm}\left(\frac{1}{\delta l}\right)^{j+m+1}\Gamma(j+m+1), \tag{11}$$

where $\Psi > 1$ and $j = 1, 2, \ldots$.

This result can be obtained by the definition

$$\mu'_j = E(Y^j) = \int_0^\infty y^j f(y)\mathrm{d}y. \tag{12}$$

The moment generating function (mgf) of the rv $Y \sim HMGOM(\delta, \Omega, \Psi, \Phi)$ is given by

$$M_y(t) = \sum_{a=0}^{\infty}\sum_{b=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{j=0}^{\infty}\varpi_{abklm}\frac{t^j}{j!}\left(\frac{1}{\delta l}\right)^{j+m+1}\Gamma(j+m+1), \Psi > 1. \tag{13}$$

Applying the definition

$$M_y(t) = E(e^{tY}) = \sum_{j=0}^{\infty}\frac{t^j E(Y^j)}{j!} = \sum_{j=0}^{\infty}\frac{t^j}{j!}\mu'_j, \tag{14}$$

gives the mgf of the HMGOM distribution.

The j^{th} incomplete moment of the rv $Y \sim HMGOM(\delta, \Omega, \Psi, \Phi)$ is given as

$$m_r(z) = \sum_{a=0}^{\infty}\sum_{b=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\varpi_{abklm}\left(\frac{1}{\delta l}\right)^{j+m+1}\gamma(j+m+1, \delta l z), \Psi > 1, \tag{15}$$

where $\gamma(\cdot,\cdot)$, the lower incomplete gamma function, $j = 1, 2, \ldots$.

By applying the definition

$$m_j(z) = E(Y^j \mid Y \le z) = \int_0^z y^j f(y)\mathrm{d}y, \tag{16}$$

the incomplete moment is obtained.

4. Estimation of Parameters

This section explores several methods that could be used to obtain parameter estimates of the HMGOM distribution. Applying these estimation techniques, we aim to determine the most suitable parameter values that best fit the HMGOM distribution in relation to a particular data.

Applying the maximum likelihood estimation (MLE) to the HMGOM distribution, researchers can obtain parameter estimates that are optimal in terms of maximizing the likelihood of the observed data and capturing the underlying characteristics of the distribution. The log-likelihood function of model is

$$l(y,\Psi,\Phi,\delta,\Omega) = n\ln\Omega + \delta\sum_{a=1}^{n} y_a - \frac{\Omega}{\delta}\Psi\sum_{a=1}^{n}(e^{\delta y_a}-1)$$
$$+\sum_{a=1}^{n}\ln[\Psi(1-\Phi)+\Phi e^{-\frac{\Omega(\Psi-1)}{\delta}(e^{\delta y_a}-1)}] - 2\sum_{a=1}^{n}\ln[1-\Phi(1-e^{-\frac{\Omega(\Psi-1)}{\delta}(e^{\delta y_a}-1)})]. \quad (17)$$

The ordinary least squares (O), weighted least squares (W), Cramér-von Mises (C) and Anderson-Darling estimation (A) methods are used to obtain estimates respectively by the minimisation of the functions

$$O(\delta,\Omega,\Psi,\Phi) = \sum_{a=1}^{n}\left[(F(y_{(a)})) - \frac{a}{n+1}\right]^2, \quad (18)$$

$$W(\delta,\Omega,\Psi,\Phi) = \sum_{a=1}^{n}\frac{(n+1)^2(n+2)}{a(n-a+1)}\left[(F(y_{(a)})) - \frac{a}{n+1}\right]^2, \quad (19)$$

$$C(\delta,\Omega,\Psi,\Phi) = \frac{1}{12n} + \sum_{a=1}^{n}\left[(F(y_{(a)})) - \frac{2a-1}{2n}\right]^2, \quad (20)$$

and

$$A(\delta,\Omega,\Psi,\Phi) = -n - \frac{1}{n}\sum_{a=1}^{n}(2a-1)\cdot[(\log F(y_{(a)})) + \log(1-F(y_{(n+1-a)}))]. \quad (21)$$

5. Simulations

The estimators are assessed in this section. We obtain the biases (ab) and mean square errors (mse) for the MLE, the O, the W, the C, and the A using the following formula:

$$ab = \frac{1}{1000}\sum_{a=1}^{N}(\hat{V}_a - V); \quad mse = \frac{1}{1000}\sum_{a=1}^{N}(\hat{V}_a - V)^2.$$

The results are presented in Tables 1 and 2. The ab and mse of the estimators generally decrease as the sample size increases, despite some fluctuations. Table 1 shows the ab and mse for the estimators with the parameters $(\Psi,\Phi,\delta,\Omega)$ = (0.2, 0.6, 0.5, 0.9). The MLE estimator produced the lowest ab and mse, suggesting that MLE is the most efficient estimator.

Table 2 displays the ab and mse for the estimators with parameters $(\Psi,\Phi,\delta,\Omega)$ = (0.3, 0.7, 0.6, 1.0). In this case, MLE generally produced the lowest ab, while O and W had the lowest mse.

Thus, MLE, O, and W can be considered effective estimators for the HMGOM distribution.

Table 1

Results of Simulations for $(\Psi, \Phi, \delta, \Omega) = (0.2, 0.6, 0.5, 0.9)$

Parameter	N	ab					mse				
		MLE	O	W	C	A	MLE	O	W	C	A
Ψ	30	0.1728	0.1205	0.1627	6.4661	8.0753	0.0334	0.0212	0.0299	111.9517	110.6284
	80	0.1318	0.1361	0.1391	4.1971	10.7874	0.0226	0.0255	0.0251	78.0884	176.0832
	200	0.1446	0.1610	0.1564	7.8795	9.9169	0.0260	0.0300	0.0291	104.5699	124.4208
	500	0.1815	0.1210	0.1449	8.1356	10.8938	0.0353	0.0208	0.0262	88.6376	156.9742
	1000	0.1158	0.1229	0.1362	9.2244	13.7704	0.0181	0.0212	0.0263	133.2229	248.2210
Φ	30	1.1881	1.5707	0.2873	5.5736	3.0961	7.3793	12.6514	0.1332	56.9280	20.4026
	80	0.3462	0.4027	0.3402	4.9603	4.2490	0.1539	0.1958	0.1498	78.0884	40.1785
	200	0.2893	0.2389	0.8352	2.8981	2.7631	0.1152	0.1101	3.1619	18.4375	12.4311
	500	0.2234	0.3470	0.3437	2.6610	4.0440	0.0784	0.1693	0.1587	9.8217	35.6852
	1000	0.4123	1.0902	0.3187	2.2845	2.4518	0.1977	6.3851	0.1316	16.7158	13.1933
δ	30	0.4276	0.3338	0.2213	13.5653	14.4995	0.2762	0.1290	0.0896	354.1072	332.2690
	80	0.2649	0.2384	0.2379	13.0908	19.3098	0.0808	0.0814	0.0817	240.0828	393.2078
	200	0.1946	0.2266	0.3252	15.6499	22.3086	0.0498	0.0666	0.1827	259.1609	513.1973
	500	0.1286	0.1755	0.2052	19.5030	28.1683	0.0271	0.0466	0.0597	423.2970	832.7475
	1000	0.1532	0.3147	0.1559	20.2267	24.1433	0.0315	0.1743	0.0410	431.5102	642.7975
Ω	30	1.3945	2.0766	0.6055	2.1069	2.7606	3.7484	6.4942	0.6311	9.2455	31.0687
	80	1.0232	1.3251	0.8992	2.6983	2.9982	2.2029	3.0258	1.2445	18.4101	17.4671
	200	0.5734	0.8506	0.8384	1.9363	3.2026	0.5469	1.6251	1.0584	6.0719	21.4858
	500	0.3220	0.7811	0.7332	1.8616	2.9325	0.2337	1.1423	0.9002	7.1347	14.5020
	1000	0.7011	0.7927	0.4509	4.9613	5.2361	0.6855	1.0064	0.3930	46.0868	43.8423

Table 2

Results of Simulations for $(\Psi, \Phi, \delta, \Omega) = (0.3, 0.7, 0.6, 1.0)$

Parameter	N	ab					mse				
		MLE	O	W	C	A	MLE	O	W	C	A
Ψ	30	0.2698	0.2086	0.2371	13.5748	11.3711	0.0765	0.0535	0.0626	268.6568	316.3286
	80	0.2434	0.2228	0.1806	9.1300	22.3308	0.0721	0.0592	0.0522	118.6071	852.4057
	200	0.1193	0.2180	0.1889	10.4509	24.6038	0.0296	0.0621	0.0534	169.6483	827.6018
	500	0.1397	0.1681	0.1404	10.3455	50.9891	0.0303	0.0464	0.0269	142.2032	2916.036
	1000	0.2252	0.1344	0.1347	14.3517	47.1094	0.0639	0.0366	0.0305	228.4411	2465.843

Contd...

Φ	30	0.6286	0.9968	0.7317	2.0432	1.7773	0.9680	1.8621	0.8903	14.6218	3.9275
	80	1.7077	0.3892	0.7359	3.8029	3.1668	10.4070	0.2265	0.9102	30.5244	33.1390
	200	0.6937	0.7383	1.4451	1.6958	2.3850	0.9408	1.6021	9.0287	6.0972	21.4689
	500	0.4377	0.4484	0.5215	1.7016	1.0061	0.2873	0.2574	0.3233	5.1822	1.0276
	1000	0.3213	0.3698	0.4537	3.0081	0.9283	0.1630	0.2007	0.2708	21.2171	0.8722
δ	30	0.2747	0.2936	0.3492	15.8928	33.1261	0.1085	0.1238	0.1403	282.8019	1818.799
	80	0.4215	0.3381	0.3341	18.4690	30.9019	0.3003	0.1424	0.1359,	362.4512	1226.404
	200	0.2751	0.3129	0.3832	20.5009	45.3442	0.0896	0.1304	0.2566	438.7893	2332.963
	500	0.2473	0.2401	0.1923	25.0683	63.0271	0.0955	0.0830,	0.0608	681.9098	4284.441
	1000	0.2464	0.1539	0.2285	27.7694	60.2618	0.0912	0.0392	0.0749	831.2347	3756.44
Ω	30	1.1085	1.7054	1.9748	2.8332	5.2564	2.3895	3.8585	6.2019	11.2405	35.7389
	80	1.0250	1.6933	1.7599	2.3893	12.2858	2.3002	5.6833	4.8018	9.5992	218.7645
	200	1.2550	0.9142	0.9964	3.5006	9.4239	2.2481	1.0905	1.3969	16.8369	147.4287
	500	1.3039	0.9597	0.9226	4.6598	8.7626	2.8990	1.4930	1.1654	23.3329	96.9149
	1000	0.7671	0.7618	0.9021	4.4774	11.0793	1.1059	1.0287	1.2151	28.0616	127.5515

6. Data Analysis

In this section, an assessment of HMGOM distribution is made by applying it to two real-world data sets. The data sets are 63 observations of the strength of 1.5 cm glass fibers [15, 16] and the failure times (10^3 hours) of 40 turbochargers from a type of diesel engine [17]

To evaluate the performance of the HMGOM distribution, we compare it with nine different modifications of the Gompertz distribution. The goodness-of-fits of the chosen distributions are assessed using the Kolmogorov-Smirnov (KS) test, Anderson-Darling (A) test, and CramÃ©r-von Mises (C) test. The model with the best performance will be identified based on the lowest values of the Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (AICc), and Bayesian Information Criterion (BIC).

These nine(9) distributions include the new odd Lindley-Gompertz distribution (OLINGOMD) [3], the Kumaraswamy Gompertz distribution (KWGO) [4],the Rayleigh Gompertz (RGO) [5], the exponentiated generalised Weibull-Gompertz distribution (EGWGD)[6], the Rayleigh Gamma Gompertz distribution (RGGOM) [7], the inverse power Gompertz (IPG)[8], the odd generalised exponential Gompertz distribution (OGEG) [9], the inverted shifted Gompertz distribution(ISGZ) [10] and the Nadarajah Haghighi Gompertz distribution (NHGD)[11].

6.1 *The 1.5 cm glass fibres*

Table 3 presents the MLEs for the fitted distributions.

We conclude from the results in Table 4 that the newly proposed HMGOM offers the best fit for the data set. It achieved the least values for the AIC, AICc, and BIC criteria as well as the least A, KS, and C statistics.

Table 3

MLEs for 1.5cm glass fibres

Model		Estimates	Errors	Z	P-value
HMGOM	α	0.0151	0.0282	0.5341	0.5933
	θ	0.0089	0.0078	1.1426	0.2532
	a	1.2271	0.0892	13.7637	2.00×10^{-16}*
	b	0.9833	0.0468	21.0250	2.00×10^{-16}*
OLINGOMD	α	3.0487	0.9701	3.1426	0.0017
	θ	0.0106	0.0066	1.6112	0.1071
	λ	2.6598	0.2773	9.5926	2.00×10^{-16}*
KWGO	γ	2.9869	0.3959	7.5455	4.51×10^{-14}*
	θ	0.0265	0.0142	1.8675	0.0618
	a	1.5130	0.3648	4.1476	3.36×10^{-5}*
	b	1.0979	0.5843	1.8790	0.0603
RGO	α	0.0265	0.0131	2.0189	0.4350
	λ	2.3014	0.2785	8.2634	2.00×10^{-16}*
	θ	1.0171	0.3288	3.0937	0.0020
EGWGD	θ	0.6915	0.3349	2.0651	0.0389
	a	0.1820	0.4658	0.3908	0.6960
	b	5.4824	7.8627	0.6973	0.4856
	c	0.1133	0.3182	0.3561	0.7217
	d	1.4886	5.1564	0.2887	0.7728
RGGOM	α	2.2785	0.7811	2.9170	0.0035
	λ	0.4422	0.1516	2.9170	0.0035
	θ	21.0836	7.0275	3.0002	0.0027
IPG	α	2.4742	0.2257	10.9619	2.00×10^{-16}*
	β	0.2177	0.0713	3.0542	0.0023
	θ	6.9429	2.6243	2.6456	0.0082
OGEG	α	11.5380	6.1085	1.8889	0.0589

Contd...

	β	4.7064	3.1616	1.4886	0.1366
	λ	0.0287	0.0133	2.1578	0.0309
	c	1.4186	0.4852	2.9234	0.0035
ISGZ	α	33.7053	9.4491	3.5670	0.0004
	θ	5.7713	0.4573	12.6200	2.00×10^{-16}*
NHGD	α	0.0937	0.0326	2.8738	0.0041
	λ	0.0919	0.0320	2.8738	0.0041
	β	2.9280	0.4258	6.8757	6.17×10^{-12}*
	δ	2.0318	0.9085	2.2366	0.0253

Table 4
Comparison criteria for 1.5cm glass fibres

Model	ℓ	AIC	AICc	BIC	KS	A	C
HMGOM	-12.0840	32.1681	32.5749	40.7406	0.0918	0.3785	0.0611
OLINGOMD	-18.7784	43.5567	43.9635	49.9861	0.1447	1.7788	0.3092
KWGO	-14.2142	36.4284	37.1180	45.0009	0.1119	0.8936	0.1441
RGO	-16.8399	41.3723	41.7791	47.8017	0.1637	1.7808	0.3074
EGWGD	-14.6649	39.3298	40.3824	50.0455	0.1457	1.0828	0.1966
RGGOM	-34.4107	74.8214	75.2281	81.2508	0.2181	4.9108	0.9045
IPG	-51.5801	109.1601	110.2127	115.5895	0.2646	7.7531	1.5242
OGEG	-16.4297	40.8594	41.5490	49.4319	0.1402	1.4076	0.2103
ISGZ	-27.2560	58.5120	58.8363	62.7983	0.2192	3.8317	0.7204
NHGD	-16.9935	41.9869	42.6766	50.5595	0.1434	1.5199	0.2678

To further validate the conclusion made earlier, the fitted pdfs and cdfs of the models shown in Figure 3 were assessed. The HMGOM model provides a better fit to the 1.5 cm glass fibres.

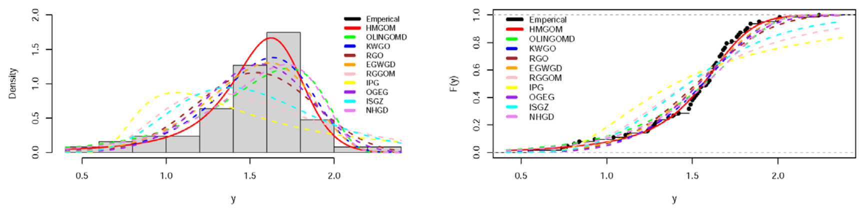

Figure 3

Fitted pdfs (Left) and cdfs (Right) for 1.5cm glass fibres

6.2 *Turbochargers failure time*

The estimates of the models that were compared are displayed in Table 5.

Table 5

MLEs for turbochargers failure time

Model		Estimates	Errors	Z	P-value
HMGOM	α	0.2402	0.2521	0.9528	0.3407
	θ	0.0014	0.0162	0.0889	0.9292
	a	0.5403	0.1061	5.0915	3.55×10^{-7}*
	b	0.0441	0.0580	0.7607	0.4469
OLINGOMD	α	19.9310	3.73×10^{-3}	5346.4782	2.00×10^{-16}*
	θ	8.14×10^{-4}	7.75×10^{-4}	1.0504	0.2936
	λ	0.5089	0.1485	3.4271	0.0006*
KWGO	γ	0.1979	0.0694	2.8521	0.0043*
	θ	0.0538	0.0245	2.1940	0.0282*
	a	2.8612	0.7997	3.5776	0.0003*
	b	4.5520	2.1160	2.1512	0.0315*
RGO	α	0.0208	0.0218	0.9547	0.3397
	λ	0.3267	0.1300	2.5130	0.0120*
	θ	1.2114	0.9450	1.2819	0.1999
EGWGD	θ	1.6283	0.3982	4.0888	4.34×10^{-5}*
	a	6.4060	0.0435	147.2327	2.20×10^{-16}*
	b	1.9263	0.0737	26.1349	2.20×10^{-16}*
	c	0.0014	0.0004	3.3205	0.0009*
	d	0.6035	0.0739	8.1687	3.12×10^{-16}*
RGGOM	α	6.2600	8.1329	0.7697	0.4415

Contd...

	λ	0.0245	0.0319	0.7697	0.4415
	θ	5.0530	1.7069	2.9603	0.0031*
IPG	α	1.8488	0.1983	9.3250	2.00×10^{-16}
	β	0.6525	0.3035	2.1500	0.0316*
	θ	24.3526	12.6022	1.9324	0.0533
OGEG	α	12.2469	0.7449	16.4421	2.00×10^{-16}*
	β	3.4789	1.8425	1.8881	0.0590
	λ	0.0089	0.0064	1.3793	0.1678
	c	0.2348	0.1027	2.2862	0.0222*
ISGZ	α	17.4313	7.3819	2.3613	0.01821*
	θ	19.3358	2.6878	7.1939	6.30×10^{-13}*
NHGD	α	0.8923	0.5065	1.7617	0.0781
	λ	0.0037	0.0021	1.7617	0.0781*
	β	0.5336	0.0871	6.1245	9.10×10^{-10}*
	δ	2.6972	1.3429	2.0085	0.0446*

According to the results in Table 6, the HMGOM model offers the best fit for the data set. It achieved the the least values for the AIC, AICc, and BIC criteria as well as the least A, KS, and C statistics.

Table 6

Comparison criteria for turbochargers failure times data set !

Model	ℓ	AIC	AICc	BIC	KS	A	C
HMGOM	-78.2964	164.5928	165.7359	171.3483	0.0537	0.0971	0.0139
OLINGOMD	-80.7292	167.4584	168.1251	172.5250	0.1157	0.7601	0.1354
KWGO	-82.1978	172.3956	173.5385	179.1511	0.0837	0.4691	0.0399
RGO	-83.8853	173.7705	174.4372	178.8372	0.1027	0.7847	0.0986
EGWGD	-85.7540	181.5080	183.2727	189.9524	0.1595	1.2566	0.1393
RGGOM	-92.6762	191.3525	192.0192	196.4191	0.1900	2.2561	0.3829
IPG	-102.4359	210.8718	211.5385	215.9384	0.2405	3.8758	0.7149
OGEG	-82.5932	173.1865	174.3293	179.9420	0.0919	0.5940	0.0582
ISGZ	-91.5428	187.0855	187.4098	190.4633	0.1456	1.9066	0.2681
NHGD	-79.3480	166.6960	167.8389	173.4515	0.0744	0.2379	0.0388

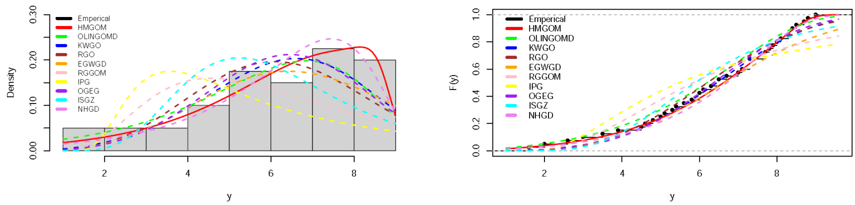

Figure 4

Fitted pdfs (Left) and cdfs (Right) for turbochargers failure time

To further validate the conclusion made earlier, the pdfs and cdfs of the models are fitted as shown in Figure 4 and assessed. The HMGOM model provides a better fit for the failure times.

7. HMGOM Regression Model

The HMGOM distribution is used to study the impact of some explanatory variables have on the response variable, the shape and scale parameters. To achieve this, we introduce the HMGOM regression model considering the parameters δ, Ω and Ψ are varying across observations using the logarithmic link functions (see [18], [19] and [20]) $\log(\delta_a) = x_a^T \delta_a$, $\log(\Omega_a) = x_a^T \Omega_a$ and $\log(\Psi_a) = x_a^T \Psi_a$, $a = 1, 2, 3\ldots, n$.

The SF of the HMGOM regression model is

$$S(y \mid x) = \frac{e^{-\frac{\exp(x_a^T\Omega_a)\exp(x_a^T\Psi_a)}{\exp(x_a^T\delta_a)}\left(e^{\exp(x_a^T\delta_a)y}-1\right)}}{\left[1-\Phi\left(1-e^{-\frac{\exp(x_a^T\Omega_a)(\exp(x_a^T\Psi_a)-1)}{\exp(x_a^T\delta_a)}\left(e^{\exp(x_a^T\delta_a)y}-1\right)}\right)\right]}. \tag{22}$$

The estimates of the HMGOM regression model are obtained by maximizing the likelihood function

$$\ell = n\sum_{a=1}^{n}\ln\Omega_a + \sum_{a=1}^{n}\delta_a x_a - \sum_{a=1}^{n}\frac{\Omega_a}{\delta_a}\Psi_a\left(e^{\delta x_a}-1\right) + \sum_{a=1}^{n}\ln\left[\Psi_a(1-\Phi)+\Phi e^{-\frac{\Omega_a(\Psi_a-1)}{\delta_a}\left(e^{\delta_a x_a}-1\right)}\right] - 2\sum_{a=1}^{n}\ln\left[1-\Phi\left(1-e^{-\frac{\Omega_a(\Psi_a-1)}{\delta_a}\left(e^{\delta_a x_a}-1\right)}\right)\right].$$

The practical application of the HMGOM regression model is made and compared to the Gompertz regression model [21]. The most appropriate model is selected based on the model with the least AIC value. Residual analysis is performed as a means to check the adequacy of the fitted model. The fitted model is assessed using the Cox Snell residuals generated. The residuals generated should exhibit approximately the same behaviour as samples from the standard exponential distribution if the models perform well [18]. The residuals of the models are diagnosed using C and KS goodness-of-fits test.

The relationship between Survival time (T) and duration of diabetes(DUR) in years of 40 male patients, was assessed using the HMGOM regression model. The was retrieved from [22]. They are:

Survival time(T) in years: 12.4, 14.1, 14.4, 14.2, 14.4, 12.4 ,12.4, 7, 13.6, 14.4 , 12.4, 9.8, 0, 12.1, 14.4, 12.4, 14.4, 14.2, 12.4, 13.7 ,13.4 ,12.6 ,14 ,12.4 , 12.9, 12.4, 14.5, 13.4, 13.9, 12, 3.6, 15.4, 10.3, 5.8, 2.5, 15, 5.5, 4.5, 6.8, and 3.6.

Duration of diabetes(DUR) in years: 3, 1, 4, 3, 7, 4, 2, 8, 14, 4, 3, 4, 6, 5, 2, 5, 9, 5, 6, 3, 15, 9, 1, 3, 1, 1, 3, 0, 6, 0, 18, 12, 3, 1, 18, 1, 18, 4, 17, and 18.

Table 7 displays the estimates of the models and some model selection criteria used. The coefficients for DUR were significant for the models fitted. However, the HMGOM regression model provides a better fit than the GZ regression model. The variable DUR in the HMGOM regression model, had a significant negative effect on the shape parameter δ and scale parameter Ω. On the other hand, the variable Dur had a positive effect on the shape parameter Ψ. The HMGOM regression model is obtained as

$$\log(\delta_a) = -0.4898 - 0.0763 DUR_a$$

$$\log(\Omega_a) = -8.2899 - 1.1961 DUR_a$$

$$\log(\Psi_a) = 0.5695 + 1.5167 DUR_a$$

Table 7

MLEs and goodness-of-fit

		Estimates	**P-value**	
HMGOM	Φ	0.6631(0.4168)	0.1116	
	δ_0	-0.4898(0.2178)	0.0245	ℓ = -91.9595
	δ_1	-0.0763(0.0179	1.983×10^{-5}	AIC=197.919
	Ω_0	-8.2899(1.5124)	4.218×10^{-8}	BIC=209.7411
	Ω_1	-1.1961(0.2763)	1.497×10^{-5}	
	Ψ_0	0.5695(0.2741)	0.7351	
	Ψ_1	1.5167(0.2741)	3.128×10^{-8}	
GZ	γ	0.3933(0.0581)	1.282×10^{-11}	ℓ = -101.2672
	λ_0	-6.3683(0.7022)	2.2×10^{-16}	AIC=208.5345
	λ_1	0.0673(0.0339)	0.04712	BIC=213.6011

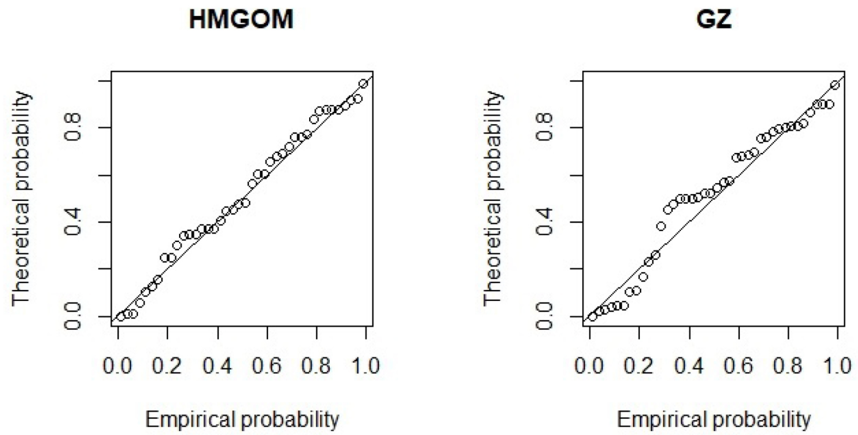

Figure 5

P-P plots

The residuals were obtained to ascertain the adequacy of both models. The probability-probability (P-P) plots for the residuals of the HMGOM and GZ regression models are observed in Figure 5 and that of the HMGOM is seen to be nearer to the diagonal line than the GZ regression model, suggesting that the HMGOM regression model performs better.

Table 8 presents the diagnostic results of the fitted models, indicating that the HMGOM regression model performs better.

Table 8

Goodness-of-fit Test Results

	KS	C
	Statistic	Statistic
HMGOM	0.0939	0.0530
GZ	0.1554	0.1685

8. Conclusion

This paper introduced the Harmonic Mixture Gompertz (HMGOM) distribution, a novel four-parameter model derived from the weighted harmonic mean of two Gompertz distributions. Key statistical properties were explored, providing a comprehensive understanding of the distribution's behaviour. Parameter estimation was performed using multiple methods. Simulation studies confirmed the effectiveness of the estimators. The HMGOM distribution demonstrated superior performance in applications to two real-life lifetime data sets, outperforming nine existing Gompertz extensions. Furthermore, the development and application of the HMGOM regression model showcased its potential as a strong alternative to the Gompertz regression model for lifetime data analysis. The findings of this study emphasize the flexibility and efficacy of the HMGOM distribution in modelling complex data sets. Future research could focus on expanding the model to exploring multi-component mixtures, or applying the HMGOM distribution in broader fields such as reliability engineering and biomedical sciences.

References

[1] O. Kharazmi, A. S. Nik, G. Hamedani, and E. Altun, "Harmonic Mixture-G Family of Distributions: Survival Regression, Simulation by Likelihood, Bootstrap and Bayesian Discussion with MCMC Algorithm," *Austrian Journal of Statistics,* vol. 51, no. 2, pp. 1–27 (2022).

[2] J. T. Eghwerido, E. Efe-Eyefia, and O. Otakore, "Performance rating of the zubair gompertz distribution: properties and applications," *Journal of Statistics and Management Systems,* vol. 24, no. 8, pp. 1601–1626 (2021).

[3] O. Atanda, T. Mabur, and G. Onwuka, "A new odd Lindley-Gompertz distribution: Its Properties and Applications," *Asian Journal of Probability and Statistics*, vol. 7, no. 4, pp. 29–47, 07 (2020).

[4] R. C. d. Silva, J. J. Sanchez, F. P. Lima, and G. M. Cordeiro, "The kumaraswamy Gompertz distribution," *Journal of Data Science,* vol. 13, no. 2, pp. 241–260 (2021).

[5] N. H. Al-Noor, M. A. Khaleel, and N. K. Assi, "The Rayleigh Gompertz distribution: Theory and real applications," *International Journal of Nonlinear Analysis and Applications,* vol. 13, no. 1, pp. 3505–3516 (2022).

[6] A. El-Bassiouny, M. El-Damcese, A. Mustafa, and M. Eliwa, "Exponentiated generalized Weibull-Gompertz distribution with application in survival analysis," *Journal of Statistics Applications and Probability,* vol. 6, no. 1, pp. 7–16 (2017).

[7] N. H. Al-Noor and N. K. Assi, "Rayleigh gamma Gompertz distribution: properties and applications," in AIP Conference Proceedings. AIP Publishing LLC (2021).

[8] D. H. Abdelhady and Y. M. Amer, "On the inverse power Gompertz distribution," *Annals of Data Science*, vol. 8, no. 3, pp. 451–473 (2021).

[9] M. El-Damcese, A. Mustafa, B. El-Desouky, and M. Mustafa, "The odd generalized exponential Gompertz," *Applied Mathematics,* vol. 6, no. 14, pp. 2340–2353 (2015).

[10] A. K. Chaudhary, L. P. Sapkota, and V. Kumar, "Inverted shifted Gompertz distribution with theory and applications," *Pravaha,* vol. 26, no. 1, pp. 1–10 (2020).

[11] A. A. Ogunde, I. O. Ajao, and G. A. Olalude, "On the application of Nadarajah Haghighi Gompertz distribution as a life time distribution," *Open Journal of Statistics,* vol. 10, no. 5, pp. 850–862 (2020b).

[12] I. B. Eraikhuemen, T. G. Ieren, F. C. Adanaá, A. I. Asongo, B. S. Yakubu, and A. H. Bala, "A transmuted power Gompertz distribution: Properties and Applications," *Theory and Practice of Mathematics and Computer Science,* vol. 9, pp. 127–145 (2021).

[13] S. Dey, F. A. Moala, and D. Kumar, "Statistical properties and different methods of estimation of gompertz distribution with application," *Journal of Statistics and Management Systems*, vol. 21, no. 5, pp. 839–876 (2018).

[14] S. K. Maurya, A. S. Yadav, S. K. Singh, and U. Singh, "A new flexible logarithmic transform heavy-tailed distribution and its applications," *Journal of Statistics and Management Systems,* vol. 24, no. 5, pp. 917–950 (2021).

[15] J. T. Eghwerido, J. O. Ogbo, and A. E. Omotoye, "The Marshall-Olkin Gompertz Distribution: Properties and Applications," *Statistica,* vol. 81, no. 2, pp. 183–215 (2021).

[16] M. A. Khaleel, N. H. Al-Noor, and M. K. Abdal-Hameed, "Marshall Olkin exponential Gompertz distribution: Properties and applications," *Periodicals of Engineering and Natural Sciences*, vol. 8, no. 1, pp. 298–312 (2020).

[17] R. R. Guerra, F. A. Peña-Ramírez, and G. M. Cordeiro, "The Weibull Burr XII distribution in lifetime and income analysis," *Anais da Academia Brasileira de Ciências,* vol. 93, no. 3 (2021).

[18] S. Nasiru, A. G. Abubakari, and I. D. Angbing, "Bounded odd inverse pareto exponential distribution: Properties, estimation, and regression," *International Journal of Mathematics and Mathematical Sciences*, vol. 2021 (2021).

[19] F. Prataviera, E. M. Ortega, G. M. Cordeiro, R. R. Pescim, and B. A. Verssani, "A new generalized odd log-logistic flexible weibull regression model with applications in repairable systems," *Reliability Engineering & System Safety*, vol. 176, pp. 13–26 (2018).

[20] F. Prataviera, J. C. S. Vasconcelos, G. M. Cordeiro, E. Hashimoto, and E. M. M. Ortega, "The exponentiated power exponential regression model with different regression structures: application in nursing data," *Journal of Applied Statistics*, vol. 46, no. 10, pp. 1792–1821 (2019).

[21] N. N. N. Azid, J. Arasan, H. S. Zulkafli, M. B. Adam et al., "Assessing the adequacy of the Gompertz regression model in the presence of right censored data," in *Journal of Physics: Conference Series.* IOP Publishing (2021).

[22] E. T. Lee and J. Wang, Statistical methods for survival data analysis. John Wiley and Sons (2003).

Received May, 2023