STATISTICA, anno LXXIX, n. 1, 2019

UNIT-GOMPERTZ DISTRIBUTION WITH APPLICATIONS

Josmar Mazucheli
Departamento de Estatística, Universidade Estadual de Maringá, Maringá, Brazil
André Felipe Menezes
Departamento de Estatística, Universidade Estadual de Maringá, Maringá, Brazil
Sanku Dey[1]
Department of Statistics, St. Anthony's College, Shillong, Meghalaya, India

1. Introduction

Benjamin Gompertz in 1825 introduced the Gompertz distribution in connection with human mortality and actuarial tables. Since then, considerable attention has received from demographers and actuaries. This distribution is a generalization of the exponential distribution and has many real life applications, especially in medical and actuarial studies. It has some nice relations with some of the well-known distributions such as exponential, double exponential, Weibull, extreme value (Gumbel distribution) or generalized logistic distribution (see Willekens, 2001). An important characteristic of the Gompertz distribution is that it has an exponentially increasing failure rate for the life of the systems. In recent past, many authors have contributed to the studies of statistical methodology and characterization of this distribution, for example Read (1983), Makany (1991), Rao and Damaraju (1992), Franses (1994), Chen (1997) and Wu and Lee (1999).

The statistics literature has numerous distributions for modeling lifetime data. But many if not most of these distributions arise either due to theoretical considerations or practical applications or both. For example, there is no apparent physical motivation for the gamma distribution. It only has a more general mathematical form than the exponential distribution with one additional parameter, so it has nicer properties and provides better fits. The same arguments apply to Weibull and many other distributions. Many generalizations of the Gompertz distribution have been attempted by researchers. Notable among them are: Bemmaor (1994) proposed the shifted Gompertz distribution, Roy and Adnan (2012) introduced the wrapped generalized Gompertz distribution, El-Gohary *et al.* (2013) studied the generalized Gompertz distribution, Jafari *et al.* (2014)

[1] Corresponding Author. E-mail: sanku_dey2k2003@anthonys.ac.in

introduced four parameter beta-Gompertz distribution, El-Damcese *et al.* (2015) proposed the odd generalized exponential Gompertz distribution, Jafari and Tahmasebi (2016) introduced the generalized Gompertz-power series distributions and Benkhelifa (2017) introduced the beta generalized Gompertz distribution.

The aim of this note is to derive a new distribution from the Gompertz distribution by a transformation of the type $X = e^{-Y}$, where Y has the Gompertz distribution. The proposed distribution encompasses the behavior of and provides better fits than some well known lifetime distributions, such as Beta and Kumaraswamy (Kumaraswamy, 1980) distributions. We are motivated to introduce the UG distribution because (i) it is capable of modeling constant, increasing, upside-down bathtub and then bathtub shaped hazard rate; (ii) it can be viewed as a suitable model for fitting the skewed data which may not be properly fitted by other common distributions and can also be used in a variety of problems in different areas such as environmental studies and industrial reliability and survival analysis; and (iii) two real data applications show that it compares well with other three competing lifetime distributions in modeling environmental and failure data.

The rest of the paper is organized as follows. In Section 2, we introduce the unit-Gompertz distribution and discuss basic properties of this family of distributions. In Section 3, maximum likelihood estimators of the unknown parameters along with the expected Fisher information matrix are obtained. Monte Carlo simulations are conducted in Section 4 to investigate the performance of the maximum likelihood estimators and the asymptotic confidence intervals of the parameters. The analysis of two real data sets have been presented in Section 5. Finally, Section 6 concludes the paper.

2. MODEL DESCRIPTION AND SOME PROPERTIES

In this section, we describe the new bounded distribution, which arises from a logarithmic transformation in the Gompertz distribution. This transformation is also considered in Grassia (1977) for the unit-Gamma distribution and Gómez-Déniz *et al.* (2014) for the log-Lindley distribution.

Let Y be a non negative random variable with Gompertz distribution, then its probability density function is given by:

$$f(y \mid \alpha, \beta) = \alpha \, \beta \, \exp\left[\beta \, y + \alpha - \alpha \, e^{\beta \, y}\right], \tag{1}$$

where $y > 0$, $\beta > 0$ and $\alpha > 0$ are scale and shape parameters, respectively.

By considering the transformation:

$$X = e^{-Y}, \tag{2}$$

we obtain a new distribution with support on $(0, 1)$, which we refer to as unit-Gompertz distribution. Its probability density function is given by:

$$f(x \mid \alpha, \beta) = \alpha \, \beta \, x^{-(\beta+1)} \exp\left[-\alpha \, (x^{-\beta} - 1)\right], \tag{3}$$

for $x \in (0,1)$. The corresponding cumulative distribution function and hazard functions are given, respectively, by:

$$F(x \mid \alpha, \beta) = \exp\left[-\alpha\,(x^{-\beta} - 1)\right], \tag{4}$$

$$h(x \mid \alpha, \beta) = \frac{\alpha\,\beta\,x^{-(\beta+1)}\exp\left[-\alpha\,(x^{-\beta}-1)\right]}{1-\exp\left[-\alpha\,(x^{-\beta}-1)\right]}. \tag{5}$$

For $x \to 1$, $\lim_{x\to 1} h(x) = \alpha\beta$ for all possible choices of the parameter β and α. From the above, it can be observed the following:

- Monotonically increasing shapes are possible for all values of $\alpha > 1$ and $\beta \geq 1$.
- Possibly bathtub shapes of the hazard rate function will happen when $\alpha \leq 0.5$.

2.1. *Shapes*

PROPOSITION 1. *The p.d.f. of the UG distribution is log-concave and unimodal.*

PROOF. The second derivative of $\log f(x \mid \alpha, \beta)$ is

$$\frac{d^2}{dx^2}\log f(x \mid \alpha, \beta) = -\frac{\beta+1}{x^2}\left(\frac{\alpha\beta}{x^{\beta}} - 1\right) < 0. \tag{6}$$

Since $\left(\frac{\alpha\beta}{x^{\beta}} - 1\right) > 0$ for all $\beta > 0$, $\alpha > 0$ and $x \in (0,1)$ then $\log f(x \mid \alpha, \beta)$ is concave for all $\beta > 0$, $\alpha > 0$. Then, $f(x)$ is log-concave and unimodal. □

2.2. *Quantile function*

The quantile function $x = Q(p) = F^{-1}(p)$, for $0 < p < 1$, of the UG distribution is obtained by inverting Equation (4) is given by:

$$Q(p \mid \alpha, \beta) = \exp\left[-\frac{1}{\beta}\log(\alpha - \log p) - \log\alpha\right]. \tag{7}$$

2.3. *Mode*

The first derivative of $\log f(x \mid \alpha, \beta)$ is

$$\frac{d}{dx}\log f(x \mid \alpha, \beta) = \alpha\beta x^{-(\beta+1)} - \frac{(\beta+1)}{x}. \tag{8}$$

Therefore, the mode of $f(x \mid \alpha, \beta)$ is the root of the equation

$$\frac{d}{dx} \log f(x \mid \alpha, \beta) = \alpha \beta x^{-(\beta+1)} - \frac{(\beta+1)}{x} = 0. \tag{9}$$

Thus, if $x = x_0$:

$$x_0 = \left(\frac{\alpha \beta}{\beta + 1} \right)^{\frac{1}{\beta}} \tag{10}$$

implies that $x_0 = \left(\frac{\alpha\beta}{\beta+1} \right)^{\frac{1}{\beta}}$ is the unique critical point at which $f(x \mid \alpha, \beta)$ is maximized.

2.4. *Hazard rate function*

LEMMA 2. *Let X be a nonnegative continuous random variable with twice differentiable p.d.f. $f(x)$ and hazard rate function $h(x)$. Let $\eta(x) = -\frac{d}{dx} \log f(x)$.*

(i) If $\eta(x)$ is decreasing (increasing) in x, then $h(x)$ is increasing (decreasing) in x.

(ii) If $\eta(x)$ has a bathtub (upside-down bathtub) shape, then $h(x)$ also has a bathtub (upside-down bathtub) shape.

From Glaser's result we can determine the shape of the hazard rate function of the UG distribution as follows.

THEOREM 3. *The hazard rate of the UG distribution is upside-down bathtub shaped.*

PROOF. Since

$$\eta(x) = \frac{\beta+1}{x} - \alpha \beta x^{-(\beta+1)}, \tag{11}$$

it follows that

$$\eta'(x) = -\frac{\beta+1}{x^2} + \alpha\beta(\beta+1) x^{-(\beta+2)}. \tag{12}$$

Then, $\eta(x) = 0$ implies that $\eta(x)$ has a global maximum at $x^* = (\alpha\beta)^{\frac{1}{\beta}}$, since

$$\eta''(x) = -\frac{\alpha \beta^2}{(x^*)^{(\beta+3)}} < 0. \tag{13}$$

This means that $\eta(x)$ is upside-down bathtub shaped. Hence, by Glaser's lemma it follows that $h(x)$ is also upside-down bathtub shaped. Note that with $\eta'(x) > 0$ for given $x \in (0,1)$, $\alpha \geq 1$ and $\beta \geq 1$. Thus $\eta(x)$ is an increasing function. □

Figures 1 and 2 show the various curves for the p.d.f. and the hazard rate function, respectively, of UG distribution with various values of α and β. Figure 1 indicates that the UG distribution can be increasing, unimodal, reversed J-shaped and positively right skewed. Figure 2 shows that the hazard rate function $h(x \mid \alpha, \beta)$ of UG distribution is constant, increasing and upside-down bathtub-shaped. One of the advantages of the UG distribution over the Gompertz distribution is that the latter cannot model phenomenon showing an upside-down bathtub-shaped hazard function.

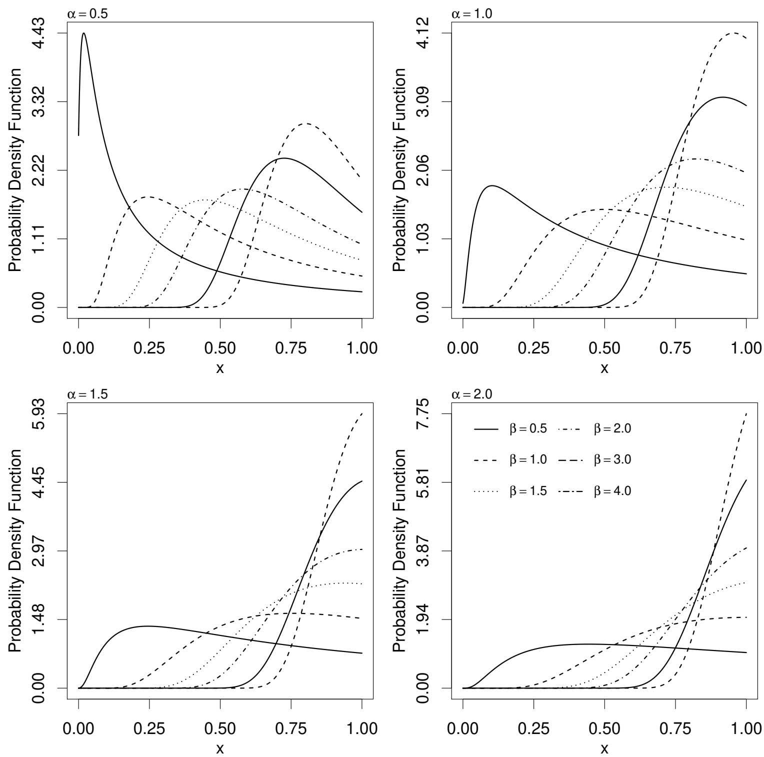

Figure 1 – Density plots of the unit-Gompertz distribution considering different values of α and β.

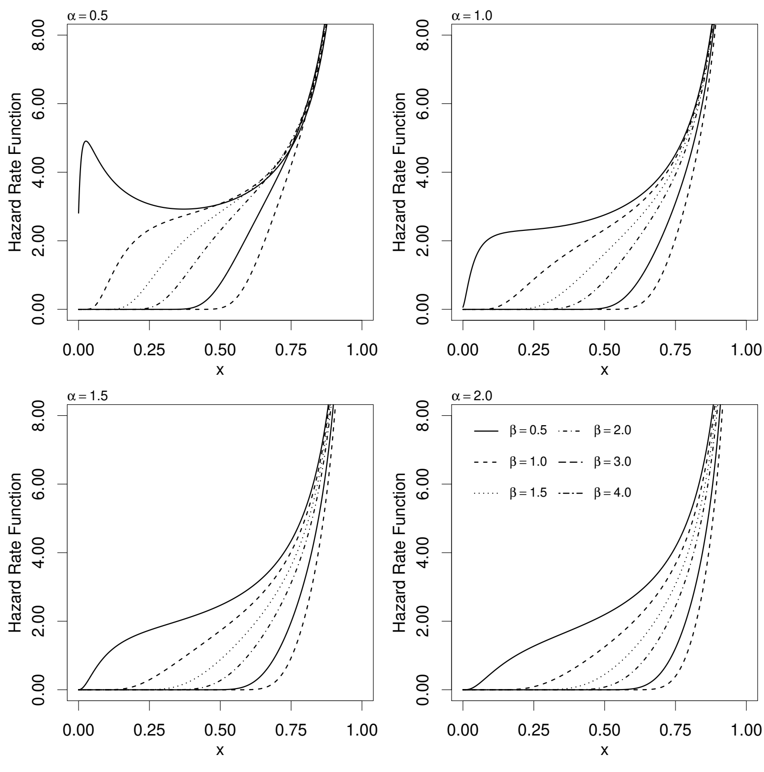

Figure 2 – Hazard rate plots of the unit-Gompertz distribution considering different values of α and β.

2.5. *Moments and associated measures*

We hardly need to emphasize the necessity and importance of the moments in any statistical analysis especially in applied work. Some of the most important features and characteristics of a distribution can be studied through moments (e.g., tendency, dispersion, skewness, and kurtosis). If the random variable X is UG distributed, then its rth moment around zero can be expressed as

$$\mu'_r = \mathbb{E}(X^r) = \int_0^1 x^r \alpha \beta x^{-(\beta+1)} \exp\left[-\alpha (x^{-\beta} - 1)\right] \mathrm{d}x = \alpha^{\frac{r}{\beta}} e^{\alpha} \Gamma\left(1 - \frac{r}{\beta}, \alpha\right), \quad (14)$$

where $\Gamma(\cdot,\cdot)$ is the upper incomplete gamma function. Note that, the moments exists only when $\frac{r}{\beta} < 1$.

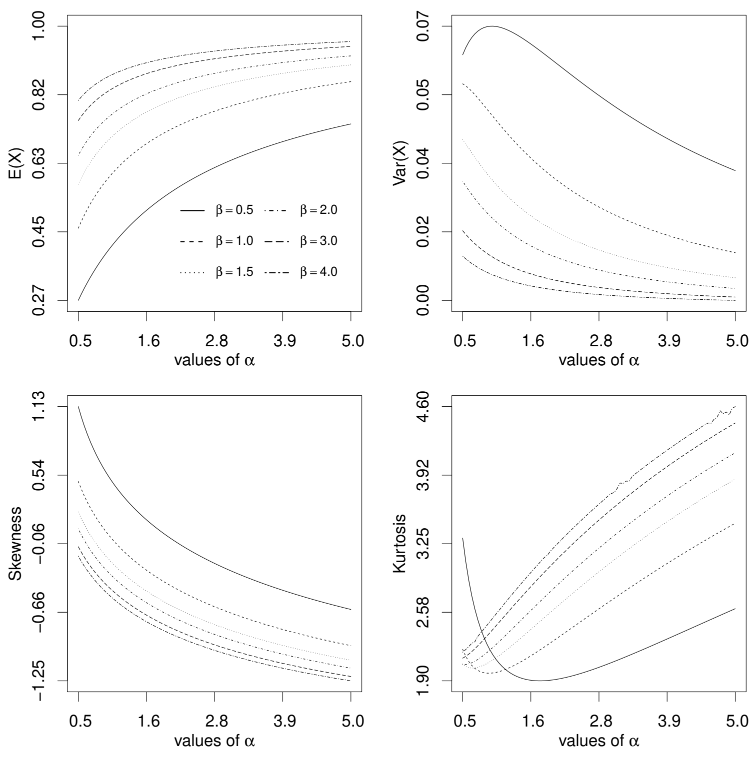

Figure 3 – Mean, variance, skewness and kurtosis of the unit-Gompertz distribution considering different values of α and β.

Thus, the mean and variance are then obtained by:

$$\mu = \alpha^{\frac{1}{\beta}}\, e^{\alpha}\, \Gamma\left(1 - \frac{1}{\beta}, \alpha\right),$$

$$\sigma^2 = \alpha^{\frac{2}{\beta}}\, e^{2\alpha} \left\{ \Gamma\left(1 - \frac{2}{\beta}, \alpha\right) - \left[\Gamma\left(1 - \frac{1}{\beta}, \alpha\right)\right]^2 \right\}.$$

The skewness and kurtosis measures can be calculated using the relations

$$\text{Skewness}(X) = \frac{\mu_3' - 3\,\mu_2'\,\mu + \mu^3}{\sigma^3},$$

$$\text{Kurtosis}(X) = \frac{\mu_4' - 4\,\mu_3'\,\mu + 6\,\mu_2'\,\mu^2 - 3\,\mu^4}{\sigma^4}.$$

Figure 3 displays the mean, variance, skewness and kurtosis of UG distribution as a function of α and β.

3. Estimation and inference

In this section, we obtain the maximum likelihood estimators and the observed Fisher information matrix from complete samples for the UG distribution. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be a random sample of size n from the unit-Gompertz distribution with unknown parameter vector $\theta = (\alpha, \beta)$. The likelihood function for θ is given by:

$$L(\mathbf{x} \mid \theta) = \alpha^n\, \beta^n \prod_{i=1}^{n} x_i^{-(\beta+1)} \exp\left[-\alpha\,(x_i^{-\beta} - 1)\right]. \tag{15}$$

Thus, the log-likelihood function, apart constant terms, can be written as:

$$\ell(\mathbf{x} \mid \theta) = n\,\log\alpha + n\,\log\beta - (\beta + 1)\sum_{i=1}^{n} \log x_i - \alpha \sum_{i=1}^{n} x_i^{-\beta} + n\,\alpha. \tag{16}$$

By taking the partial derivatives with of the log-likelihood function with respect to α and β we have the Score vector $\mathbf{U}_\theta = [\,U_\alpha\; U_\beta]$, which their components are

$$U_\alpha = n\left(1 + \frac{1}{\alpha}\right) - \sum_{i=1}^{n} x_i^{-\beta} \tag{17}$$

and

$$U_\beta = \frac{n}{\beta} + \alpha \sum_{i=1}^{n} \frac{\log x_i}{x_i^{\beta}} - \sum_{i=1}^{n} \log x_i. \tag{18}$$

The maximum likelihood estimates $\widehat{\theta} = (\widehat{\alpha}, \widehat{\beta})$ of $\theta = (\alpha, \beta)$ is obtained by setting $U_\alpha = U_\beta = 0$ and solving these equations simultaneously. From (17), it follows immediately that the maximum likelihood estimator of α can be get by:

$$\widehat{\alpha} = \frac{n}{\sum_{i=1}^{n} x_i^{-\widehat{\beta}} - n}. \tag{19}$$

The Hessian matrix corresponds to the second derivatives of the log-likelihood with respect to the unknown parameters is given by:

$$\mathbf{H} = \left[\begin{array}{c|c} -\dfrac{n}{\alpha^2} & \displaystyle\sum_{i=1}^{n} \frac{\log x_i}{x_i^{\beta}} \\ \hline \displaystyle\sum_{i=1}^{n} \frac{\log x_i}{x_i^{\beta}} & -\alpha \displaystyle\sum_{i=1}^{n} \frac{(\log x_i)^2}{x_i^{\beta}} + \frac{n}{\beta^2} \end{array}\right] \tag{20}$$

So, the expected Fisher information matrix is given by:

$$\mathbf{I}(\theta) = n \left[\begin{array}{c|c} \dfrac{1}{\alpha^2} & \dfrac{e^{\alpha} E_1(\alpha) + 1}{\beta\,\alpha} \\ \hline \dfrac{e^{\alpha} E_1(\alpha) + 1}{\beta\,\alpha} & \alpha^2 \beta I_{22} + \dfrac{1}{\beta^2} \end{array}\right], \tag{21}$$

where $E_n(\cdot)$ represents the exponential integral function defined as

$$E_k(x) = \int_1^{\infty} \frac{e^{-xt}}{t^k}\, dt = x^{k-1}\Gamma(1-k, x) \tag{22}$$

and

$$I_{22} = \int_0^1 \left[\frac{(\log x)^2}{x^{\beta}}\right] x^{-(\beta+1)} \exp\left[-\alpha\,(x^{-\beta} - 1)\right] dx. \tag{23}$$

It is well known that under mild regularity conditions (see Lehmann and Casella, 1998), the asymptotic distribution of the maximum likelihood estimator $\widehat{\theta}$ of θ is such that

$$\sqrt{n}\,(\widehat{\theta} - \theta) \xrightarrow{D} N_2(0, \mathbf{I}^{-1}(\theta)), \tag{24}$$

where $\xrightarrow{D}$ denotes convergence in distribution and $\mathbf{I}^{-1}(\theta)$ is the inverse of expected Fisher information matrix.

Lindsay and Li (1997) have shown that the observed Fisher information is a consistent estimator of the expected Fisher information. Therefore the asymptotic behavior remains if $\mathbf{I}(\theta) = \lim_{n\to\infty} n^{-1}\mathbf{J}(\theta)$, where $\mathbf{J}(\theta)$ denotes the observed Fisher information matrix.

4. Simulation study

In this section we evaluate the performance of the maximum likelihood estimators and the asymptotic confidence intervals of the parameters that index the unit-Gompertz distribution discussed in the previous sections through Monte Carlo simulation. We fixed the sample size at $n = 20, 30, 50$ and 100 and the parameters at $\alpha = 0.5, 1.0, 1.5$ and 2.0 and $\beta = 0.5, 1.0, 1.5, 2.0, 3.0$ and 4.0. For each combination, we generate $M = 10,000$ pseudo-random samples from the unit-Gompertz distribution using the inverse cumulative distribution function method, that is:

$$x = \exp\left[-\frac{\log(\alpha - \log u) - \log \alpha}{\beta}\right], \tag{25}$$

where u is a uniform $(0, 1)$ observation.

To assess the performance of the maximum likelihood estimators and their asymptotic confidence intervals we calculate the bias, root mean-squared error and coverage probabilities of 90% and 95% confidence levels. The following observations can be drawn from Tables 1-4:

1. All the estimators show the property of consistency i.e., the RMSE decreases as sample size increases.
2. The bias of $\widehat{\alpha}$ decreases with increasing n.
3. The bias of $\widehat{\beta}$ decreases with increasing n for all the method of estimations.
4. The RMSE of $\widehat{\alpha}$ increases with increasing α.
5. The RMSE of $\widehat{\beta}$ increases with increasing β.
6. The RMSE of $\widehat{\beta}$ increases with increasing α.
7. As n increases the CPs of α and β become very close to the nominal levels 90% and 95%.

TABLE 1

Estimated bias, root mean-squared and coverage probability α and β ($\alpha = 0.5$).

		Bias		RMSE		$CP_{90\%}$		$CP_{95\%}$	
β	n	α	β	α	β	α	β	α	β
0.5	20	0.0618	0.0923	0.5979	0.2422	0.7698	0.8920	0.8068	0.9469
	30	0.0390	0.0600	0.4468	0.1838	0.8054	0.8960	0.8396	0.9488
	50	0.0205	0.0351	0.2909	0.1354	0.8387	0.9008	0.8761	0.9468
	100	0.0086	0.0178	0.1886	0.0919	0.8724	0.8959	0.9061	0.9495
1.0	20	0.0593	0.1892	0.5945	0.4888	0.7625	0.8918	0.8022	0.9469
	30	0.0439	0.1194	0.4527	0.3743	0.8014	0.8902	0.8381	0.9442
	50	0.0220	0.0706	0.2967	0.2749	0.8373	0.8936	0.8691	0.9415
	100	0.0074	0.0371	0.1873	0.1847	0.8679	0.8948	0.9031	0.9461
1.5	20	0.0763	0.2784	0.6613	0.7406	0.7671	0.8908	0.8024	0.9456
	30	0.0427	0.1838	0.4574	0.5630	0.7990	0.8923	0.8389	0.9463
	50	0.0173	0.1108	0.2966	0.4064	0.8357	0.8950	0.8729	0.9517
	100	0.0061	0.0566	0.1837	0.2740	0.8694	0.9006	0.9051	0.9509
2.0	20	0.0706	0.3622	0.6292	0.9704	0.7713	0.8896	0.8086	0.9477
	30	0.0415	0.2372	0.4338	0.7439	0.8029	0.8916	0.8387	0.9441
	50	0.0215	0.1390	0.2911	0.5404	0.8387	0.8978	0.8748	0.9497
	100	0.0093	0.0695	0.1862	0.3669	0.8702	0.8975	0.9062	0.9504
3.0	20	0.0680	0.5542	0.6054	1.4510	0.7723	0.8926	0.8096	0.9455
	30	0.0463	0.3537	0.4553	1.1124	0.8046	0.8934	0.8430	0.9454
	50	0.0247	0.2028	0.2948	0.8083	0.8430	0.8980	0.8778	0.9480
	100	0.0101	0.1036	0.1895	0.5466	0.8745	0.8991	0.9079	0.9502
4.0	20	0.0770	0.6680	0.6521	1.8171	0.7793	0.8994	0.8140	0.9554
	30	0.0449	0.4605	0.4681	1.4558	0.8055	0.8927	0.8415	0.9472
	50	0.0201	0.2798	0.2914	1.0791	0.8419	0.8971	0.8760	0.9469
	100	0.0117	0.1246	0.1881	0.7266	0.8781	0.8983	0.9118	0.9491

TABLE 2

Estimated bias, root mean-squared and coverage probability α and β ($\alpha = 1.0$).

		Bias		RMSE		$CP_{90\%}$		$CP_{95\%}$	
β	n	α	β	α	β	α	β	α	β
0.5	20	0.1227	0.1475	1.2322	0.3405	0.7404	0.8977	0.7789	0.9479
	30	0.1156	0.0948	1.0260	0.2602	0.7784	0.8969	0.8131	0.9488
	50	0.0762	0.0547	0.7457	0.1887	0.8138	0.8981	0.8523	0.9489
	100	0.0308	0.0272	0.4519	0.1267	0.8595	0.8987	0.8913	0.9492
1.0	20	0.1331	0.2923	1.2212	0.6825	0.7464	0.8952	0.7810	0.9444
	30	0.1341	0.1804	1.0415	0.5155	0.7845	0.8952	0.8213	0.9487
	50	0.0877	0.1034	0.7575	0.3732	0.8250	0.8996	0.8607	0.9487
	100	0.0405	0.0508	0.4666	0.2536	0.8610	0.8992	0.8962	0.9461
1.5	20	0.1227	0.4380	1.2015	1.0190	0.7410	0.9012	0.7775	0.9481
	30	0.1263	0.2732	1.0473	0.7628	0.7801	0.9053	0.8212	0.9523
	50	0.0815	0.1605	0.7576	0.5618	0.8165	0.8944	0.8518	0.9484
	100	0.0327	0.0816	0.4636	0.3771	0.8566	0.9008	0.8950	0.9510
2.0	20	0.1118	0.6000	1.1865	1.3790	0.7448	0.8983	0.7823	0.9444
	30	0.1194	0.3756	1.0277	1.0406	0.7826	0.8943	0.8204	0.9460
	50	0.0774	0.2200	0.7451	0.7568	0.8156	0.8943	0.8555	0.9473
	100	0.0326	0.1116	0.4582	0.5091	0.8527	0.8964	0.8887	0.9496
3.0	20	0.1373	0.8354	1.2156	1.9506	0.7428	0.9044	0.7793	0.9534
	30	0.1278	0.5612	1.0591	1.5515	0.7807	0.8940	0.8196	0.9461
	50	0.0899	0.3211	0.7738	1.1384	0.8230	0.8921	0.8585	0.9466
	100	0.0411	0.1522	0.4599	0.7648	0.8572	0.9000	0.8968	0.9466
4.0	20	0.1774	0.8789	1.2629	2.2344	0.7719	0.9383	0.8148	0.9850
	30	0.1421	0.6758	1.0647	1.9487	0.7841	0.9039	0.8221	0.9576
	50	0.0870	0.4219	0.7629	1.4985	0.8169	0.8962	0.8518	0.9469
	100	0.0354	0.2130	0.4592	1.0135	0.8551	0.8956	0.8907	0.9467

TABLE 3
Estimated bias, root mean-squared and coverage probability α and β ($\alpha = 1.5$).

		Bias		RMSE		$CP_{90\%}$		$CP_{95\%}$	
β	n	α	β	α	β	α	β	α	β
0.5	20	0.0848	0.2165	1.6318	0.4439	0.7005	0.8911	0.7367	0.9436
	30	0.1462	0.1354	1.4851	0.3287	0.7540	0.9011	0.7908	0.9444
	50	0.1487	0.0746	1.2200	0.2346	0.8004	0.9032	0.8395	0.9566
	100	0.0873	0.0362	0.8104	0.1584	0.8502	0.8989	0.8842	0.9517
1.0	20	0.0005	0.4566	1.5378	0.9033	0.6890	0.8867	0.7288	0.9374
	30	0.0827	0.2909	1.4206	0.6673	0.7402	0.9005	0.7782	0.9448
	50	0.1166	0.1647	1.2175	0.4808	0.7958	0.8981	0.8314	0.9508
	100	0.0782	0.0769	0.8320	0.3201	0.8396	0.8952	0.8748	0.9467
1.5	20	0.0336	0.6511	1.5329	1.3232	0.7066	0.8968	0.7441	0.9431
	30	0.1178	0.4024	1.4466	0.9659	0.7553	0.9087	0.7949	0.9528
	50	0.1346	0.2267	1.1989	0.7021	0.7989	0.9044	0.8338	0.9553
	100	0.0962	0.1040	0.8392	0.4751	0.8432	0.8979	0.8800	0.9492
2.0	20	0.0421	0.8527	1.5534	1.7308	0.7116	0.8933	0.7485	0.9424
	30	0.1108	0.5515	1.4403	1.3153	0.7573	0.9047	0.7937	0.9467
	50	0.1575	0.2981	1.2628	0.9388	0.8028	0.8999	0.8417	0.9532
	100	0.0967	0.1351	0.8397	0.6264	0.8468	0.9029	0.8831	0.9525
3.0	20	0.0985	1.0942	1.6018	2.2813	0.7322	0.9207	0.7684	0.9672
	30	0.1465	0.7630	1.4603	1.8860	0.7618	0.9092	0.8006	0.9525
	50	0.1526	0.4395	1.2360	1.4052	0.8056	0.9036	0.8410	0.9529
	100	0.1031	0.1986	0.8459	0.9465	0.8457	0.9010	0.8809	0.9511
4.0	20	0.1703	1.0804	1.5934	2.4808	0.7818	0.9748	0.8174	0.9975
	30	0.1711	0.8544	1.4581	2.2091	0.7793	0.9368	0.8199	0.9795
	50	0.1499	0.5874	1.2407	1.8267	0.7984	0.9053	0.8364	0.9582
	100	0.0962	0.2780	0.8455	1.2707	0.8414	0.8978	0.8794	0.9484

TABLE 4
Estimated bias, root mean-squared and coverage probability α and β ($\alpha = 2.0$).

		Bias		RMSE		$CP_{90\%}$		$CP_{95\%}$	
β	n	α	β	α	β	α	β	α	β
0.5	20	-0.1570	0.2913	1.7606	0.5391	0.6766	0.8887	0.7169	0.9368
	30	0.0111	0.1877	1.7161	0.3970	0.7238	0.8959	0.7642	0.9439
	50	0.1231	0.1053	1.5482	0.2828	0.7802	0.9094	0.8160	0.9523
	100	0.1572	0.0445	1.2237	0.1874	0.8338	0.9034	0.8700	0.9555
1.0	20	-0.1716	0.5958	1.7730	1.0895	0.6777	0.8818	0.7138	0.9361
	30	-0.0017	0.3853	1.7145	0.8059	0.7165	0.8968	0.7571	0.9416
	50	0.1477	0.2108	1.5957	0.5694	0.7789	0.9075	0.8143	0.9533
	100	0.1426	0.0965	1.2289	0.3792	0.8282	0.8997	0.8620	0.9511
1.5	20	-0.1587	0.8742	1.7863	1.5993	0.6758	0.8890	0.7148	0.9382
	30	-0.0049	0.5649	1.7164	1.1868	0.7239	0.9024	0.7670	0.9453
	50	0.1082	0.3229	1.5554	0.8470	0.7790	0.9100	0.8158	0.9540
	100	0.1431	0.1431	1.2260	0.5646	0.8302	0.9038	0.8669	0.9535
2.0	20	-0.1814	1.1585	1.7289	2.1048	0.6793	0.8884	0.7165	0.9392
	30	0.0073	0.7641	1.7342	1.6076	0.7253	0.8997	0.7611	0.9410
	50	0.1620	0.4153	1.6278	1.1367	0.7788	0.9107	0.8128	0.9516
	100	0.1495	0.1828	1.2140	0.7483	0.8292	0.9054	0.8673	0.9556
3.0	20	-0.0925	1.4311	1.7841	2.6231	0.6998	0.9266	0.7398	0.9785
	30	0.0189	1.0427	1.7108	2.1899	0.7396	0.9092	0.7743	0.9576
	50	0.1329	0.6228	1.5720	1.6689	0.7813	0.9109	0.8178	0.9531
	100	0.1463	0.2837	1.2266	1.1307	0.8304	0.9024	0.8662	0.9507
4.0	20	0.1014	1.2108	1.8102	2.6171	0.7869	0.9951	0.8275	0.9997
	30	0.1248	1.0255	1.7349	2.4043	0.7818	0.9672	0.8239	0.9955
	50	0.1905	0.7097	1.6204	2.0623	0.7928	0.9250	0.8299	0.9677
	100	0.1791	0.3450	1.2812	1.4977	0.8305	0.8988	0.8649	0.9497

5. APPLICATIONS

In this section, we compare the fits of the UG distribution with beta, Kumaraswamy and McDonald (McDonald, 1984) distributions since they are used for modeling bounded data by means of two data sets to illustrate the potentiality of the UG model.

The beta, Kumaraswamy and McDonald distribution have p.d.f. written, respectively, as

(i) beta distribution:

$$f(x;\alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\,x^{\alpha-1}(1-x)^{\beta-1}.$$

(ii) Kumaraswamy distribution:

$$f(x;\alpha,\beta)=\alpha\,\beta\,x^{\alpha-1}(1-x^{\alpha})^{\beta-1}.$$

(iii) McDonald distribution:

$$f(x;\mu,\sigma,\nu,\tau)=\frac{\tau\nu^{\beta}x^{\tau\alpha-1}(1-x^{\tau})^{\beta-1}}{\text{Beta}(\alpha,\beta)[\nu+(1-\nu)x^{\tau}]^{\alpha+\beta}},$$

where in (iii) we have $\alpha=\mu\left(\frac{1}{\sigma^2}-1\right)$ and $\beta=\alpha\left(\frac{1-\mu}{\mu}\right)$ (see, Stasinopoulos *et al.*, 2017).

The data sets are reported in Table 5. The first data set represents 20 observations of the maximum flood level (in millions of cubic feet per second) for Susquehanna River at Harrisburg, Pennsylvania and is reported in Dumonceaux and Antle (1973). The second data set related to 30 measurements of tensile strength of polyester fibers Quesenberry and Hales (1980).

TABLE 5
Flood level and tensile strength data sets.

Data Set I
0.26, 0.27, 0.30, 0.32, 0.32, 0.34, 0.38, 0.38, 0.39, 0.40, 0.41, 0.42, 0.42, 0.42, 0.45, 0.48, 0.49, 0.61, 0.65, 0.74
Data Set II
0.023, 0.032, 0.054, 0.069, 0.081, 0.094, 0.105, 0.127, 0.148, 0.169, 0.188, 0.216, 0.255, 0.277, 0.311, 0.361, 0.376, 0.395, 0.432, 0.463, 0.481, 0.519, 0.529, 0.567, 0.642, 0.674, 0.752, 0.823, 0.887, 0.926

In Table 6 we present the maximum likelihood estimates and their corresponding standard errors in parentheses of the fitted distributions.

For model comparison, we have taken into account the likelihood-based statistics Akaike's Information Criterion (AIC) and Bayesian information criterion (BIC), and

TABLE 6
Maximum likelihood estimates (standard-error).

	Data Set I			
Distribution	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\nu}$	$\widehat{\tau}$
Unit-Gompertz	0.02 (0.02)	4.14 (0.74)	-	-
Beta	6.76 (2.10)	9.11 (2.85)	-	-
Kumaraswamy	3.36 (0.60)	11.79 (5.36)	-	-
McDonald	17.05 (148.57)	2.13 (3.23)	0.01 (0.04)	2.57 (4.38)

	Data Set II			
Distribution	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\nu}$	$\widehat{\tau}$
Unit-Gompertz	1.04 (0.77)	0.43 (0.19)	-	-
Beta	0.97 (0.22)	1.62 (0.41)	-	-
Kumaraswamy	0.96 (0.20)	1.61 (0.41)	-	-
McDonald	103.08 (873.34)	1.20 (0.45)	0.02 (0.16)	0.30 (0.32)

the goodness-of-fit measures Kolmogorov-Smirnov statistic (KS), Anderson -Darling statistic (AD) and Cramér-von Mises statistic (CvM). The statistics with smaller values and larger p-values are said to be better fit. The results for the two data sets are reported in Table 7.

The results in Table 7 revels that both the data sets may be modeled by the four distributions. The superiority of the unit-Gompertz model, in terms of goodness-of-fit statistics, in comparison to the beta and Kumaraswamy distributions, is observed for both the data sets. However, McDonal distribution performs little bit better than unit-Gompertz distribution in terms all considered criteria. It is important to note that the McDonald distribution has four parameters and they are not estimated accurately.

TABLE 7
Goodness-of-fit measures (p*-values) and likelihood-based statistics.*

	Data Set I				
Distribution	KS	CvM	AD	AIC	BIC
Unit-Gompertz	0.15 (0.78)	0.05 (0.88)	0.29 (0.95)	-28.72	-26.73
Beta	0.20 (0.41)	0.12 (0.49)	0.73 (0.53)	-24.13	-22.13
Kumaraswamy	0.21 (0.34)	0.16 (0.35)	0.93 (0.39)	-21.73	-19.74
McDonald	0.13 (0.87)	0.05 (0.90)	0.28 (0.95)	-24.77	-20.79

	Data Set II				
Distribution	KS	CvM	AD	AIC	BIC
Unit-Gompertz	0.07 (0.99)	0.02 (1.00)	0.11 (1.00)	-3.90	-1.10
Beta	0.07 (1.00)	0.02 (1.00)	0.17 (1.00)	-2.61	0.19
Kumaraswamy	0.07 (1.00)	0.02 (1.00)	0.16 (1.00)	-2.62	0.18
McDonald	0.07 (1.00)	0.01 (1.00)	0.10 (1.00)	-0.04	5.57

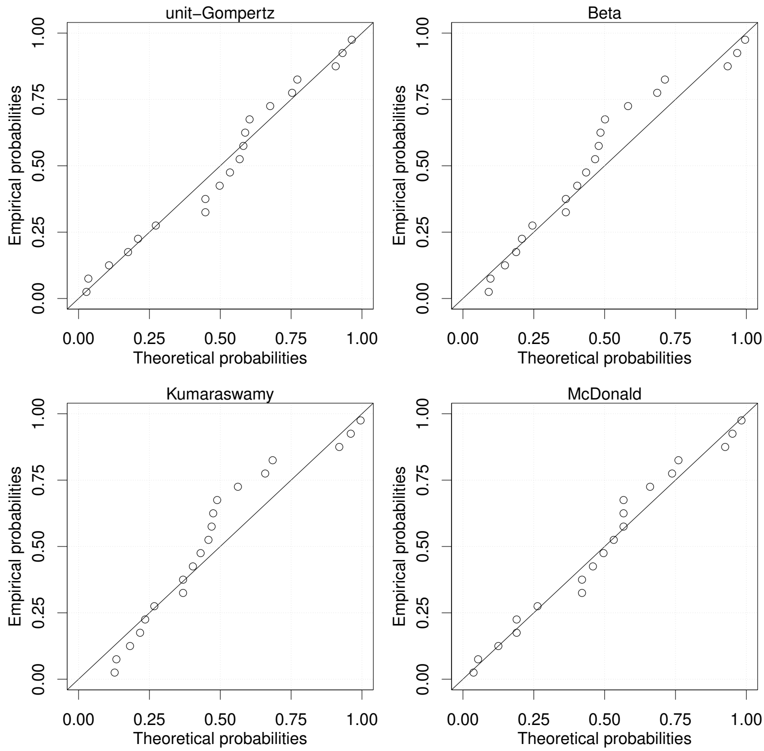

Figure 4 – PP-Plots of the fitted distributions — Data Set I.

Finally, the probability-probability plots for all fitted models are shown in Figures 4 and 5. These plots show that the unit-Gompertz distribution provides the good fit to these data sets compared to the other two models, beta and Kumaraswamy distributions.

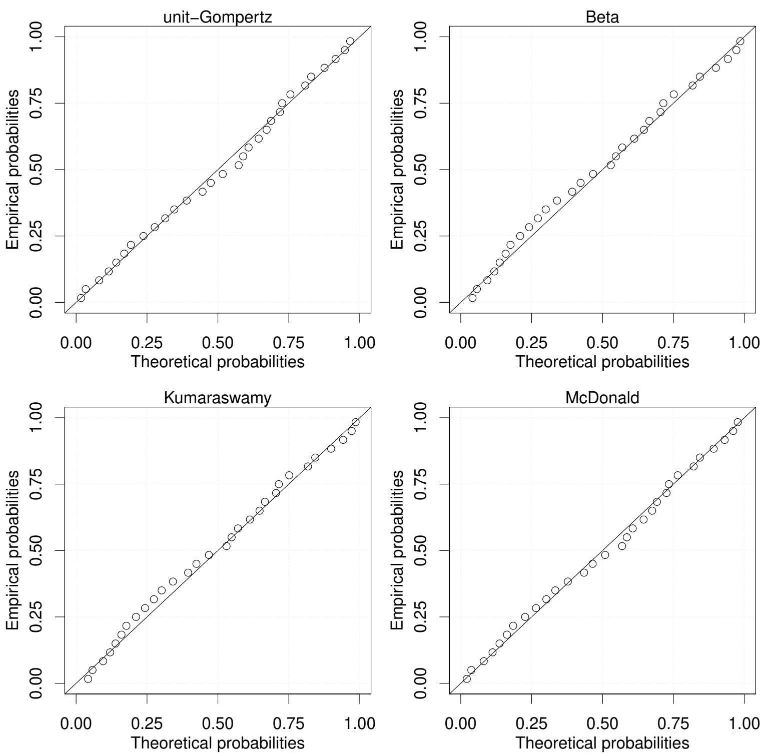

Figure 5 – PP-Plots of the fitted distributions — Data Set II.

6. Conclusion

In this article, a new distribution is proposed which can serve as an alternatives to beta and Kumaraswamy distributions having two shape parameters. The estimation of model parameters are obtained by maximum likelihood estimation. Two applications to real data sets are presented as an illustration of the potentiality of the new model as compared to beta, Kumaraswamy and McDonald models. After comparing the values of five popular goodness-of-fit statistics, we may say that our model UG is better as compared to beta and Kumaraswamy distributions for these two data sets. We expect the utility of the newly proposed model in different fields especially in life-time and reliability when hazard rate is increasing, unimodal (upside-down bathtub) or bathtub.

Acknowledgements

The authors would like to thank the Referees and the Editor for several helpful comments which had improved the earlier versions of the manuscript.

References

A. C. Bemmaor (1994). *Modeling the diffusion of new durable goods: Word-of-mouth effect versus consumer heterogeneity*. In G. Laurent, G. L. Lilien, B. Pras (eds.), *Research Traditions in Marketing*, Springer Netherlands, Dordrecht, pp. 201–229.

L. Benkhelifa (2017). *The beta generalized Gompertz distribution*. Applied Mathematical Modelling, 52, pp. 341–357.

Z. Chen (1997). *Parameter estimation of the Gompertz population*. Biometrical Journal, 39, no. 1, pp. 117–124.

R. Dumonceaux, C. E. Antle (1973). *Discrimination between the log-normal and the Weibull distributions*. Technometrics, 15, no. 4, pp. 923–926.

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

A. El-Gohary, A. Alshamrani, A. N. Al-Otaibi (2013). *The generalized Gompertz distribution*. Applied Mathematical Modelling, 37, no. 1–2, pp. 13 – 24.

P. H. Franses (1994). *Fitting a Gompertz curve*. The Journal of the Operational Research Society, 45, no. 1, pp. 109–113.

E. Gómez-Déniz, M. A. Sordo, E. Calderín-Ojeda (2014). *The log-Lindley distribution as an alternative to the Beta regression model with applications in insurance*. Insurance: Mathematics and Economics, 54, pp. 49 – 57.

A. GRASSIA (1977). *On a family of distributions with argument between 0 and 1 obtained by transformation of the Gamma distribution and derived compound distributions*. Australian Journal of Statistics, 19, no. 2, pp. 108–114.

A. A. JAFARI, S. TAHMASEBI (2016). *Gompertz-power series distributions*. Communications in Statistics - Theory and Methods, 45, no. 13, pp. 3761–3781.

A. A. JAFARI, S. TAHMASEBI, M. ALIZADEH (2014). *The Beta-Gompertz distribution*. Revista Colombiana de Estadística, 37, no. 1, pp. 141–158.

P. KUMARASWAMY (1980). *A generalized probability density function for double-bounded random processes*. Journal of Hydrology, 46, no. 1, pp. 79 – 88.

E. J. LEHMANN, G. CASELLA (1998). *Theory of Point Estimation*. Springer Verlag, New York.

B. G. LINDSAY, B. LI (1997). *On second-order optimality of the observed Fisher information*. The Annals of Statistics, 25, no. 5, pp. 2172–2199.

R. MAKANY (1991). *A theoretical basis for Gompertz's curve*. Biometrical Journal, 33, no. 1, pp. 121–128.

J. B. MCDONALD (1984). *Some generalized functions for the size distribution of income*. Econometrica, 52, no. 3, pp. 647–665.

C. QUESENBERRY, C. HALES (1980). *Concentration bands for uniformity plots*. Journal of Statistical Computation and Simulation, 11, no. 1, pp. 41–53.

B. R. RAO, C. V. DAMARAJU (1992). *New better than used and other concepts for a class of life distributions*. Biometrical Journal, 34, no. 8, pp. 919–935.

C. B. READ (1983). *Gompertz distribution*. In S. KOTZ, N. JOHNSON (eds.), *Encyclopedia of Statistical Sciences*, John Wiley & Sons, Inc., New York, p. 446.

S. ROY, M. A. S. ADNAN (2012). *Wrapped generalized Gompertz distribution: An application to ornithology*. Journal of Biometrics & Biostatistics, 3, no. 6, pp. 153–156.

M. D. STASINOPOULOS, R. A. RIGBY, G. Z. HELLER, V. VOUDOURIS, F. DE BASTIANI (2017). *Flexible Regression and Smoothing: Using GAMLSS in R*. Chapman & Hall/CRC, Boca Raton.

F. WILLEKENS (2001). *Gompertz in context: The Gompertz and related distributions*. In E. TABEAU, A. VAN DEN BERG JETHS, C. HEATHCOTE (eds.), *Forecasting Mortality in Developed Countries: Insights from a Statistical, Demographic and Epidemiological Perspective*, Springer Netherlands, Dordrecht, pp. 105–126.

J.-W. WU, W.-C. LEE (1999). *Characterization of the mixtures of Gompertz distributions by conditional expectation of order statistics*. Biometrical Journal, 41, no. 3, pp. 371–381.

SUMMARY

The transformed family of distributions are sometimes very useful to explore additional properties of the phenomenons which non-transformed (baseline) family of distributions cannot. In this paper, we introduce a new transformed model, called the unit-Gompertz (UG) distribution which exhibit right-skewed (unimodal) and reversed-J shaped density while the hazard rate has constant, increasing, upside-down bathtub and then bathtub shaped hazard rate. Some statistical properties of this new distribution are presented and discussed. Maximum likelihood estimation for the parameters that index UG distribution are derived along with their corresponding asymptotic standard errors. Monte Carlo simulations are conducted to investigate the bias, root mean squared error of the maximum likelihood estimators as well as the coverage probability. Finally, the potentiality of the model is presented and compared with three others distributions using two real data sets.

Keywords: Gompertz distribution; Maximum likelihood estimators; Monte Carlo simulation.