International Journal of Mathematical Analysis and Optimization: Theory and Applications
Vol. 2019, No. 1 , pp. 443 - 454

Properties and Applications of the Gompertz Distribution

J. A., Adewara, [1*] J. S., Adeyeye[2] and C. P., Thron[3]

1 *, Distance Learning Institute, University of Lagos.
2, Department of Mathematics, University of Lagos.
3, Department of Science and Mathematics,Texas A& M University-Central Texas.
*Corresponding Author email: jadewara@unilag.edu.ng

Article Info
Received: 12 January 2019 Revised: 26 February 2019
Accepted: 01 March 2019 Available online: 23 April 2019

Abstract

The importance of statistical distributions in describing and predicting real world events cannot be over-emphasized. The Gompertz distribution is one example of a widely-used distribution, with many applications to survival analysis. In this paper, several properties of the Gompertz distribution are studied. The two-parameter Gompertz distribution is shown to be identical to the three-parameter Gompertz exponential distribution. Functions used in reliability analysis related to the Gompertz distribution are reviewed. Properties of maximum likelihood estimate (MLE) parameter estimates for the Gompertz distribution are studied: the bias and root mean squared error of parameter estimates are expressed as a function of sample size and parameter values. When the Gompertz shape parameter is large, MLE parameter estimates may fail to exist because of parameter degeneracy, as the two-parameter Gompertz distribution approaches a 1-parameter exponential distribution. The distribution is fitted to real life data sets from both industrial and biological applications. Compared to several 3-parameter distributions, the Gompertz distribution provides significantly better fits to the industrial data sets chosen, but the 3-parameter generalized Gompertz distribution gives a better fit to guinea pig lifetime data.

Keywords: Gompertz distribution, Skewed data, Maximum likelihood, Parameters, Reliability, Survival function, Hazard function, Quantile function.
MSC2010: 62N86

1 Introduction

Statistical distributions and their properties are used in modeling naturally occurring phenomena. A large number of distributions have been defined and studied in the literature, which are found to be applicable in real life. The normal distribution addresses real-valued variables that tend to cluster at a single mean value. The Poisson distribution models discrete rare events. [1] studied the Gompertz distribution and calculated the moment generating function in terms of incomplete or complete gamma functions, and their results are either approximate or left in an integral form. The

IJMAO

INTERNATIONAL JOURNAL OF MATHEMATICAL ANALYSIS AND OPTIMIZATION: THEORY AND APPLICATIONS
VOL. 2019, NO. 1, PP. 443 - 454

development of new compound distributions that are more flexible than existing distributions is an important new trend in the theory and application of distributions. For instance, the beta-Gompertz distribution [2] and the generalized Gompertz distribution (GGD) [3] were both introduced to model skewed data; while the Exponentiated Generalized Weibull-Gompertz distribution and the Gompertz- Lomax distribution (which extends the Lomax distribution using the Gompertz family of distributions) were introduced to take care of non-normal data [4]
The exponential distribution is perhaps the most widely applied statistical distribution for reliability studies. The exponentiated Gompertz distribution defined and studied by [5] is obtained by raising the cumulative distribution function (cdf) of the Gompertz distribution to a parameter θ.
The current paper focuses on the two-parameter Gompertz distribution. The paper examines various properties of the distribution, including its relation to the three-parameter Gompertz exponential distribution and the accuracy of maximum likelihood estimates of parameter values. The Gompertz distribution is also applied to both industrial and biological data, and resulting fits are compared to fits for the three-parameter Gompertz distribution, as well as other three-parameter distributions.

1.1 Reduction of Gompertz-exponential distribution to Gompertz distribution

The cumulative distribution function and probability density function of the exponential distribution with parameter λ are given by

$$G(x) = 1 - exp(-\lambda x); \lambda > 0 \tag{1.1}$$

$$G(x) = \lambda exp(-\lambda x); \lambda > 0 \tag{1.2}$$

respectively where λ is referred to as the rate parameter. According to [?] the cdf and pdf of the Gompertz generalized family of distributions are given by

$$F(x) = 1 - exp\left((\theta/\gamma)\left(1 - (1 - G(x))^{-\gamma}\right)\right); \theta > 0, \gamma > 0 \tag{1.3}$$

$$f(x) = \theta g(x)\left[1 - G(x)\right]^{-\gamma - 1} exp\left((\theta/\gamma)\left(1 - (1 - G(x))^{-\gamma}\right)\right); \theta > 0, \gamma > 0 \tag{1.4}$$

where θ and γ are additional shape parameters which are introduced to vary tail weights. G(x) and g(x) are the cdf and pdf of the parent (or baseline) distribution respectively. The pdf of the Gompertz-exponential distribution is derived by inserting the densities in equation (1.1) and (1.2) into equation (1.4)

$$f(x) = \theta\lambda exp(\lambda\gamma x) exp\left((\theta/\gamma)\left(1 - exp\lambda\gamma x\right)\right); x > 0, \theta > 0, \gamma > 0, \lambda > 0 \tag{1.5}$$

The cdf of Gompertz-exponential distribution is derived by inserting the density in equation (1.1) into (1.3),

$$F(x) = 1 - exp\left((\theta/\gamma)\left(1 - exp\lambda\gamma x\right)\right); x > 0, \theta > 0, \gamma > 0, \lambda > 0 \tag{1.6}$$

γ, λ and θ in equation (1.5) and (1.6) may be combined into two independent parameters t, z defined as follows: $z = \theta/\gamma$, $t = \lambda\gamma$. Therefore,

$$f(x) = tzexp(tx)exp\left(z\left[1 - exp(tx)\right]\right); x > 0, t > 0, z > 0 \tag{1.7}$$

$$F(x) = 1 - \exp\left(z\left[1 - \exp(tx)\right]\right); x > 0, t > 0, z > 0 \tag{1.8}$$

equations (1.7) and (1.8) are the pdf and cdf of Gompertz distribution with parameters t and z respectively. From equation (1.8) it follows that $t^{-1}f(x/t)$ is independent of t. This implies that the parameter t only changes the horizontal and vertical scaling of the pdf without affecting the shape of the distribution. The possible shapes of the Gompertz probability may thus be obtained by plotting $t^{-1}f(x)$ as a function of tx for different values of z, as shown in Figure 1

International Journal of Mathematical Analysis and Optimization: Theory and Applications
Vol. 2019, No. 1 , pp. 443 - 454

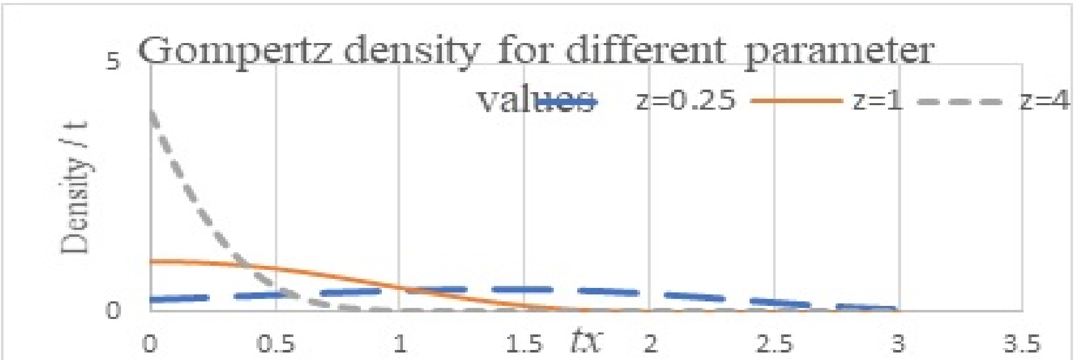

Figure 1: Possible Gompertz density shapes

1.2 Functions used in reliability analysis

The expressions for the reliability function, hazard function (or failure rate), and quantile function for the Gompertz distribution are derived as follows.

1.2.1 Reliability function

The reliability or survival function can be obtained from

$$S(x) = 1 - F(x) \tag{1.9}$$

Therefore, the survival function of the Gompertz distribution is given by

$$S(x) = exp\left(z\left[1 - exp(tx)\right]\right); x > 0, t > 0, z > 0 \tag{1.10}$$

1.2.2 Hazard Function

The hazard function can be obtained from

$$h(x) = f(x)/S(x) \tag{1.11}$$

which implies

$$h(x) = tzexp(tx); x > 0, t > 0, z > 0 \tag{1.12}$$

1.2.3 Quantile Function and Median

The quantile function can be derived from $Q(u) = F^{-1}(u)$. Letting $F(x) = u$, so that $u = F(x) = 1 - exp\left(z\left[1 - \exp(tx)\right]\right)$ and solving for x, we obtain

$$x = (1/t)log\left[1 - (1/z)log(1 - u)\right] \tag{1.13}$$

Therefore,

$$Q(u) = (1/t)log\left[1 - (1/z)log(1 - u)\right] \tag{1.14}$$

where $0 \leq u \leq 1$. The median of the Gompertz distribution can be derived by substituting $u = 0.5$ in equation(1.14)

$$Median = (1/t)log\left[1 - (1/z)log(0.5)\right] \tag{1.15}$$

Other quantiles can also be derived from equation(1.14) by substituting the appropriate values of$u$.

International Journal of Mathematical Analysis and Optimization: Theory and Applications
Vol. 2019, No. 1, pp. 443 - 454

1.3 Maximum likelihood parameter estimation for the Gompertz distribution

The parameters of the Gompertz distribution can be estimated using the method of Maximum Likelihood Estimation (MLE) as follows: let $x_1, x_2, ..., x_n$ denote random samples each having the pdf of the Gompertz distribution, then the likelihood function is given by

$$f(x_1, x_2, ..., x_n; t, z) = \prod_{i=1}^{n} \{tzexp(tx)exp\left(z\left[1 - exp(tx)\right]\right)\}. \tag{1.16}$$

Let l denote the log-likelihood function:

$$l = logf(x_1, x_2, ..., x_n; t, z)$$

Then,

$$l = nlogt + nlogz + t\sum_{i=1}^{n} x_i + z\sum_{i=1}^{n}[1 - exp(tx_i)]$$

Solving $dl/dt = 0$ and $dl/dz = 0$ simultaneously gives the maximum likelihood estimates of parameters t and z.

$$0 = \partial l/\partial t = n/t + \sum_{i=1}^{n} x_i + z\sum_{i=1}^{n} x_i exp(tx_i) \tag{1.17}$$

$$0 = \partial l/\partial z = n/t + \sum_{i=1}^{n}[1 - exp(tx_i)] \tag{1.18}$$

Solving (18) for z leads to

$$z = \frac{-n}{\sum_{i=1}^{n}[exp(tx_i) - 1]} \tag{1.19}$$

Substituting this expression $\frac{-n}{\sum_{i=1}^{n}[1-exp(tx_i)]}$ for z into equation(1.18), we obtain

$$\frac{\sum_{i=1}^{n} tx_i exp(tx_i)}{\sum_{i=1}^{n}[1 - exp(tx_i)]} - (1/n)\sum_{i=1}^{n} tx_i - 1 = 0. \tag{1.20}$$

Through algebraic manipulation, equation (1.21) may be shown to be equivalent to

$$([tx_i exp(tx_i)] - [exp(tx_i)]\,[tx_i]) - [exp(tx_i) - 1 - tx_i] = 0 \tag{1.21}$$

where the square brackets $[\dots]$ denote average: $[y_i] \equiv \sum_{i=1}^{n} y_i$. The solution to equation (1.22) may be found numerically, and then then z may be obtained using (1.20) using the value for t obtained from (1.22). The limiting behavior of the left-hand side of (1.22) when $t \longrightarrow \infty$ and when $t \longrightarrow 0$ may be characterized as follows.
i. When $t \longrightarrow \infty$ the first term in (1.22) dominates all the other terms, so the left side of (1.22) is always positive when $t \longrightarrow \infty$.
ii. When $t \longrightarrow 0$, we may replace the exponential terms in (1.22) with their Taylor series and solve for the lowest order terms. This gives the following expression:

$$\left[tx_i + t^2x_i^2 + O(t^3)\right] - [tx_i]\left[1 + tx_i + O(t^2)\right] - \left[\frac{1}{2}t^2x_i^2 + O(t^3)\right] \tag{1.22}$$

which simplifies to

$$\frac{1}{2}t^2\left(\left([x_i^2] - [x_i]^2\right) - [x_i]^2\right) + O(t^3) \tag{1.23}$$

We may recognize $\left(\left([x_i^2] - [x_i]^2\right)\right.$ as the variance of the sample x_i. Accordingly, when $t \longrightarrow 0$ the left-hand side of (21) is positive or negative depending on whether or not the standard deviation of the sample x_i is larger or smaller than the sample mean.
The above results imply that the left-hand side of (1.22) has a sign change somewhere on the positive t axis as long as the standard deviation of the sample x_i is smaller than the sample mean. Since the expression in (1.22) is a continuous function of t, we conclude that in this case there must be a root. where the sign of (1.22) changes from negative to positive, which implies that the root corresponds to a local minimum of the log likelihood. It follows that if the standard deviation of the sample x_i is smaller than the mean, then positive MLE parameters will always exist; but if the standard deviation is not smaller than the mean, then MLE parameters may not exist. (Note that $t = 0$ is also a solution to (1.22), but this leads to the result that $z = \infty$ from (1.20).)
In our empirical studies (see Section 2.1) we observed that when simulated data was generated using a Gompertz distribution with large values of z, then Equation (1.22) frequently failed to have a solution. This may be explained as follows. When $z >> 1$, then the exponent $z(1 - exp(tx))$ in the Gompertz cdf expression (1.8) will be large in magnitude and negative in sign unless $tx << 1$. In this case, $exp(tx)$ is closely approximated by its Taylor series: $exp(t_i) \simeq 1 + tx + O(t^2x^2)$. Replacing this in (1.8) gives the result that

$$F(x) \simeq 1 - exp(-ztx), x > 0, t > 0(when z >> 1) \tag{1.24}$$

It follows from (1.24) that there is degeneracy in the distribution parameters when $z >> 1$: if $z >> 1$, then any parameter set (t',z') with $z'/t' = z/t$ will produce nearly the same distribution as the parameter set (t, z). This degeneracy is what leads to the non-existence of solutions to the MLE equation. Notice that for an exponential distribution, the standard deviation is equal to the mean: so according to the result of the previous paragraph, then MLE parameter estimates may not exist for samples drawn from a distribution that is exponential (or nearly exponential).

2 Materials and Methods

2.1 Empirical study of Gompertz distribution maximum likelihood estimates

In order to evaluate the bias and variance of MLE parameter estimates as a function of sample size and actual parameter values, R software was used to generate samples of size $50, 100, 150, 200, 300, and 400$ with a replication $m = 100000$, from Gompertz distributions with parameter values $t = 1$ and $z = 0.125, 0.25, 0.5, 1, 2, and 4$. In the simulations t was not varied: since, t changes only the scale and not the shape of the distribution, it follows that results for different t values may be directly inferred from the results for $t = 1$. As mentioned in Section 2.3, for some samples the MLE equations had no solutions: when this occurred, the random sample was regenerated so that a total of 100000 estimates of t and z were used in each bias and variance calculation.

2.2 Comparison of fits to industrial data

A comparison was made of different distribution fits using data on the strengths of $1.5cm$ glass fibres of workers at the UK national Physical Laboratory. The data has previously been used by [7], [8] and [4]. The second data represents the lifetimes of 50 devices, drawn from [9] and also used by [3]. For both datasets, the 2-parameter Gompertz fit was compared with 3-parameter Kumaraswamy-Exponential, Generalized Gompertz, and three-parameter Lindley fits.

2.3 Survival analysis of biological lifetime data

The flexibility of Gompertz distribution in analyzing lifetime data was compared to that of the Generalized Gompertz distribution. The parameters of each distribution were first estimated using

R software via the maximum likelihood method. The values of the parameters were then inserted into the survival and hazard functions of each distribution together with the values of the random variable X. The data representing the survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli, observed and reported by [10] and subsequently used by [11].

3 Results

3.1 Gompertz maximum likelihood estimators' statistics

This section presents results for the empirical study of MLE parameter estimation described in Section 2.1

Empirical biases in MLE estimates for t and z shown in the log-log plots in Figure 2. Each line in the plot corresponds to a different actual z value for the underlying sampled distribution: as noted in Section 2.1, $t = 1$ was used for all sampled distributions. The straight-line dependence of the biases in z and t indicates that bias is inversely proportional to sample size.

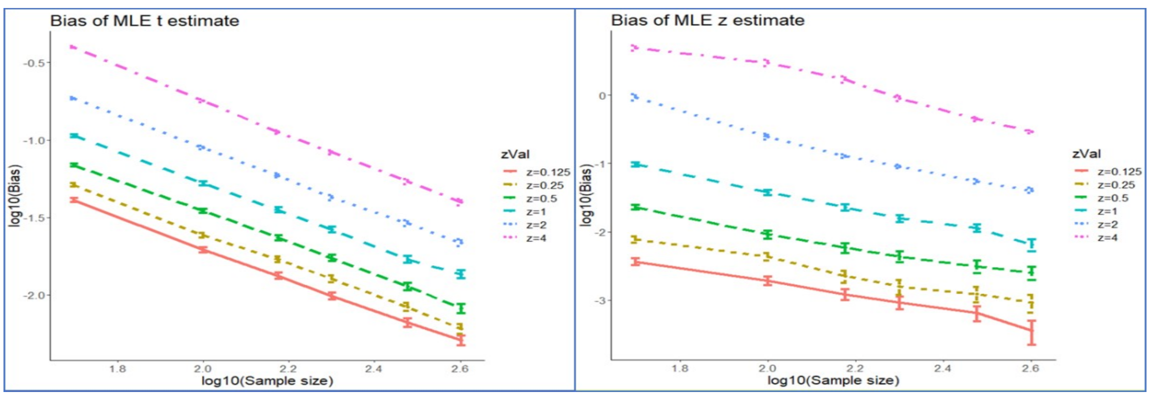

Figure 2: Log-log plots of empirical biases for MLE estimates of t and z as a function of sample size. Different curves are for different actual values of z for the sampled distribution (indicated by 'zVal' in the legend), while the actual value of t was 1 for all cases. Error bars show plus or minus two standard deviations in the bias estimates.

Figure 3 gives log-log plots of bias as a function of the parameter z. It is evident that biases increase as the parameter z increases: as indicated in Section 2.3, increasing values of z correspond to increasing parameter degeneracy, so that different parameter sets give rise to very similar distributions. The slopes of the curves for t biases start less than 1 but tend towards 1 for larger values of z, indicating sublinear dependence; while the slopes of the curves for z bias are larger, indicating superlinear dependence of z bias on the parameter z.

Figure 4 shows the root mean squared error (RMSE) and standard deviations for MLE estimates for t and z as a function of sample size. Error bars at each data point show RMSE (top value) and standard deviation (lower value): the error bars are scarcely visible, indicating little difference between the two values. In the log-log plots, the lines have a slope of $-1/2$, indicating that the RMSEs vary as the inverse square root of the sample size (as is usually the case with standard deviations of estimators). However, for large values of z the RMSE of the MLE z estimates for small sample size become very large: this reflects the instability of z parameter estimation when z is large and sample size is small, due to the parameter degeneracy discussed in Section 2.3.

Figure 5 shows the root mean squared error (RMSE) for MLE estimates for t and z as a function of parameter z. The plots show that parameter estimates become increasingly inaccurate as the

International Journal of Mathematical Analysis and Optimization: Theory and Applications
Vol. 2019, No. 1 , pp. 443 - 454

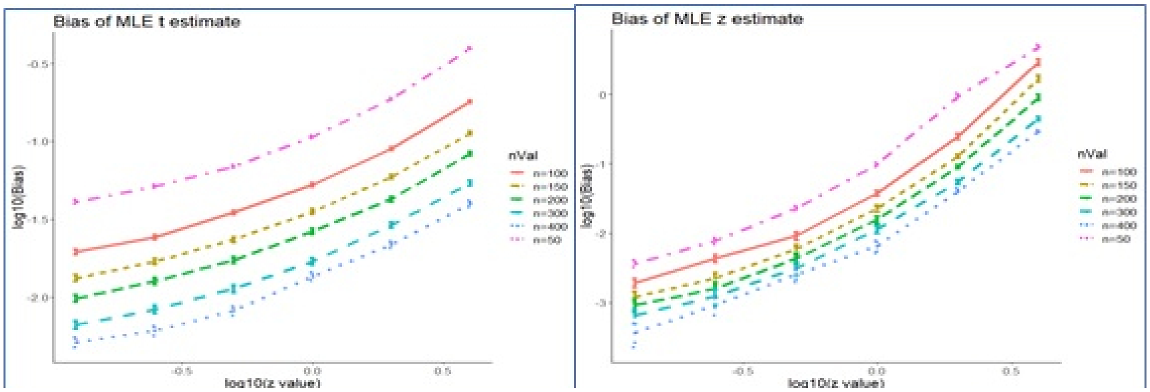

Figure 3: Log-log plots of empirical bias for MLE estimates of t and z as a function of z. Error bars show plus or minus two standard deviations in the empirical bias estimates.

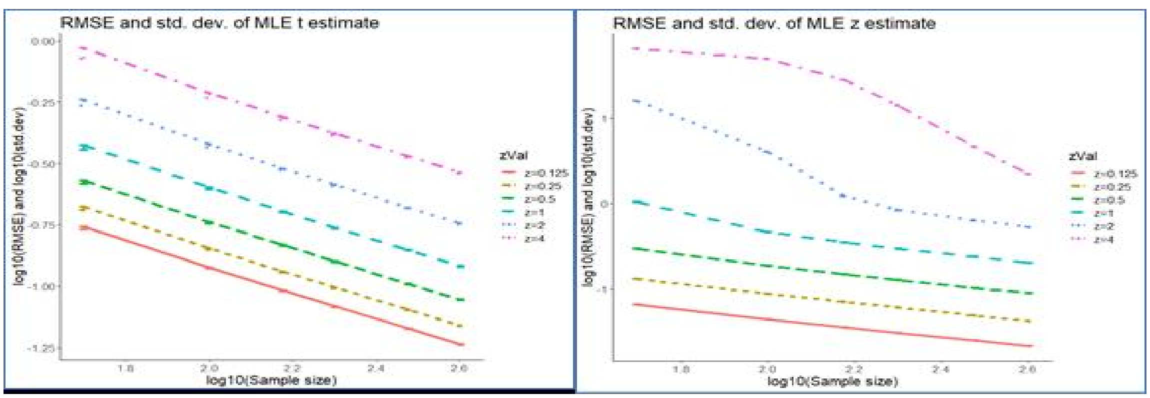

Figure 4: Log-log plots of empirical RMSE and standard deviations for MLE estimates for t and z, as a function of sample size. Error bars at each data point show RMSE (top value) and standard deviation (lower value): the error bars are scarcely visible, indicating little difference between the two.

value of z increases. MLE estimates of z for large values of z are very inaccurate, due to the parameter degeneracy alluded to earlier.

In general, the figures show that the bias for MLE parameter estimates is a small faction of the RMSE, so bias correction would provide minimal improvement over the MLE estimates of t and z.

3.2 Distribution fits to industrial data

This section presents results of the comparisons between distributional fits of industrial data described in Section 3.2.

3.2.1 Descriptive statistics for industrial datasets

Tables 1 and 2 show the descriptive statistics of the two industrial datasets used, from [7] and [9], respectively.

Table 1 shows that the data has a moderate negative skewness (by comparison, the exponential

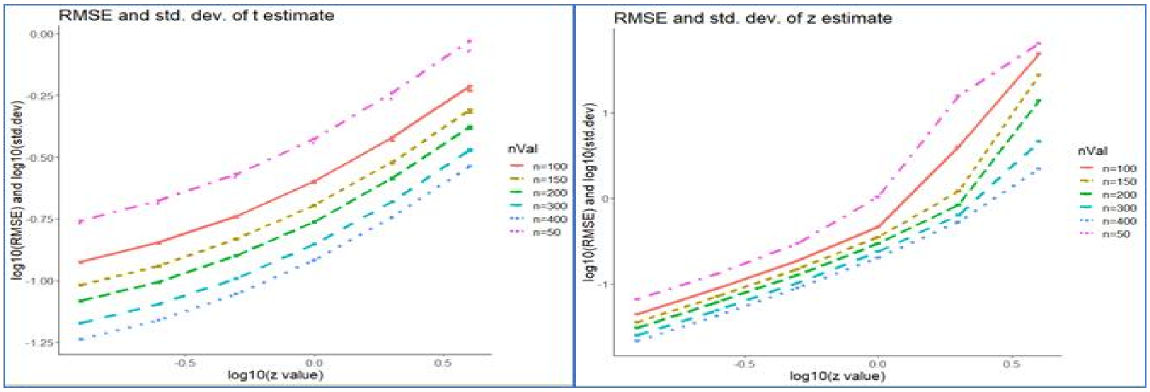

Figure 5: Log-log plots of empirical RMSE errors of MLE estimates for t and z, as a function of sample size.

Table 1: Descriptive statistics for strengths of 1.5cm glass fibres (from [7]

Min.	Q_1	Q_2	Q_3	Mean	Max.	Variance	Skewness	Kurtosis
0.55	1.375	1.590	1.685	1.507	2.240	0.1051	-0.8999	3.9238

distribution has a skewness of 2), and excessive kurtosis of about 0.9, implying that the tails are somewhat heavier than a normal distribution.

Table 2: Descriptive Statistics for lifetimes of 50 devices (from [9]

Min.	Q_1	Q_2	Q_3	Mean	Max.	Variance	Skewness	Kurtosis
0.10	13.50	48.50	81.25	45.69	86.00	1078.15	-0.1378	1.4139

Table 2 shows that the data has excessive kurtosis of about -1.6, implying thinner tails than the normal distribution.

3.2.2 Performance comparisons for different distribution fits to industrial datasets

Table 3 shows the performance of different distribution fits to the glass fibre data.

The distribution with the lowest AIC is judged to be the best out of the competing distributions. With this regard, the competing distributions can be ranked in the following order (best to the least): Two- Parameter Gompertz exponential distribution, Kumaraswamy Exponential distribution, Generalized Gompertz distribution and Three-Parameter Lindley distribution.
Table 4 shows the performance of different distribution fits to the device lifetime data.

Once again, the Gompertz distribution has the lowest AIC, this time followed in order by Generalized Gompertz, Kumaraswamy Exponential, and Three-Parameter Lindley.

3.3 Biological survival analysis results

This section presents results from distributional fits of biological data described in Section 3.3. Using the guinea pig survival data from [10] the maximum likelihood Gompertz parameters are$t =$

Table 3: Performance of Compared Distributions using data on 1.5cm glass fibres

Distributions	Parameter Estimates	-Log likelihood	AIC
Gompertz	$t = 3.6474$ $z = 0.002417$	14.8081	33.6162
Kumaraswamy-Exponential	$\theta = 1756$ $\lambda = 7.001$ $\gamma = 0.2599$	15.9137	37.8274
Generalized Gompertz	$\theta = 0.49258$ $\lambda = 0.48005$ $\gamma = 0.54581$	83.20132	172.4026
Three-Parameter Lindley	$\theta = 0.499886$ $\lambda = 0.002048$ $\gamma = 0.01903$	102.9163	211.8325

Table 4: Performance of Compared Distributions using data on Lifetimes of 50 devices

Distributions	Parameter Estimates	-Log likelihood	AIC
Gompertz	$t = 0.47858$ $z = 0.02030$	235.3308	474.6617
Generalized Gompertz	$\theta = 0.00143$ $\lambda = 0.044$ $\gamma = 0.2599$	235.3920	476.7840
Kumaraswamy-Exponential	$\theta = 0.12631$ $\lambda = 0.46598$ $\gamma = 0.15839$	238.4378	482.8756
Three-Parameter Lindley	$\theta = 0.004952$ $\lambda = 0.010351$ $\gamma = 0.0001513$	322.0525	650.105

$0.0044273, z = 0.6729702$, while for the Generalized Gompertz distribution, the maximum likelihood estimated parameters are $\theta = 3.37335, \lambda = 0.010394451$ and $\gamma = 0.000421141$. The results of the survival analysis of the guinea pigs are tabulated in Table 5 and displayed graphically in Figure 6.

Figure 6 shows that the generalized Gompertz fit is visibly superior to the Gompertz distribution fit in the case of guinea pig data. The Gompertz fit fails to detect the inflection point around 150 days, while the generalized Gompertz fit does a much better job.
The hazard functions estimated from Gompertz and generalized Gompertz distributions are tabulated in Table 6, and displayed graphically in Figure 7.

.

Figure 7 shows that the hazard functions for Gompertz and generalized Gompertz fits are quite different. The Gompertz distribution estimates much greater hazards at small and large times, while in mid-range the two hazard functions are in fairly good agreement.

4 Discussion

The results of the research described in this paper may be summarized as follows:
The Gompertz distribution was shown to be identical with the 3-parameter Gompertz-exponential

International Journal of Mathematical Analysis and Optimization: Theory and Applications
Vol. 2019, No. 1, pp. 443 - 454

Table 5: Survival analysis for Gompertz and Generalized Gompertz distributions using guinea pig data

Time x(days)	Survivial counts			Survival rates S(x)		
	Actual	Gomp.	Gen. Gomp.	S(x)	Gomp.	Gen. Gomp.
10	71	70	72	0.98611	0.9699955	0.9995919
50	68	61	68	0.94444	0.8464131	0.9511673
100	61	49	55	0.84722	0.6874185	0.7611461
150	36	38	38	0.5	0.5302375	0.5289144
200	23	28	24	0.31944	0.3835151	0.3359761
250	14	18	14	0.19444	0.255997	0.2015553
300	8	11	8	0.11111	0.1545932	0.1163765
350	5	6	5	0.06944	0.08238942	0.06537859
400	4	3	3	0.05556	0.03756901	0.03595709
450	2	1	1	0.02778	0.01410215	0.01942711
500	1	0	1	0.01389	0.00415241	0.01033027

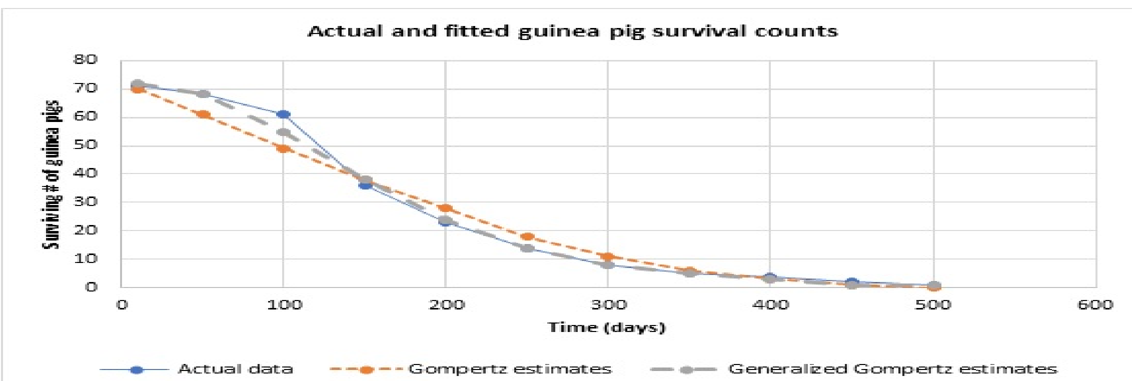

Figure 6: Actual and fitted guinea pig survival count data

distribution. Typically, introducing additional parameters into a distribution typically increases its flexibility, but at the cost of increased mathematical complexity and higher AIC and BIC values that are used in model fitting. Our result shows that no benefit is gained in utilizing the 3-parameter Gompertz-exponential distribution, and the simpler 2-parameter formula should be used instead. Results on MLE parameter estimation for Gompertz distributions included the following:

- When Gompertz distributions with large values of z are sampled, the MLE parameter estimates for z and t may fail to exist. We explained this result by showing mathematically that Gompertz distributions for large z values approach exponential distributions, and the two parameters z and t become degenerate since the exponential distribution is a 1-parameter family.
- There are statistically significant biases in MLE parameter estimates based on samples taken from Gompertz distributions. Biases increase with decreasing sample sizes, and with increasing values of z. Because of these biases, the RMSE values for the parameter estimates are larger than the corresponding standard deviations. However these differences are very slight, indicating that bias correction would produce very little improvement in the parameter

International Journal of Mathematical Analysis and Optimization: Theory and Applications
Vol. 2019, No. 1 , pp. 443 - 454

Table 6: Hazard functions of Gompertz and generalized Gompertz distribution

Time x (days)	h(x) of Gompertz	h(x) of generalized Gompertz
10	0.003114313	0.00001295406
50	0.003717683	0.001087316
100	0.004638846	0.003969643
150	0.005788254	0.006653175
200	0.007222461	0.008616968
250	0.009012033	0.009961473
300	0.01124502	0.01087898
350	0.0140313	0.01152673
400	0.01750796	0.01201241
450	0.02184606	0.01240443
500	0.02725905	0.0127446

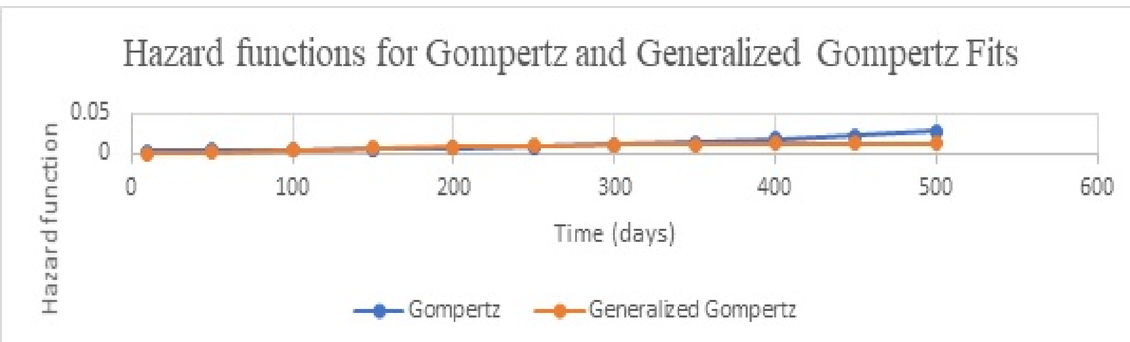

Figure 7: Hazard functions for Gompertz and generalized Gompertz fits

estimates.

- The RMSE values for parameter estimates decrease with increasing sample size as expected, and also increase with increasing z, which reflects the degeneracy into the 1-parameter family of exponential distributions alluded to above.

Results on fitting performance of the Gompertz distribution compared to various 3-parameter distributions were mixed. For two industrial data sets, The Gompertz fits were superior as determined by AIC, showing that the Gompertz distribution can provide good fits with a reduced number of parameters. For guinea pig lifetime data, the generalized Gompertz distribution was better able to approximate the S-shaped distribution. The Gompertz distribution is limited in the shapes it can assume, because as shown in Figure 1 only the z parameter affects the shape of the distribution, while the t distribution only changes the scale.

5 Conclusions

The Gompertz distribution serves as a useful intermediate alternative between the 1-parameter exponential distribution and more sophisticated 3-parameter distributions. In some cases, the

Gompertz distribution gives better fits to lifetime data than 3-parameter alternatives. When the shape parameter z is large, MLE parameter estimates are unreliable, and the exponential distribution should be used instead. Although MLE parameter estimates are biased, the bias is much smaller than the RMSE of the parameter estimates, so bias correction will not effectively improve the estimates.

Acknowledgments

We wish to thank the anonymous reviewer of this paper, whose careful reading and thoughtful comments have contributed greatly to the presentation of this paper.

Competing financial interests

The authors declare no competing financial interests.

References

[1] Pollard, J. H. & Valkovics, E. J. The Gompertz distribution and its applications. *Genus*, 40(3), pp. 15-28, (1992).

[2] Jafari A.A., Tahmasebi S, & Alizadeh M . The Beta Gompertz distribution ,*Revista Colombiana de Estadistica*, 37(1), pp 141 -158, (2014).

[3] El-Gohary A., Alshamrani A, & Naif Al-Otaibi A. The generalized Gompertz distribution, *Applied Mathematics Modelling* 37(1-2) pp. 13- 24,(2013).

[4] Oguntunde, P. E., Khaleel, M. A., Ahmed, M. T., Adejumo, A. O. & Odetunmibi, O. A. A New Generalization of the Lomax Distribution with Increasing, Decreasing, and Constant Failure Rate. Hindawi: *Modelling and Simulation in Engineering (online journal)*, Article ID 6043169, 6 pages, (2017).

[5] Abu-Zinadah H. H. Some Characterizations of the Exponentiated Gompertz Distribution, *International Mathematics Forum*, Volume 9, no 30, 1427-1439, (2014).

[6] Alizadeh, M., Cordeiro, G.M., Pinho, L.G.B. & Ghosh, I. The Gompertz − G family of Distributions, *Journal of Statistics Theory and Practice*,11(1), 179-207, DOI:10.1080/15598608.2016.1267668,(2016).

[7] Smith, R. L. & Naylor, J. C. A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *Applied Statistics* 36(3) 358-369, (1987).

[8] Bourguignon, M., Silva, R. B. & Cordeiro, G. M. The Weibull-G family of probability distributions. *Journal of Data Science*, 12(1), 53-68, (2014).

[9] Aarset, M. V. How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, 36(1), 106-108, (1987).

[10] Bjerkedal, T. Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Am.J. Hygiene*, 72:130-148, (1960).

[11] Rama S., Kamlesh K. S., Ravi S. & Tekie A. L. A Three- Parameter Lindley Distribution, *American Journal of Mathematics and Statistics* 7(1): 15 − 26, DOI: 10.5923/j.ajms.20170701.03, (2017).