On Generalized Lindley Distribution and Its Applications to Model Lifetime Data from Biomedical Science and Engineering

Shanker R

Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea, Africa

Corresponding author:
Dr. Rama Shanker

✉ shankerrama2009@gmail.com

Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea, Africa.

Tel: 2918371120

Abstract

In the present paper, moments and moments based characteristics including expressions for coefficient of variation, skewness, kurtosis and index of dispersion of the three-parameter generalized Lindley distribution (GLD) introduced by Zakerzadeh and Dolati have been derived and discussed. The expressions for the hazard rate function and the mean residual life function of the distribution have been obtained. The applications and goodness of fit of the GLD have been discussed with several lifetime data sets from biomedical science and engineering and the fit has been compared with the three –parameter generalized gamma distribution (GGD).

Keywords: Generalized Lindley distribution; Moments; Coefficient of variation; Skewness; Kurtosis; Index of dispersion; Lifetime data; Generalized gamma distribution; Goodness of fit

Introduction

The probability density function of three-parameter generalized Lindley distribution (GLD) introduced by Zakerzadeh and Dolati (2009) having parameters $\alpha, \beta, \text{and } \theta$ is given by:

$$f(x;\alpha,\beta,\theta) = \frac{\theta^{\alpha+1}}{(\beta+\theta)\,\Gamma(\alpha+1)} x^{\alpha-1}(\alpha+\beta x)e^{-\theta x}; \; x>0, \alpha>0, \beta>0, \theta>0 \quad (1.1)$$

Where

$$\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} dy; \; \alpha>0$$

It is the complete gamma function.

It can be easily verified that the gamma distribution, the Lindley (1958) distribution and the exponential distribution are particular cases of (1.1) for $(\beta=0)$, $(\alpha=\beta=1)$ and $(\alpha=1, \beta=0)$, respectively. Ghitany et al. have detailed study about various properties of Lindley distribution, estimation of parameter and application for modeling waiting time data in a bank. Detailed and comparative study on modeling of lifetime data using one parameter Lindley and exponential distributions [1].

Further, the p.d.f. (1.1) can be easily expressed as a two-component mixture of gamma (α,θ) and gamma $(\alpha+1,\theta)$ distributions. We have,

$$f(x;\theta,\alpha,\beta) = p\, f_1(x;\theta,\alpha) + (1-p)\, f_2(x;\theta,\alpha+1), \quad (1.2)$$

where

$$p = \frac{\theta}{\beta+\theta},$$

$$f_1(x;\theta,\alpha) = \frac{\theta^\alpha}{\Gamma(\alpha)} e^{-\theta x} x^{\alpha-1},$$

$$f_2(x;\theta,\alpha+1) = \frac{\theta^{\alpha+1}}{\Gamma(\alpha+1)} e^{-\theta x} x^{\alpha+1-1}.$$

The corresponding distribution function of the GLD can be obtained as:

$$F(x;\alpha,\beta,\theta) = 1 - \frac{\alpha(\beta+\theta)\Gamma(\alpha,\theta x) + \beta(\theta x)^\alpha e^{-\theta x}}{(\beta+\theta)\Gamma(\alpha+1)}; \; x>0, \alpha>0, \beta>0, \theta>0 \quad (1.3)$$

Where $\Gamma(\alpha,z)$ is the upper incomplete gamma function defined as:

$$\Gamma(\alpha,z) = \int_z^\infty e^{-y} y^{\alpha-1} dy; \alpha>0, \, z\geq 0 \quad (1.4)$$

The probability density function of three-parameter generalized gamma distribution (GGD) introduced by Stacy [2] having parameters $\alpha, \beta, \text{and }$ is given by:

$$f_2(x;\alpha,\beta,\theta) = \frac{\beta\,\theta^\alpha}{\Gamma(\alpha)} x^{\beta\alpha-1} e^{-\theta x^\beta}; \; x>0, \alpha>0, \beta>0, \theta>0 \tag{1.5}$$

Where α and β are the shape parameter and θ is the scale parameter. Clearly the gamma distribution, the Weibull distribution and the exponential distribution are particular cases of (3.1) for $(\beta=1)$, $(\alpha=1)$ and $(\alpha=\beta=1)$ respectively. Detailed discussion about GGD is available in [2] and parametric estimation for the GGD is available in [3]. In fact, GGD is the power gamma distribution.

The cumulative distribution function of the GGD is thus obtained as:

$$F(x;\alpha,\beta,\theta) = 1 - \frac{\Gamma(\alpha,\theta x^\beta)}{\Gamma(\alpha)} \tag{1.6}$$

Where $\Gamma(\alpha,z)$ is the upper incomplete gamma function defined in (1.4).

Some of properties of GLD including nature of its p.d.f for varying values of its parameters, distribution of the sums of GLD, stochastic ordering, maximum likelihood estimates of parameters, distribution of the bivariate cases, and applications for modeling lifetime data [4]. It seems that some of its important properties based on moments including coefficient of variation, skewness, kurtosis, and index of dispersion has not been studied. Further, expressions for survival function, hazard rate function and mean residual life function of the distribution have not been obtained. Finally, the goodness of fit of GLD with two- parameter gamma, Weibull and lognormal distributions and concluded that the GLD is competing with these two-parameter distributions [4]. In fact, GLD should have been compared with other three-parameter lifetime distributions to test its goodness of fit. Recently Shanker and Shukla have detailed and critical study about the modeling of lifetime data from various fields of knowledge using three-parameter GLD and GGD and concluded that in majority of data sets GGD gives better fit [5].

In the present paper, the moments and moments based expressions including expressions for coefficient of variation, skewness, kurtosis, and index of dispersion have been given. The expressions for survival function, hazard rate function and mean residual life function have been obtained. The goodness of fit of GLD has been compared with the goodness of fit obtained by GGD and found that GLD does not give satisfactory fit in all data sets.

Moments and Associated Measures

Using mixture representation (1.2), the r th moment about origin of GLD (1.1) can be obtained as:

$$\mu_r' = E(X^r) = p\int_0^\infty x^r f_1(x;\theta,\alpha)\,dx + (1-p)\int_0^\infty x^r f_2(x;\theta,\alpha+1)\,dx$$

$$= \frac{[\alpha(\beta+\theta)+\beta]\Gamma(\alpha+r)}{\theta^r(\beta+\theta)\Gamma(\alpha+1)}; \; r=1,2,3,\ldots \tag{2.1}$$

Substituting $r=1,2,3,$ and 4 in (2.1), the first four moments about origin of GLD (1.1) are obtained as:

$$\mu_1' = \frac{\alpha(\beta+\theta)+\beta}{\theta(\beta+\theta)}$$

$$\mu_2' = \frac{(\alpha+1)\{\alpha(\beta+\theta)+2\beta\}}{\theta^2(\beta+\theta)}$$

$$\mu_3' = \frac{(\alpha+1)(\alpha+2)\{\alpha(\beta+\theta)+3\beta\}}{\theta^3(\beta+\theta)}$$

$$\mu_4' = \frac{(\alpha+1)(\alpha+2)(\alpha+3)\{\alpha(\beta+\theta)+4\beta\}}{\theta^4(\beta+\theta)}$$

Again, using relationship between central moments and moments about origin, the central moments of GLD are obtained as:

$$\mu_2 = \frac{(\alpha+1)\beta^2 + 2\theta(\alpha+1)\beta + \alpha\theta^2}{\theta^2(\beta+\theta)^2}$$

$$\mu_3 = \frac{2\left[(\alpha+1)\beta^3 + 3\theta(\alpha+1)\beta^2 + 3\theta^2(\alpha+1)\beta + \alpha\theta^3\right]}{\theta^3(\beta+\theta)^3}$$

$$\mu_4 = \frac{3\left[\begin{array}{l}(\alpha^2+4\alpha+3)\beta^4 + 4\theta(\alpha^2+4\alpha+3)\beta^3 + 2\theta^2(3\alpha^2+11\alpha+6)\beta^2 \\ +4\theta^3(\alpha^2+3\alpha+2)\beta + \alpha(\alpha+2)\theta^4\end{array}\right]}{\theta^4(\beta+\theta)^4}$$

The expressions for coefficient of variation (C.V.) coefficient of skewness $(\sqrt{\beta_1})$, coefficient of kurtosis (β_2), and index of dispersion (γ) of GLD are thus obtained as:

$$C.V. = \frac{\sigma}{\mu_1'} = \frac{\sqrt{(\alpha+1)\beta^2 + 2\theta(\alpha+1)\beta + \alpha\theta^2}}{\alpha(\beta+\theta)+\beta}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2\left[(\alpha+1)\beta^3 + 3\theta(\alpha+1)\beta^2 + 3\theta^2(\alpha+1)\beta + \alpha\theta^3\right]}{\left\{(\alpha+1)\beta^2 + 2\theta(\alpha+1)\beta + \alpha\theta^2\right\}^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\left[\begin{array}{l}(\alpha^2+4\alpha+3)\beta^4 + 4\theta(\alpha^2+4\alpha+3)\beta^3 + 2\theta^2(3\alpha^2+11\alpha+6)\beta^2 \\ +4\theta^3(\alpha^2+3\alpha+2)\beta + \alpha(\alpha+2)\theta^4\end{array}\right]}{\left\{(\alpha+1)\beta^2 + 2\theta(\alpha+1)\beta + \alpha\theta^2\right\}^2}$$

$$\gamma = \frac{\sigma^2}{\mu_1'} = \frac{(\alpha+1)\beta^2 + 2\theta(\alpha+1)\beta + \alpha\theta^2}{\theta(\beta+\theta)\{\alpha(\beta+\theta)+\beta\}}$$

Hazard Rate Function and Mean Residual Life Function

Hazard rate function

Using the mixture representation (1.2), the survival (reliability) function of GLD can be obtained as:

$$S(x) = P(X > x) = p \int_x^\infty f_1(y;\theta,\alpha)dy + (1-p)\int_x^\infty f_2(y;\theta,\alpha+1)dy$$

$$= \frac{\alpha(\beta+\theta)\Gamma(\alpha,\theta x) + \beta(\theta x)^\alpha e^{-\theta x}}{(\beta+\theta)\Gamma(\alpha+1)} \qquad (3.1.1)$$

Where $\Gamma(\alpha,z)$ is the upper incomplete gamma function defined in (1.4).

The hazard (or failure) rate function, $h(x)$ of GLD is thus obtained as:

$$h(x) = \frac{f(x)}{S(x)} = \frac{\theta^{\alpha+1}x^{\alpha-1}(\alpha+\beta x)e^{-\theta x}}{\alpha(\beta+\theta)\Gamma(\alpha,\theta x)+\beta(\theta x)^\alpha e^{-\theta x}} ; x>0, \theta>0, \alpha>0, \beta>0 \qquad (3.1.2)$$

The shape of the hazard rate function, $h(x)$ of the GLD is difficult to study because it includes the upper incomplete gamma function.

Mean residual life function

Using the mixture representation (1.2), the mean residual life function $\mu(x) = E(X-x \mid X>x)$ of the GLD can be obtained as:

$$\mu(x) = \frac{1}{S(x)}\int_x^\infty y f(y)dy - x$$

$$= \frac{1}{S(x)}\left[p\int_x^\infty f_1(y;\theta,\alpha)dy + (1-p)\int_x^\infty f_2(y;\theta,\alpha+1)dy \right] - x$$

$$= \frac{\{\alpha(\beta+\theta)+1\}(\theta x)^\alpha e^{-\theta x} + [\alpha\{\alpha(\beta+\theta)+\beta\} - \alpha(\beta+\theta)(\theta x)]\Gamma(\alpha,\theta x)}{\theta[\alpha(\beta+\theta)\Gamma(\alpha,\theta x)+\beta(\theta x)^\alpha e^{-\theta x}]}$$

Where $\Gamma(\alpha,z)$ is the upper incomplete gamma function defined in (1.4).

The shape of the mean residual life function, $\mu(x)$ of the GLD is difficult to study because it includes the upper incomplete gamma function.

Maximum Likelihood Estimation

Maximum likelihood estimate (MLE) of GLD

Let $(x_1, x_2, x_3, \dots, x_n)$ be a random sample of size n from GLD (1.1). The likelihood function, L of GLD is given by

$$L = \left(\frac{\theta^{\alpha+1}}{\beta+\theta}\right)^n \frac{1}{(\Gamma(\alpha+1))^n}\prod_{i=1}^n x_i^{\alpha-1}(\alpha+\beta x_i)e^{-n\theta\bar{x}} ; \ \bar{x}$$ being the sample mean

The natural log likelihood function is thus obtained as:

$$\ln L = n\left[(\alpha+1)\ln\theta - \ln(\beta+\theta) - \ln(\Gamma(\alpha+1))\right] + (\alpha-1)\sum_{i=1}^n \ln(x_i) + \sum_{i=1}^n \ln(\alpha+\beta x_i) - n\theta\bar{x}$$

The MLE $\hat{\theta}, \hat{\alpha}, \hat{\beta}$ of parameters θ, α, β of GLD can be obtained by solving the natural log likelihood equation using R software (Package Stat 4).

Maximum likelihood estimate (MlE) of GGD

Assuming $(x_1, x_2, x_3, \dots, x_n)$ a random sample of size n from GGD (1.5), the likelihood function, L of GGD can be given as

$$L = \left(\frac{\beta\theta^\alpha}{\Gamma(\alpha)}\right)^n \prod_{i=1}^n x_i^{\beta\alpha-1} e^{-\theta\sum_{i=1}^n x_i^\beta} ;$$

$\bar{x}$ being the sample mean.

The natural log likelihood function is thus obtained as:

$$\ln L = n\left[\ln\beta + \alpha\ln\theta - \ln(\Gamma(\alpha))\right] + (\beta\alpha-1)\sum_{i=1}^n \ln(x_i) - \theta\sum_{i=1}^n x_i^\beta$$

The MLE $\hat{\theta}, \hat{\alpha}, \hat{\beta}$ of parameters θ, α, β of GGD can be obtained by solving the natural log likelihood equation using R software (Package Stat 4).

Applications and Goodness of Fit

In this section, the applications and goodness of fit of the GLD have been discussed for several lifetime data from biomedical science and engineering and the fit is compared GGD. The following eight lifetime data sets have been considered for testing the goodness of fit of GLD and GGD.

Data set 1

This data set used by Bhaumik et al. is vinyl chloride data obtained from clean up gradient monitoring wells in mg/l [6]:

5.1	1.2	1.3	0.6	0.5	2.4	0.5	1.1	8	0.8	0.4	0.6
0.9	0.4	2	0.5	5.3	3.2	2.7	2.9	2.5	2.3	1	0.2
0.1	0.1	1.8	0.9	2	4	6.8	1.2	0.4	0.2		

Data set 2

This data set represents the waiting times (in minutes) before service of 100 Bank customers and examined and analyzed by Ghitany et al. [7] for fitting the Lindley distribution [8].

0.8	0.8	1.3	1.5	1.8	1.9	1.9	2.1	2.6	2.7	2.9	3.1
3.2	3.3	3.5	3.6	4.0	4.1	4.2	4.2	4.3	4.3	4.4	4.4
4.6	4.7	4.7	4.8	4.9	4.9	5.0	5.3	5.5	5.7	5.7	6.1
6.2	6.2	6.2	6.3	6.7	6.9	7.1	7.1	7.1	7.1	7.4	7.6
7.7	8.0	8.2	8.6	8.6	8.6	8.8	8.8	8.9	8.9	9.5	9.6
9.7	9.8	10.7	10.9	11.0	11.0	11.1	11.2	11.2	11.5	11.9	12.4
12.5	12.9	13.0	13.1	13.3	13.6	13.7	13.9	14.1	15.4	15.4	17.3
17.3	18.1	18.2	18.4	18.9	19.0	19.9	20.6	21.3	21.4	21.9	23.0
27.0	31.6	33.1	38.5								

Data Set 3

This data is for the times between successive failures of air conditioning equipment in a Boeing 720 airplane [9].

74	57	48	29	502	12	70	21	29	386	59	27
153	26	326									

Data set 4

This data set represents the lifetime's data relating to relief times (in minutes) of 20 patients receiving an analgesic [10].

1.1	1.4	1.3	1.7	1.9	1.8	1.6	2.2	1.7	2.7	4.1	1.8
1.5	1.2	1.4	3	1.7	2.3	1.6	2				

Data set 5

This data set is the strength data of glass of the aircraft window reported by [11].

18.83 20.8 21.657 23.03 23.23 24.05 24.321 25.5 25.52 25.8 26.69 26.77 26.78
27.05 27.67 29.9 31.11 33.2 33.73 33.76 33.89 34.76 35.75 35.91 36.98
37.08 37.09 39.58 44.045 45.29 45.381

Data set 6

The following data represent the tensile strength, measured in GPa, of 69 carbon fibers tested under tension at gauge lengths of 20 mm [12].

1.312 1.314 1.479 1.552 1.700 1.803 1.861 1.865 1.944 1.958 1.966 1.997 2.006
2.021 2.027 2.055 2.063 2.098 2.140 2.179 2.224 2.240 2.253 2.270 2.272 2.274
2.301 2.301 2.359 2.382 2.382 2.426 2.434 2.435 2.478 2.490 2.511 2.514 2.535
2.554 2.566 2.570 2.586 2.629 2.633 2.642 2.648 2.684 2.697 2.726 2.770 2.773
2.800 2.809 2.818 2.821 2.848 2.880 2.954 3.012 3.067 3.084 3.090 3.096 3.128
3.233 3.433 3.585 3.858

Data set 7

The following data set represents the failure times (in minutes) for a sample   of 15 electronic components in an accelerated life test [13].

1.4 5.1 6.3 10.8 12.1 18.5 19.7 22.2 23.0 30.6 37.3 46.3
53.9 59.8 66.2

Data set 8

The following data set represents the number of cycles to failure for 25 100-cm specimens of yarn, tested at a particular strain level [13,14].

15 20 38 42 61 76 86 98 121 146 149 157
175 176 180 180 198 220 224 251 264 282 321 325 653

In order to compare the goodness of fit of GLD and GGD, values of $-2\ln L$, K-S Statistics (Kolmogorov-Smirnov Statistics) and p-values for the above data sets have been computed and presented in **Table 1.** The formulae for computing K-S Statistics are as follows:

$$K-S = \underset{x}{\text{Sup}} \left| F_n(x) - F_0(x) \right|,$$

Where k = the number of parameters, n = the sample size and $F_n(x)$ is the empirical distribution function. The best distribution corresponds to lower values of $-2\ln L$ and K-S statistics and higher p- value.

It is obvious from the fitting of GLD and GGD that both are competing. But in data sets 5 and 8, GLD gives better fit than GGD and in all other data sets GGD gives better fit than GLD. It should be noted that [4] have considered data sets 7 and 8 for testing the goodness of fit of GLD and compared it with two - parameter gamma, Weibull and Lognormal distributions and concluded that GLD gives slightly better fit than these distributions.

Conclusion

In this paper moments and moments based characteristics including expressions for coefficient of variation, skewness, kurtosis and index of dispersion of the three-parameter generalized Lindley distribution (GLD) introduced by [4] have been derived and discussed. The expressions for the hazard rate function and the mean residual life function of the distribution have been obtained. The applications and goodness of fit of the distribution have been discussed with several lifetime data sets and the fit has been compared with the three –parameter generalized gamma distribution (GGD). The goodness of fit of the GLD and GGD shows that GGD gives better fit in majority of lifetime data sets and hence GGD can be considered as an important model over GLD for modeling lifetime data from biomedical science and engineering.

Acknowledgement

Table 1 ML Estimates, -2ln L, K-S Statistics and p-values of the fitted distributions of data sets 1 to 8.

	Model	ML Estimates			$-2\ln L$	K-S Statistic	*P*-value
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$			
Data 1	GLD	1.0628	0.0006	0.5647	110.826	0.936	0.000
	GGD	5.9538	0.3802	5.2747	109.721	0.927	0.000
Data 2	GLD	2.0093	0.0007	0.2038	634.600	0.043	0.994
	GGD	3.8037	0.7017	0.8028	634.035	0.036	0.999
Data 3	GLD	0.9427	0.0003	0.0081	173.873	0.726	0.000
	GGD	26.6637	0.1736	12.7036	170.488	0.726	0.000
Data 4	GLD	9.6686	0.0029	5.0891	35.637	0.609	0.000
	GGD	51.4619	0.4350	39.4639	34.376	0.600	0.000
Data 5	GLD	17.9881	14.6111	0.6150	208.233	0.135	0.580
	GGD	19.6720	0.9814	0.6800	208.225	0.136	0.562
Data 6	GLD	22.7198	4.7710	9.3907	101.959	0.056	0.979
	GGD	3.5861	2.6483	0.3044	100.581	0.044	0.999
Data 7	GLD	1.2025	0.0832	0.0641	128.161	0.095	0.997
	GGD	0.8597	1.4152	0.0068	127.931	0.095	0.997
Data 8	GLD	0.8186	3.9740	0.0101	304.883	0.132	0.769
	GGD	1.9916	0.9426	0.0152	304.928	0.139	0.719

References

1 Shanker R, Hagos F, Sujatha S (2015) On modeling of lifetimes data using exponential and Lindley distributions. Biometrics & Biostatistics International Journal 2(5): 1-9.

2 Stacy EW (1962) A generalization of the gamma distribution, Annals of Mathematical Statistics 33: 1187–1192.

3 Stacy EW, Mihram GA (1965) Parametric estimation for a generalized gamma distribution. Technometrics 7: 349-358.

4 Zakerzadeh H, Dolati A (2009) Generalized Lindley distribution, Journal of Mathematical extension 3(2): 13–25.

5 Shanker R, Shukla KK (2016) On modeling of lifetime data using three parameters generalized Lindley and generalized gamma distributions. Biometrics & Biostatistics International Journal.

6 Bhaumik DK, Kapur K, Gibbons RD (2009) Testing parameters of a gamma distribution for small samples. Technometrics 51: 326-334.

7 Ghitany ME, Atieh B, Nadarajah S (2008) Lindley distribution and its Application. Mathematics Computing and Simulation 78: 493-506.

8 Lindley DV (1958) Fiducial distributions and Bayes' theorem. Journal of the Royal Statistical Society. Series B 20: 102–107.

9 Proschan F (1963) Theoretical explanation of observed decreasing failure rate. Technometrics 5: 375-383.

10 Gross AJ, Clark VA (1975) Survival distributions: Reliability applications in the biometrical sciences. John Wiley, New York, USA.

11 Fuller EJ, Frieman S, Quinn J, Quinn G, Carter W (1994) Fracture mechanics approach to the design of glass aircraft windows: A case study. SPIE Proc 2286: 419-430.

12 Bader MG, Priest AM (1982) Statistical aspects of fiber and bundle strength in hybrid composites. In: Progressing Science in Engineering Composites. Tokyo, Japan.

13 Lawless JF (2003) Statistical models and methods for lifetime data. John Wiley and Sons, New York, USA.

14 Smith RL, Naylor JC (1987) A comparison of maximum likelihood and Bayesian estimators for the three parameter Weibull distribution. Applied Statistics 36: 358-369.