

**A COMPARATIVE STUDY OF VARIOUS METHODS
OF ESTIMATION FOR GOMPERTZ-LINDLEY DISTRIBUTION**

Fatmah Almathkour¹, A. Alothman¹, M.E. Ghitany¹ §,
Ramesh C. Gupta², J. Mazucheli³

¹Department of Statistics and Operations Research
Faculty of Science, Kuwait University, KUWAIT

²Department of Mathematics and Statistics
University of Maine, Orono, Maine, USA

³Department of Statistics
Universidade Estadual de Maringà
Maringà, PR, BRAZIL

Abstract: This paper deals with a comparative study of different methods of estimation for Gompertz-Lindley distribution. Simulation studies are carried out and the most efficient estimator is the one whose bias is close to zero with smaller mean-square error. A real data set is analyzed to illustrate the different procedures.

AMS Subject Classification: 62N02, 62N03, 62N05, 62F99

Key Words: estimation methods; simulations; reliability

1. Introduction

In order to estimate the parameters of a model, different methods of estimation have been proposed in the literature. The most popular is the method of maximum likelihood. Comparative studies of various methods of estimation have been carried out for different models. It has been observed that a particular estimation procedure outperforms the others for a particular model.

Received: March 3, 2022

© 2022 Academic Publications

§Correspondence author

For example, [12] proposed a new estimation method for the Weibull distribution based on TL moments and L moments. Their estimation procedure gives the best overall performance with respect to the biases of the estimators of the shape and scale parameters. The use of a particular estimation method depends on the area of application and the criteria of performance. For example, a user may like to prefer a minimum variance unbiased estimator even though it does not have a close form expression and it is minimum variance among the unbiased estimators. [2] compared various estimation procedures for the weighted exponential distribution and recommend the use of the maximum likelihood or Bayes estimators for the model. [3] compared different estimators for Rayleigh Distribution mainly with respect to the biases and mean-squared errors. They conclude that Bayes estimators with non informative priors work very well. Other estimation methods for the generalized exponential distribution, weighted Lindley distribution, and Marshall-Olkin extended exponential distribution can be found in [7], [9], and [8], respectively.

Compound distributions provide a tool for obtaining new parametric families of distributions in terms of existing ones, see for example [1], [5] and [6].

In the present work, we have compared different methods of estimation for Gompertz-Lindley (GL) distribution. The GL distribution was proposed by [4] by compounding the frailty parameter of the Gompertz distribution by Lindley distribution. Different methods of estimation are considered in Section 3. They include method of moments, method of maximum likelihood, method of maximum product of spacings, method of ordinary least squares, method of weighted least squares, method of percentiles, method of L_2 distance and method of Kullbeck-Lieber divergence of survival functions. Monte Carlo simulation studies are carried out in Section 4 to compare the performance of these estimators. In Section 5, we analyze a real data reported in [10] consisting of 213 observations on the number of successive failures of the air conditioning system of a fleet of 13 Boeing 720 jet planes. [4] fitted the GL distribution to this data set using only the method of maximum likelihood. Finally, some comments and conclusion are provided in Section 6.

2. The Gompertz-Lindley distribution

A continuous random variable X is said to have a Gompertz-Lindley (GL) distribution with shape parameter α and scale parameter λ , denoted by $GL(\alpha, \lambda)$,

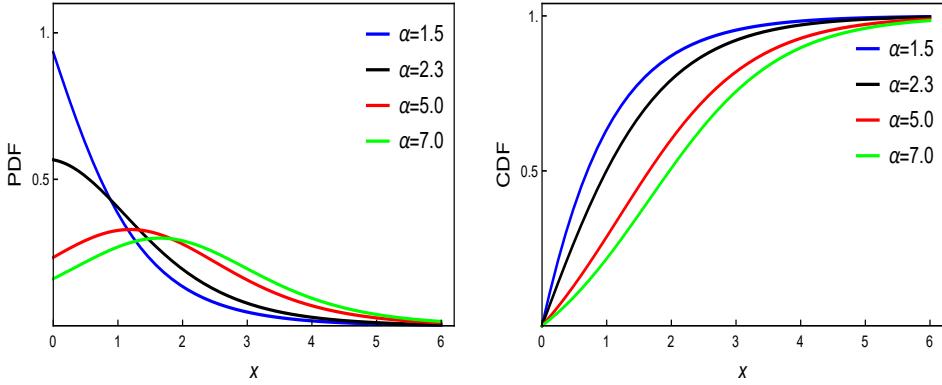


Figure 1: PDF and CDF of the GL distribution as a function of α when $\lambda = 1$.

if its probability density function (PDF) is given by

$$f(x; \alpha, \lambda) = \frac{\alpha^2 \lambda}{\alpha + 1} \frac{e^{\lambda x} (e^{\lambda x} + \alpha + 1)}{(e^{\lambda x} + \alpha - 1)^3}, \quad x > 0, \quad (1)$$

where $\alpha, \lambda > 0$.

The following is a summary of some properties of the GL distribution presented in [4].

- (i) The PDF (1) is decreasing in x for all $\alpha \leq 1$ and unimodal for $\alpha > 1$.
- (ii) The cumulative distribution function (CDF) of the GL distribution is given by

$$F(x; \alpha, \lambda) = 1 - \frac{\alpha^2 (e^{\lambda x} + \alpha)}{(\alpha + 1)(e^{\lambda x} + \alpha - 1)^2}, \quad x > 0. \quad (2)$$

- (iii) The quantile function (QF) of GL distribution is given by

$$F^{-1}(q; \alpha, \lambda) = \frac{1}{\lambda} \ln \left(\frac{\alpha^2 + 2(1 - \alpha^2)\bar{q} + \alpha \sqrt{\alpha^2 + 4(1 + \alpha)\bar{q}}}{2(1 + \alpha)\bar{q}} \right), \quad (3)$$

where $\bar{q} = 1 - q$ and $0 < q < 1$.

In particular, the median of GL distribution is given by

$$\text{median}(X) = \frac{1}{\lambda} \ln \left(\frac{1 + \alpha \sqrt{\alpha^2 + 2\alpha + 2}}{1 + \alpha} \right). \quad (4)$$

- (iv) The r th moment (about 0) of GL distribution is given by:

$$E(X^r) = \frac{r! \alpha^2}{\lambda^r (\alpha + 1)(1 - \alpha)^2} [\text{Li}_{r-1}(1 - \alpha) - \alpha \text{Li}_r(1 - \alpha)]. \quad (5)$$

where

$$\text{Li}_s(z) = \frac{z}{\Gamma(s)} \int_0^\infty \frac{t^{s-1}}{e^t - z} dt, \quad s > 0, \quad -\infty < z < 1, \quad (6)$$

is the *polylogarithm function*. Note that,

$$\text{Li}_1(z) = -\ln(1-z), \quad \text{Li}_0(z) = z \frac{\partial \text{Li}_1(z)}{\partial z} = \frac{z}{1-z}.$$

In particular, the first two moments, respectively, are given by

$$E(X) = \frac{\alpha [1 - \alpha + \alpha^2 \ln(\alpha)]}{\lambda(\alpha + 1)(1 - \alpha)^2}, \quad (7)$$

$$E(X^2) = \frac{2\alpha^2}{\lambda^2(\alpha + 1)(1 - \alpha)^2} [-\ln(\alpha) - \alpha \text{Li}_2(1 - \alpha)]. \quad (8)$$

The R functions `dGL`, `pGL` and `qGL` in Appendix can be used to calculate the PDF, CDF and QF of the GL distribution, respectively.

3. Methods of Estimation

In this section, we describe the eight estimation methods considered in this paper for estimating the unknown parameters α and λ of the GL distribution.

Let x_1, x_2, \dots, x_n be a random sample from the $\text{GL}(\alpha, \lambda)$ distribution with PDF (1).

3.1. Method of Maximum Likelihood

The maximum likelihood (ML) estimates $\hat{\alpha}_{ML}$ and $\hat{\lambda}_{ML}$ of the parameters α and λ , are obtained by maximizing, with respect to α and λ , the log-likelihood function

$$\ell(\alpha, \lambda) = \sum_{i=1}^n \ln f(x_i; \alpha, \lambda). \quad (9)$$

3.2. Method of Moments

The method of moments (MM) estimates $\hat{\alpha}_{MM}$ and $\hat{\lambda}_{MM}$ of the parameters α and λ , are obtained by solving the system of equations:

$$E(X)|_{(\alpha, \lambda) \leftrightarrow (\hat{\alpha}_{MM}, \hat{\lambda}_{MM})} = m_1,$$

$$E(X^2)|_{(\alpha, \lambda) \leftrightarrow (\hat{\alpha}_{MM}, \hat{\lambda}_{MM})} = m_2,$$

where $m_1 = \frac{1}{n} \sum_{i=1}^n x_i$ and $m_2 = \frac{1}{n} \sum_{i=1}^n x_i^2$ are the first and second sample moments, respectively.

It follows that

$$\hat{\lambda}_{MM} = \frac{\hat{\alpha}_{MM}[1 - \hat{\alpha}_{MM} + \hat{\alpha}_{MM}^2 \ln(\hat{\alpha}_{MM})]}{m_1(1 + \hat{\alpha}_{MM})(1 - \hat{\alpha}_{MM})^2}, \quad (10)$$

where $\hat{\alpha}_{MM}$ is the solution of the non-linear equation

$$2 m_1^2 (\alpha + 1)(1 - \alpha)^2 [\alpha \operatorname{Li}_2(1 - \alpha) + \ln \alpha] + m_2 [1 - \alpha + \alpha^2 \ln \alpha]^2 = 0. \quad (11)$$

3.3. Method of maximum product of spacings

Cheng and Amin (1979,1983) introduced the maximum product of spacings (MPS) method as an alternative to MLE for the estimation of parameters of continuous univariate distributions.

Define the uniform spacings of a random sample from the GL distribution as

$$D_i(\alpha, \lambda) = F(x_{i:n}; \alpha, \lambda) - F(x_{i-1:n}; \alpha, \lambda), \quad i = 1, 2, \dots, n, \quad (12)$$

where $x_{i:n}$, $i = 1, 2, \dots, n$, is the i th order statistic of a random sample x_1, x_2, \dots, x_n . Note that $x_{0:n} = 0$ and $x_{n+1:n} = \infty$. The maximum product of spacings (MPS) estimates $\hat{\alpha}_{MPS}$ and $\hat{\lambda}_{MPS}$ of the parameters α and λ , can be obtained by maximizing, with respect to α and λ , the log-geometric mean of the spacings function:

$$M(\alpha, \lambda) = \frac{1}{n+1} \sum_{i=1}^{n+1} \ln D_i(\alpha, \lambda). \quad (13)$$

Cheng and Amin (1983) showed that maximizing H as a method of parameter estimation is as efficient as MLE estimation and the MPS estimators are consistent under more general conditions than the MLE estimators.

3.4. Method of Ordinary Least-Squares

The ordinary least-squares estimates $\hat{\alpha}_{OLS}$ and $\hat{\lambda}_{OLS}$ of the parameters α and λ can be obtained by minimizing, with respect to α and λ , the function

$$S(\alpha, \lambda) = \sum_{i=1}^n \left[F(x_{i:n}; \alpha, \lambda) - \frac{i}{n+1} \right]^2. \quad (14)$$

3.5. Method of Weighted Least-Squares

The weighted least-squares estimates $\hat{\alpha}_{WLS}$ and $\hat{\lambda}_{WLS}$ of the parameters α and λ can be obtained by minimizing, with respect to α and λ , the function

$$W(\alpha, \lambda) = \sum_{i=1}^n \frac{(n+1)^2 (n+2)}{i(n-i+1)} \left[F(x_{i:n}; \alpha, \lambda) - \frac{i}{n+1} \right]^2. \quad (15)$$

3.6. Method of Percentiles

The percentile estimates $\hat{\alpha}_{PC}$ and $\hat{\lambda}_{PC}$ of the parameters α and λ can be obtained by minimizing, with respect to α and λ , the function

$$C(\alpha, \lambda) = \sum_{i=1}^n \left[x_{i:n} - F^{-1} \left(\frac{i}{n+1}; \alpha, \lambda \right) \right]^2. \quad (16)$$

3.7. Method of L_2 distance

The L_2 distance estimates $\hat{\alpha}_{L2}$ and $\hat{\lambda}_{L2}$ of the parameters α and λ can be obtained by minimizing, with respect to α and λ , the function

$$L(\alpha, \lambda) = \int_0^\infty f^2(t; \alpha, \lambda) dt - \frac{2}{n} \sum_{i=1}^n f(x_i; \alpha, \lambda). \quad (17)$$

For more details about this estimation method, see [13].

3.8. Method of Kullback-Leibler divergence of survival function

The Kullback-Leibler divergence of Survival function (KLS) is given by

$$\begin{aligned} K(\alpha, \lambda) &= \int_0^\infty S_n(x) \ln \frac{S_n(x)}{S(x; \alpha, \lambda)} - [S_n(x) - S(x; \alpha, \lambda)] dx, \\ &= \sum_{i=1}^{n-1} \left(1 - \frac{i}{n} \right) \ln \left(1 - \frac{i}{n} \right) (x_{i+1:n} - x_{i:n}) \\ &\quad - \frac{1}{n} \sum_{i=1}^n \int_0^{x_i} \ln S(y; \alpha, \lambda) dy - (\bar{x} - \mu), \end{aligned} \quad (18)$$

where $S(x; \alpha, \lambda) = 1 - F(x; \alpha, \lambda)$ and

$$S_n(x) = \sum_{i=0}^{n-1} \left(1 - \frac{i}{n} \right) I_{[x_{i:n}, x_{i+1:n}]}(x), \quad (19)$$

is the empirical survival function, with $I_A(\cdot)$ as the indicator function.

The KLS estimates $\hat{\alpha}_{KLS}$ and $\hat{\lambda}_{KLS}$ of the parameters α and λ can be obtained by minimizing, with respect to α and λ , the function $K(\alpha, \lambda)$, see [14].

4. Monte Carlo simulation study

The following algorithm for generating a random sample of size n , X_1, X_2, \dots, X_n , from the GL distribution is based on the quantile function given by equation (3).

Algorithm

1. Generate $U_i \sim Uniform(0, 1)$, $i = 1, 2, \dots, n$.
2. Set

$$X_i = \frac{1}{\lambda} \ln \left(\frac{\alpha^2 + 2(1 - \alpha^2)(1 - U_i) + \alpha\sqrt{\alpha^2 + 4(1 + \alpha)(1 - U_i)}}{2(1 + \alpha)(1 - U_i)} \right),$$

where $i = 1, 2, \dots, n$.

The R function `rGL` in Appendix is used to generate random data from the GL distribution.

In this section, we evaluate the performance of the eight estimation methods presented in Section 3 through a Monte Carlo simulation study. The simulation experiment was repeated $M = 10,000$ times each with sample size $n = 50, 100, 150, 200, 250, 300$. The true values of the parameters are given by $(\alpha, \lambda) : (2, 1), (2, 4), (4, 1), (4, 6)$.

Two quantities, the bias and mean squared error (MSE), were examined in this Monte Carlo study, i.e.

$$\text{Bias}(\hat{\nu}) = \frac{1}{M} \sum_{j=1}^M (\hat{\nu}_j - \nu), \quad \nu = \alpha, \lambda,$$

$$\text{MSE}(\hat{\nu}) = \frac{1}{M} \sum_{j=1}^M (\hat{\nu}_j - \nu)^2, \quad \nu = \alpha, \lambda,$$

where $\hat{\nu}_j$ is the estimate of the parameter ν in j th iteration, using a particular estimation method, for each sample size n . All computations were performed using the R software, [11], version 4.0.5.

Considering the above quantities, the most efficient estimator will be the one whose bias is closer to zero with smaller MSE.

Figures 2-3 show the biases of the estimates of the parameters versus sample size n . From these figures, we observe that:

1. biases of the estimates of the parameters tend to zero as n increases, i.e. the estimators are asymptotically unbiased.
2. estimation methods 1: MMO, 3: MPS, and 6: PCE produce negative bias while the remaining 5 estimation methods produce positive bias.
3. estimation methods 3: MPS and 6: PCE produce smaller *absolute* bias than other estimation methods, in particular for smaller values of n .

Figures 4-5 show the MSE of the estimates of the parameters versus sample size n . From these figures, we observe that:

1. MSEs of the estimates of the parameters decrease as n increases
2. estimation methods 2: MLE and 7: L2D have the smallest MSE for all sample sizes..
3. estimation methods 6: PCE and 8: KLS have the largest MSE for all sample sizes.

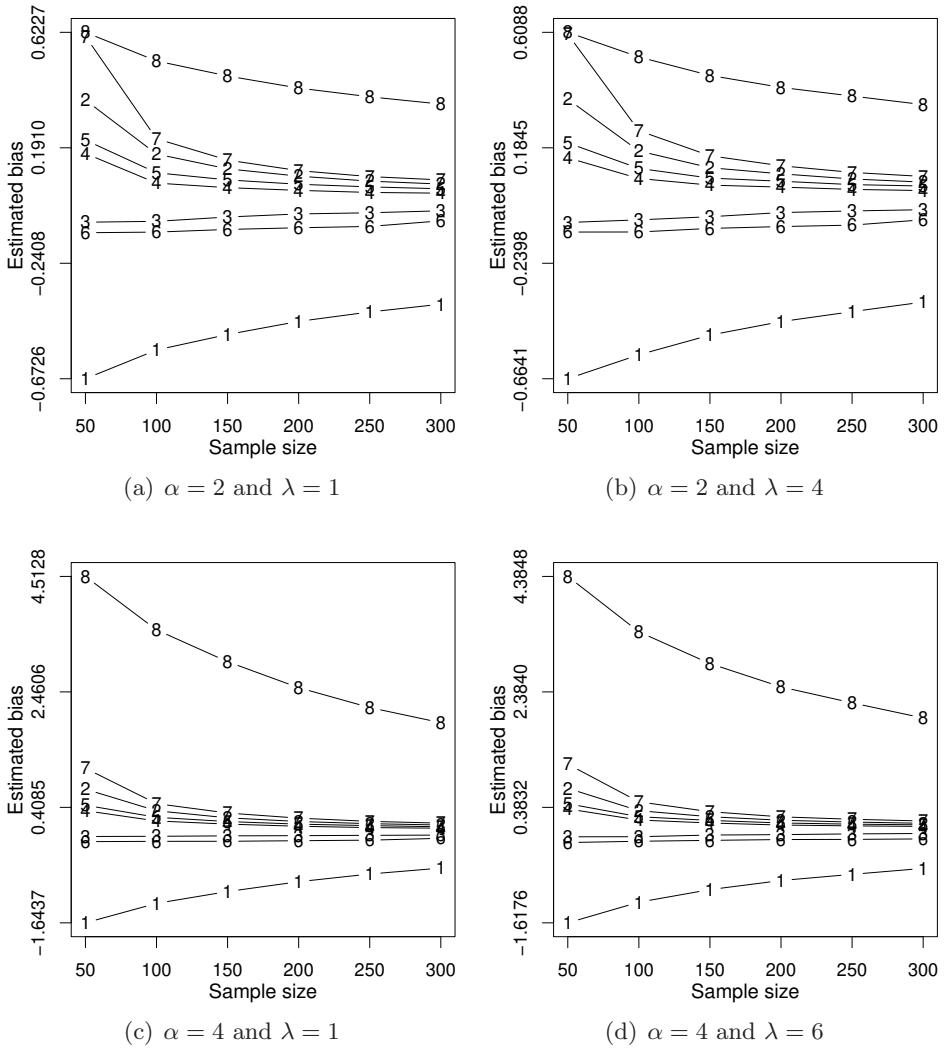


Figure 2: Bias of $\hat{\alpha}$ (1: MMO, 2: MLE, 3: MPS, 4: OLS, 5: WLS, 6: PCE, 7: L2D and 8: KLS).

5. Data analysis

In this section, we analyze a real data set representing 213 observations on the number of successive failures of the air conditioning system of a fleet of 13 Boeing 720 jet airplanes, see [10]. [4] fitted the Gompertz-Lindley distribution

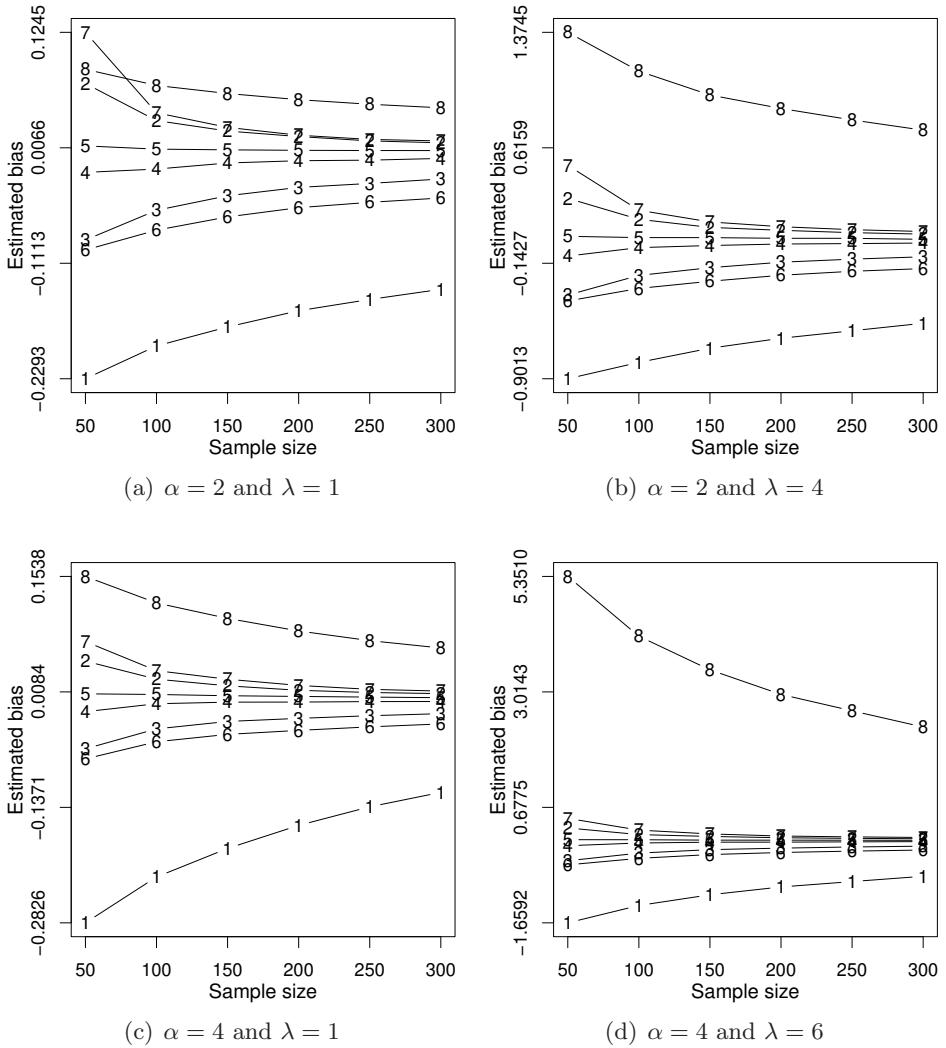


Figure 3: Bias of $\hat{\lambda}$ (1: MMO, 2: MLE, 3: MPS, 4: OLS, 5: WLS, 6: PCE, 7: L2D and 8: KLS).

to this data set using only the method of maximum likelihood.

Table 1 shows the different estimates of α and λ , using the eight methods presented in this paper, Anderson-Darling and Cramér-von Mises goodness-of-fit tests for the corresponding fitted GL distribution.

Table 1 shows that the WLS estimation method produces the smallest test

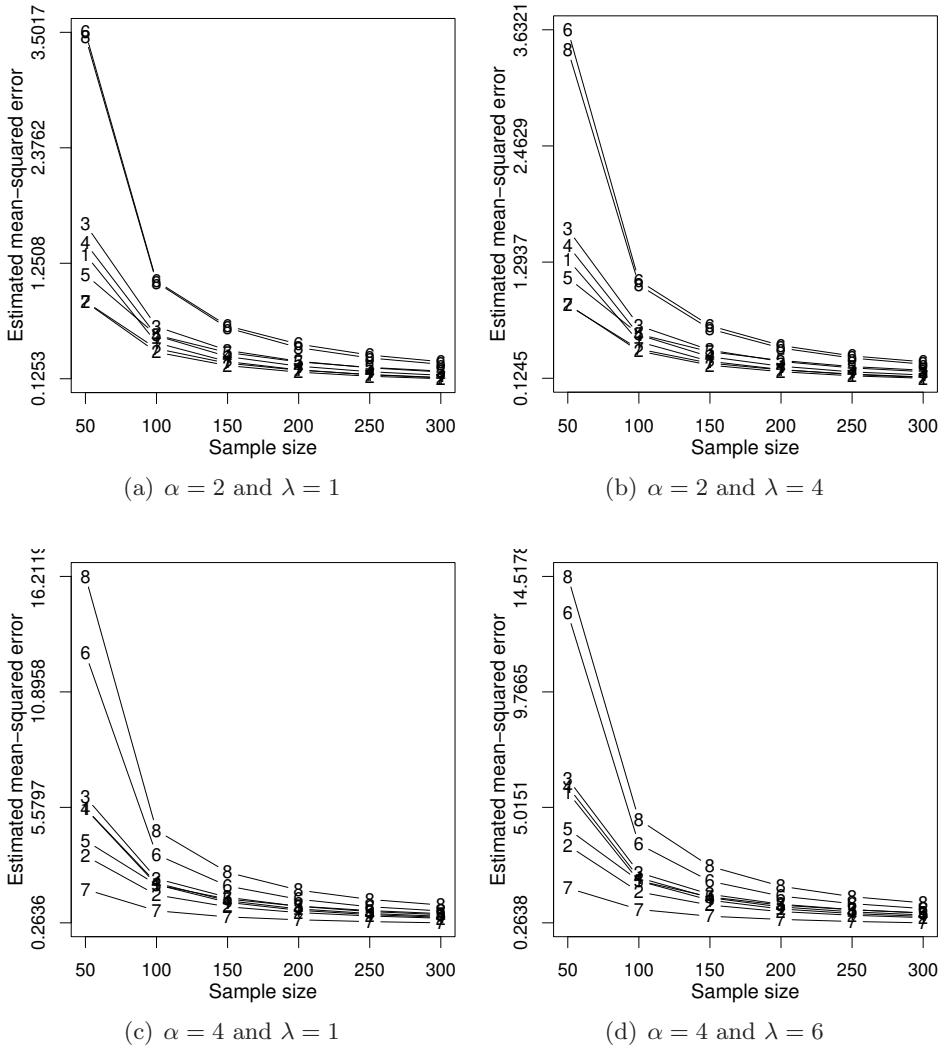


Figure 4: MSE of $\hat{\alpha}$ (1: MMO, 2: MLE, 3: MPS, 4: OLS, 5: WLS, 6: PCE, 7: L2D and 8: KLS).

statistics and largest p -values of both goodness-of-fit tests. Thus, for this data set, we can conclude that the WLS estimation method provides the best fit (Rank 1) for this data set. This conclusion is also supported by the Probability-Probability (PP) plots in Figure 6 and QQ-plots in Figure 7.

Finally, it is worth mentioning that by examining Table 1 and Figures 6-7,

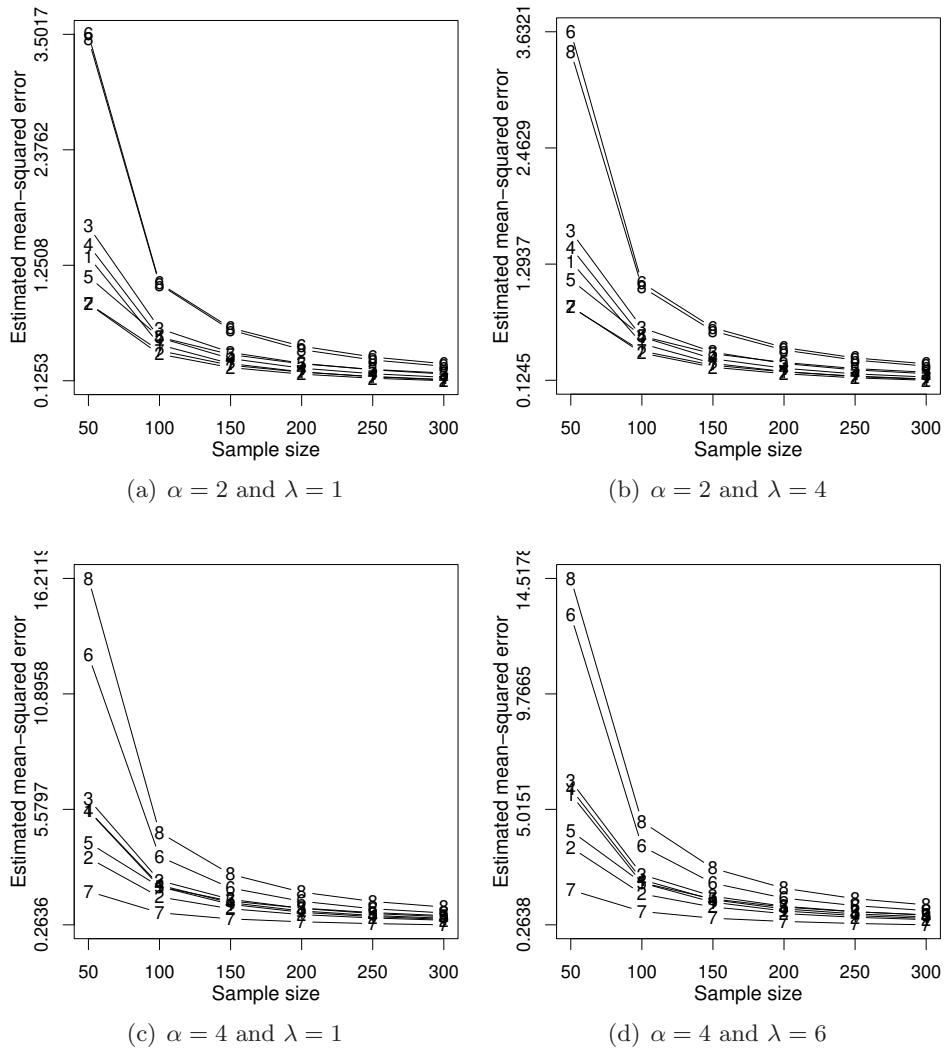


Figure 5: MSE of $\hat{\lambda}$ (1: MMO, 2: MLE, 3: MPS, 4: OLS, 5: WLS, 6: PCE, 7: L2D and 8: KLS).

the MLE estimation method comes second (Rank 2) to WLS method.

Table 1: Fitted distributions and goodness-of-fit tests for Proschan data set.

Method	$\hat{\alpha}$	$\hat{\lambda}$	AD	<i>p</i> -value	CvM	<i>p</i> -value
1. MMO	0.8009	0.0074	0.883	0.424	0.103	0.570
2. MLE	0.8122	0.0069	0.440	0.808	0.052	0.867
3. MPS	0.7343	0.0065	0.504	0.742	0.053	0.858
4. OLS	0.7075	0.0064	0.621	0.628	0.070	0.754
5. WLS	0.7580	0.0065	0.434	0.814	0.049	0.880
6. PCE	0.8080	0.0067	0.491	0.756	0.072	0.737
7. L2D	1.1117	0.0096	0.902	0.413	0.082	0.681
8. KLS	0.9758	0.0078	0.699	0.560	0.124	0.478

6. Concluding remarks

Different methods of estimators of the parameters are considered to perform a comparative study for the estimation of the parameters of Gompertz-Lindley distribution. It has been pointed out that the performance of a particular method depends on the model as well as on the criteria of comparison. It is hoped that our study will prove helpful to the data analysts.

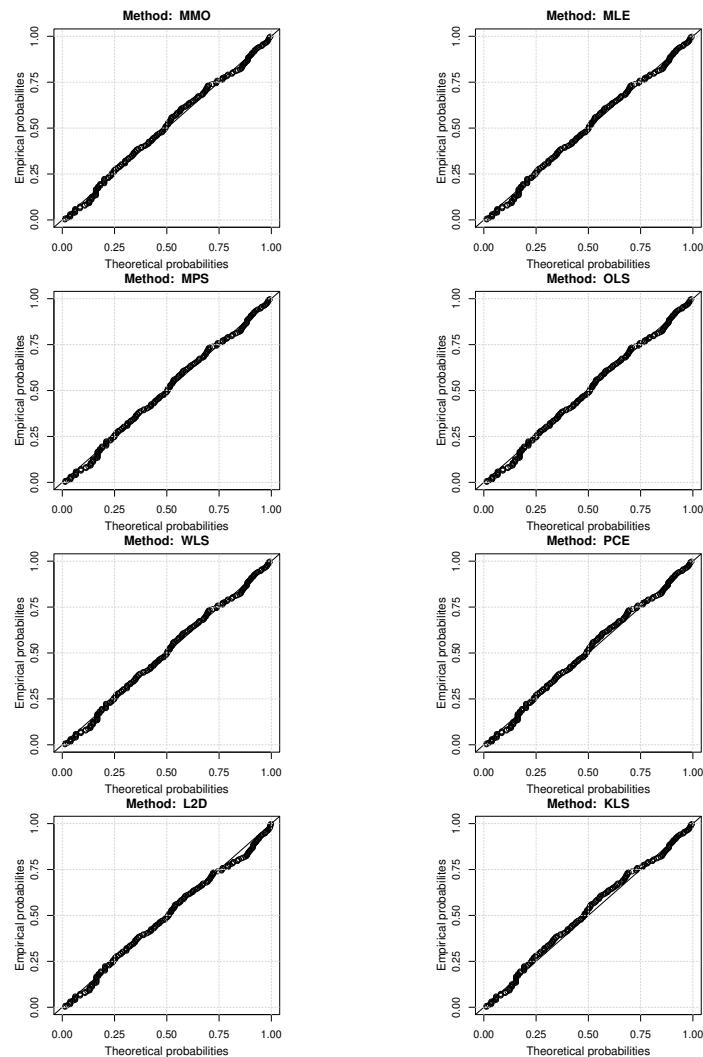


Figure 6: PP-plots under eight estimation methods for Proschan data set.

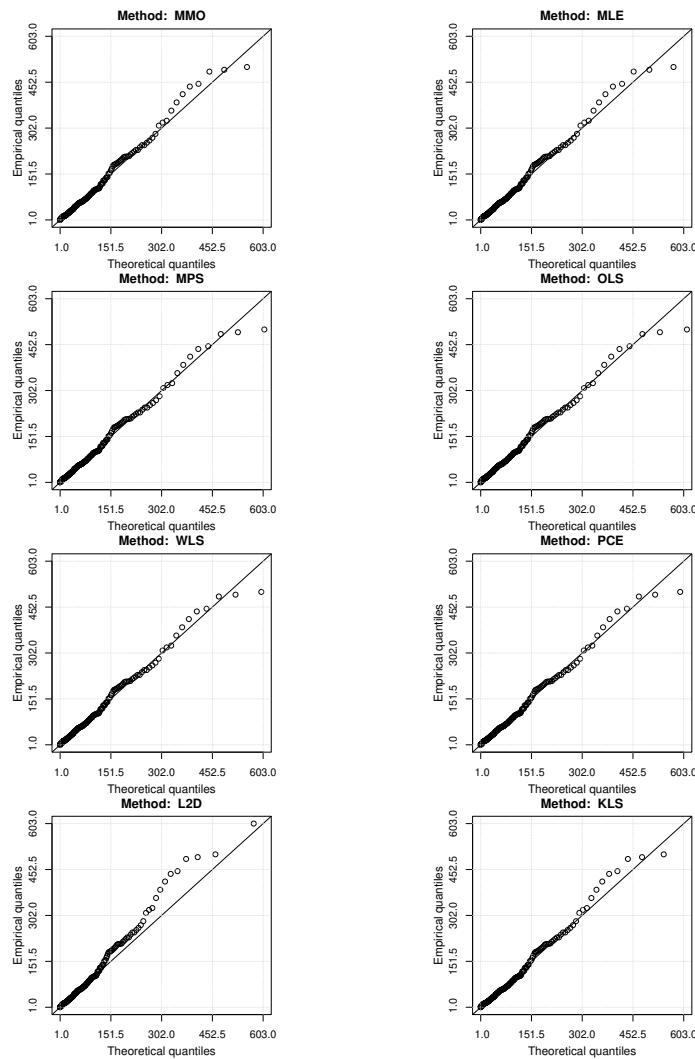


Figure 7: QQ-plots under eight estimation methods for Proschan data set.

Appendix

This appendix presents the R codes used to calculate the density function (`dGL`), cumulative distribution function (`pGL`), quantile function (`qGL`) and random number generation (`rGL`) for the Gompertz-Lindley distribution.

```
## GL pdf --- Equation 1
dGL <- function(x, a, l) {
  stopifnot(x > 0)
  stopifnot(a > 0)
  stopifnot(l > 0)
  (a^2*l)/(a+1) * exp(l*x)*(exp(l*x)+a+1)/(exp(l*x)+a-1)^3
}

## GL cdf --- Equation 2
pGL <- function(x, a, l) {
  stopifnot(x > 0)
  stopifnot(a > 0)
  stopifnot(l > 0)
  1 - a^2/(a+1) * (exp(l*x)+a)/(exp(l*x)+a-1)^2
}

## GL quantile function --- Equation 3
qGL <- function(u, a, l) {
  stopifnot(u > 0)
  stopifnot(u < 1)
  stopifnot(a > 0)
  stopifnot(l > 0)
  1/l * log(((a^2) + 2*(1-a^2)*(1-u)
  + a * sqrt(a^2+4*(1+a)*(1-u)))/(2*(1+a)*(1-u)))
}

## GL random deviates --- (see Section 4)
rGL <- function(n, a, l) {
  stopifnot(a > 0)
  stopifnot(l > 0)
  F.x = runif(n, 0, 1)
  qGL(F.x, a, l)
}
```

References

- [1] F. Alqallaf, M.E. Ghitany, M.F. Alfahad, A new two-parameter distribution with decreasing, increasing, bathtub hazard rate function and its applications, *International Journal of Applied Mathematics*, **33**, No 3 (2020), 523–548; DOI: 10.12732/ijam.v33i3.11
- [2] S. Dey, S. Ali, C. Park, Weighted exponential distribution: properties and different methods of estimation, *Journal of Statistical Computation and Simulation* (2015), 1–21.
- [3] S. Dey, T. Dey, D. Kundu, Two-parameter Rayleigh distribution: Different methods of estimation, *American Journal of Mathematical and Management Sciences*, **33** (2014), 55-74.
- [4] M.E. Ghitany, S.M. Aboukhamseen, A.A. Baqer, R.C. Gupta, Gompertz-Lindley distribution and associated inference, *Communications in Statistics - Simulation and Computation* (2019), 1-20; DOI: 10.1080/03610918.2019.1699113
- [5] M.E. Ghitany, D.K. Al-Mutairi, F.A. Al-Awadhi, M.M. Al-Burais, Marshall-Olkin extended Lindley distribution and its applications, *International Journal of Applied Mathematics*, **25**, No 5 (2012), 709-721.
- [6] M. E. Ghitany, Marshall-Olkin extended Pareto distribution and its application, *International Journal of Applied Mathematics*, **18** (1) (2005), 17-31.
- [7] R.D. Gupta, D. Kundu, Generalized exponential distribution: different method of estimations, *Journal of Statistical Computation and Simulation*, **69** (2001), 315-337.
- [8] J. Mazucheli, M. E. Ghitany, F. Louzada, Comparisons of ten estimation methods for the parameters of Marshall-Olkin extended exponential distribution, *Communications in Statistics - Simulation and Computation*, **46** (2017), 5627-5645.
- [9] J. Mazucheli, F. Louzada, M. E. Ghitany, Comparison of estimation methods for the parameters of the weighted Lindley distribution, *Applied Mathematics and Computation*, **220** (2013), 463-471.
- [10] F. Proschan, Theoretical explanation of observed decreasing failure rate. *Technometrics*, **5** (1963), 375-383.

- [11] R Core Team, *R: A Language and Environment for Statistical Computing*, Vienna (2021).
- [12] M. Teimouri, S.M. Hoseini, S. Nadarajah, Comparison of estimation methods for the Weibull distribution, *Statistics*, **47** (2013), 93-109.
- [13] J. Yang, D.W. Scott, Robust fitting of a Weibull model with optional censoring, *Computational Statistics and Data Analysis*, **67** (2013), 149-161.
- [14] G. Yari, A. Mirhabibi, A. Saghafi, Estimation of the Weibull parameters by Kullback-Leibler divergence of survival functions, *Applied Mathematics and Information Sciences*, **7** (2013), 187-192.