ICASTOR Journal of Mathematical Sciences
Vol. 3, No. 2 (2009) 227 – 235

SOME PROPERTIES AND APPLICATION OF THE NEGATIVELY SKEWED EXTENDED GENERALIZED LOGISTIC DISTRIBUTION

A. K. Olapade

Department of Statistics and Actuarial Science
University of Stellenbosch
Stellenbosch, South Africa

ABSTRACT

In this paper, some results on properties of negatively skewed extended generalized logistic distribution of Olapade (2005) is presented. The survival and the hazard function, the mean, the median, the mode and the 100k-percentage point of the distribution are presented. The study of order statistics is also considered in addition to the estimation of the parameters of the distribution and an application of the distribution to model a chemical data.

KEYWORDS: Negatively skewed extended generalized logistic distribution, mean, median, mode, 100k-percentage point, order statistics, estimation and application.

2000 Mathematics Subject Classifcations: Primary 62E15; Secondary 62E10.

INTRODUCTION

The logistic model has been useful in the analyses of biological assay and quantal response data. George and Ojo (1980), Ojo (1997), Ojo (2002), Olapade (2002) are few of many literatures on logistic model. Olapade (2005) presented a form of logistic distribution with probability density function

$$f_X(x;\lambda,p)=\frac{\lambda p e^{-px}}{(\lambda+e^{-x})^{p+1}},\quad -\infty<x<\infty,\ \lambda>0,\ p>0, \tag{1}$$

for a continuous random variable X, with corresponding cumulative distribution function

$$F_X(x;\lambda,p)=1-\frac{e^{-px}}{(\lambda+e^{-x})^{p}},\quad -\infty<x<\infty,\lambda>0,p>0. \tag{2}$$

Correspondence: Department of Mathematics, Obafemi Awolowo University, Ile-Ife, Nigeria.
E-mail: akolapad@oauife.edu.ng; olapade@sun.ac.za

We refer to the function in the equation (1) as the negatively skewed extended generalized logistic (NSEGL) density function. The case $\lambda = 1$ in the equation (2) corresponds to the type II generalized logistic distribution of Balakrishnan and Leung (1988) and Olapade (2002). As expected, the case $\lambda = p = 1$ corresponds to the standard logistic distribution density function.

If X is the lifetime of an object, then the survival function $S(x; \lambda, p)$ of X is

$$S(x; \lambda, p) = pr(X > x) = e^{-px}(\lambda + e^{x})^{-p} \tag{3}$$

This is the probability that an object whose life time is a random variable X that follows the NSEGL distribution will survive beyond time x.

The hazard function of the random variable X is $H(x) = f(x)/(1 - F(x))$. Hence, if X follows the NSEGL distribution, its hazard function is

$$H(x; \lambda, p) = \lambda p(\lambda + e^{x})^{-1}. \tag{4}$$

1. THE MOMENTS OF THE NSEGL DISTRIBUTION

The moment generating function of the NSEGL distribution as presented in Olapade (2005) is

$$M_X(t) = \frac{\lambda^{-t}\Gamma(p-t)\Gamma(1+t)}{\Gamma(p)} \tag{1.1}$$

From equation (1.1), the cumulant generating function of the distribution is obtained as

$$\ln \phi_X(t) = t \ln \lambda + \ln \Gamma(p - t) + \ln \Gamma(1 + t) - \ln \Gamma(p) \tag{1.2}$$

The r^{th} cumulant is obtained as

$$\kappa_r(X) = \frac{d^r}{dt^r}[-t \ln \lambda]_{t=0} + \frac{d^r}{dt^r}[\ln \Gamma(p - t) + \ln \Gamma(1 + t)]_{t=0}. \tag{1.3}$$

The 2nd or the 3rd term of the right hand side of equation (1.3) is called di-gamma function. The series expansion of this function as in Copson (1962) is given as

$$\psi^{r-1}(x) = (r-1)!\left[(-1)^r \sum_{j=0}^{\infty}(j+x)^{-r}\right]_{r\geq 2}$$

and

$$\psi(x) = \sum_{j=0}^{\infty}(j+x)^{-1}, \quad r = 1$$

where

$$\psi^{r-1}(z) = \frac{d^r}{dz^r}[\ln(z+1)]. \tag{1.4}$$

Therefore

$$\kappa_r(X) = (r-1)!\,(-1)^r \left[\sum_{j=0}^{\infty}(j+x)^{-r} + (-1)^r \sum_{j=0}^{\infty}(j+p)^{-r} \right]_{r\geq 2} \tag{1.5}$$

and

$$\kappa_r(X) = -\ln\lambda - \sum_{j=0}^{\infty}(j+1)^{-1} + \sum_{j=0}^{\infty}(j+p)^{-1}, \quad r=1 \tag{1.6}$$

When $r = 1$, we obtained the mean μ_1 of the NSEGL distribution as

$$\mu_1 = \kappa_1(X) = -\ln\lambda - \sum_{j=0}^{\infty}(j+1)^{-1} + \sum_{j=0}^{\infty}(j+p)^{-1} = -\ln\lambda - \sum_{j=1}^{p-1} j^{-1} \tag{1.7}$$

The second cumulant of the distribution is obtained when $r = 2$ as

$$\kappa_2(X) = \sum_{j=0}^{\infty}(j+p)^{-2} + \sum_{j=0}^{\infty}(j+1)^{-2} = \sum_{j=p}^{\infty} j^{-2} + \frac{\pi^2}{6} \tag{1.8}$$

Therefore, the variance of the random variable X can be obtained from $\kappa_2(X) - \{\kappa_1(X)\}^2$ as

$$\mu_2 = \sigma_X^2 = \sum_{j=p}^{\infty} j^{-2} + \frac{\pi^2}{6} - \left(\sum_{j=1}^{p-1} j^{-1} + \ln\lambda \right)^2 \tag{1.9}$$

The third cumulant of the distribution when $r = 3$ in equation (1.5) is

$$\kappa_3(X) = -2\left\{ \sum_{j=0}^{\infty}(j+1)^{-3} - \sum_{j=0}^{\infty}(j+p)^{-3} \right\} = -2\sum_{j=1}^{p-1} j^{-3} \tag{1.10}$$

The fourth cumulant of the random variable X that has an NSEGL distribution is obtained when $r = 4$ as

$$\kappa_4(X) = 6\left\{ \sum_{j=0}^{\infty}(j+1)^{-4} + \sum_{j=0}^{\infty}(j+p)^{-4} \right\} = 12\sum_{j=p}^{\infty} j^{-4} + 6\sum_{j=1}^{p-1} j^{-4} \tag{1.11}$$

So, the third moment μ_3 of the distribution could be obtained from

$$\mu_3 = \kappa_3(X) - 3\kappa_2(X)\kappa_1(X) + 2\{\kappa_1(X)\}^3 \quad (1.12)$$

while the fourth moment could be obtained from

$$\mu_4 = \kappa_4(X) - 4\kappa_3(X)\kappa_1(X) + 6\kappa_2\{\kappa_1(X)\}^2 - 3\{\kappa_1(X)\}^4 \quad (1.13)$$

Hence, the coefficient of skewness and that of kurtosis for X with NSEGL distribution are $\beta_1(X) = \mu_3^2/\mu_2^3$ and $\beta_2(X) = \mu_4/\mu_2^2$ respectively.

2. THE MEDIAN OF THE NSEGL DISTRIBUTION

The median of a probability density function $f(x)$ is a point x_m on the real line such that

$$\int_{-\infty}^{x_m} f(x)\,dx = 1/2 \quad (2.1)$$

This implies that $F(x_m) = 1/2$. For the NSEGL distribution with probability distribution function given in equation (2), equating the distribution function to $1/2$ and replacing x by x_m implies

$$1 - \frac{e^{-px_m}}{(\lambda + e^{-x_m})^p} = 1/2 \Rightarrow \frac{e^{-px_m}}{(\lambda + e^{-x_m})^p} = 1/2$$

By simplification

$$\Rightarrow x_m = \ln[\frac{\sqrt[p]{2}-1}{\lambda}] \quad (2.2)$$

This reduces to zero when $\lambda = 1$ and $p = 1$ as the median of the standard logistic distribution.

3. THE $100k$-PERCENTAGE POINT OF THE NSEGL DISTRIBUTION

Consider the NSEGL distribution, the $100k$-percentage point is obtained by equating the probability distribution function to k. That is

$$F(x_{(k)}) = k \Rightarrow 1 - \frac{e^{-px_{(k)}}}{\left(\lambda + e^{-x_{(k)}}\right)^p} = k \quad (3.1)$$

Solving for $x_{(k)}$ gives

$$x_{(k)} = \ln\left[\frac{1-\sqrt[p]{1-k}}{\lambda\sqrt[p]{1-k}}\right] . \quad (3.2)$$

This gives the value of the point $x_{(k)}$on the real'line that produces a percentage k of the distribution. We can easily test this by checking the value of $x_{(k)}$when $k = 0.5$ which corresponds to the median when $p = \lambda = 1$. This gives the value of the median for a standard logistic distribution which is already known to be $\ln 1 = 0$.

4. THE MODE OF THE NSEGL DISTRIBUTION

The mode of a probability density function is obtained by equating the derivative of the density function to zero and solve for the variable. Therefore, for the NSEGL distribution density function in equation (1)

$$f'_X(x;\lambda,p) = p\lambda\left[\frac{-pe^{-px}(\lambda+e^{-x})+(p+1)e^{-px}e^{-x}}{(\lambda+e^{-x})^{p+2}}\right] \tag{4.1}$$

By equating the derivative to zero and solve for x we have

$$x = \ln \lambda p \tag{4.2}$$

This gives the mode of the distribution as confirmed when we put $p = \lambda = 1$ which gives the mode of standard logistic distribution.

5. ORDER STATISTICS FROM THE NSEGL DISTRIBUTION

Let $X_1, X_2, \ldots, X_n$ be n independently continuous random variables from the NSEGL distribution and let $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ be the corresponding order statistics. Let $F_{X_{r:n}}(x); r = 1, 2, \ldots, n$ be the cumulative distribution function of the r^{th} order statistics $X_{r:n}$ and $f_{X_{r:n}}(x)$ denotes its probability density function. David (1970) gives the probability density function of $X_{r:n}$ as

$$f_{X_{r:n}}(x) = \frac{1}{B(r,n-r+1)} P^{r-1}(x)[1-P(x)]^{n-r}p(x) \tag{5.1}$$

For the NSEGL distribution with probability density function and cumulative distribution function given in equations (1) and (2) respectively,

$$f_{X_{r:n}}(x) = \frac{1}{B(r,n-r+1)}\left[1-\frac{e^{-px}}{(\lambda+e^{-x})^p}\right]^{r-1}\left[1-1+\frac{e^{-px}}{(\lambda+e^{-x})^p}\right]^{n-r}\frac{p\lambda e^{-px}}{(\lambda+e^{-x})^{p+1}}$$

$$= \frac{p\lambda}{B(r,n-r+1)}\frac{e^{-px}}{(\lambda+e^{-x})^{p=1}}\left[1-\frac{e^{-px}}{(\lambda+e^{-x})^p}\right]^{r-1}\left[\frac{e^{-px}}{(\lambda+e^{-x})^p}\right]^{n-r} \tag{5.2}$$

5.1. THE MINIMUM ORDER STATISTICS

Consider the probability density of the r^{th} order statistics from the NSEGL distribution in the equation above. Let $r = 1$, then the probability density function of the minimum order statistics is

$$f_{X_{1:n}}(x) = \frac{np\lambda e^{-npx}}{(\lambda + e^{-x})^{pn+1}} \tag{5.3}$$

This is another NSEGL distribution with parameters (np, λ). This distribution shares all the properties of the NSEGL distribution with np replacing p.

5.2. THE MAXIMUM ORDER STATISTICS FROM THE NSEGL DISTRIBUTION

Consider the probability density of the r^{th} order statistics from the NSEGL distribution in equation (5.2). Let $r = n$, then the probability density function of the maximum order statistics is obtained as

$$\begin{aligned}
f_{X_{n:n}}(x) &= \frac{np\lambda e^{-px}}{(\lambda + e^{-x})^{p+1}}[1 - \frac{e^{-px}}{(\lambda + e^{-x})^{p}}]^{n-1} \\
&= \frac{np\lambda e^{-px}}{(\lambda + e^{-x})^{np+1}}[(\lambda + e^{-x})^{p} - e^{-px}]^{n-1} \\
&= \frac{np\lambda e^{-px}}{(\lambda + e^{-x})^{np+1}} \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} (\lambda + e^{-x})^{pk} (e^{-px})^{n-k-1} \\
&= np\lambda \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} (\lambda + e^{-x})^{pk-np-1} (e^{-px})^{n-k} \\
&= np\lambda \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} (e^{-x})^{np-kp} \left[\sum_{j=0}^{pk-pn-1} \binom{pk-pn-1}{j} \lambda^{pk-pn-j-1} e^{-jx} \right] \\
&= np \sum_{k=0}^{n-1} \sum_{j=0}^{pk-pn-1} (-1)^k \binom{n-1}{k} \binom{pk-pn-1}{j} \lambda^{pk-np-j} e^{-(np-pk+j)x}
\end{aligned} \tag{5.4}$$

The q^{th} moment of $X_{n:n}$ is obtained as

$$E[X_{n:n}^q] = np \sum_{k=0}^{n-1} \sum_{j=0}^{pk-pn-1} (-1)^k \binom{n-1}{k} \binom{pk-pn-1}{j} \lambda^{pk-np-j} \int_{-\infty}^{\infty} x^q \, e^{-(np-pk+j)x} dx \tag{5.5}$$

The equation (5.5) above can be used to obtain the moments of the maximum observations from the NSEGL distribution when n, p and λ are known. This can be done numerically for $q = 1, 2, \ldots$.

6. ESTIMATION OF THE PARAMETERS OF THE NSEGL DISTRIBUTION

Given a sample $X_1, X_2, \ldots, X_n$ of size n from an NSEGL distribution with probability density function

$$f_X(x;\mu,\sigma,\lambda,p) = \frac{p\lambda}{\sigma}\frac{e^{-p\left(\frac{x_i-\mu}{\sigma}\right)}}{\left[\lambda + e^{-\left(\frac{x_i-\mu}{\sigma}\right)}\right]^{(p+1)}} \tag{6.1}$$

where μ is the location parameter, σ is the scale parameter, p is the shape parameter and λ is a shift parameter. The likelihood function of the NSEGL distribution is given as

$$L(X;\mu,\sigma,\lambda,p) = \frac{p^n\lambda^n}{\sigma^n}\frac{e^{-p\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)}}{\prod_{i=1}^{n}\left[\lambda + e^{-\left(\frac{x_i-\mu}{\sigma}\right)}\right]^{(p+1)}}, \tag{6.2}$$

Taking the natural logarithm of both sides of $L(X;\mu,\sigma,\lambda,p)$, we have

$$\ln L(X;\mu,\sigma,\lambda,p) = n\ln p + n\ln\lambda - n\ln\sigma - p\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right) - (p+1)\sum_{i=1}^{n}\ln\left[\lambda + e^{-\left(\frac{x_i-\mu}{\sigma}\right)}\right] \tag{6.3}$$

To obtain the estimates of the parameters that maximize the likelihood function, we differentiate the logarithm of the likelihood function partially with respect to each of the parameters and equate the derivatives to zero and solve for the parameters. Hence by differentiation

$$\frac{\partial \ln L(X;\mu,\sigma,\lambda,p)}{\partial\mu} = \frac{np}{\sigma} - \frac{(p+1)}{\sigma}\sum_{i=1}^{n}\frac{e^{-\left(\frac{x_i-\mu}{\sigma}\right)}}{\left[\lambda + e^{-\left(\frac{x_i-\mu}{\sigma}\right)}\right]} \tag{6.4a}$$

$$\frac{\partial \ln L(X;\mu,\sigma,\lambda,p)}{\partial\sigma} = -\frac{n}{\sigma} + \frac{p\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2} + \frac{(p+1)}{\sigma^2}\sum_{i=1}^{n}\frac{x_i e^{-\left(\frac{x_i-\mu}{\sigma}\right)}}{\left[\lambda + e^{-\left(\frac{x_i-\mu}{\sigma}\right)}\right]} \tag{6.4b}$$

$$\frac{\partial \ln L(X;\mu,\sigma,\lambda,p)}{\partial\lambda} = \frac{n}{\lambda} - (p+1)\sum_{i=1}^{n}\frac{1}{\left[\lambda + e^{-\left(\frac{x_i-\mu}{\sigma}\right)}\right]} \tag{6.4c}$$

$$\frac{\partial \ln L(X;\mu,\sigma,\lambda,p)}{\partial p} = \frac{n}{p} - \sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right) - \sum_{i=1}^{n}\ln\left[\lambda + e^{-\left(\frac{x_i-\mu}{\sigma}\right)}\right] \tag{6.4d}$$

Since the equations (6.4a,b,c) above are nonlinear in the parameters, we used numerical iterative method with the aid of optimization computer programme to estimate the parameters from a given sample.

6.1. APPLICATION OF THE NSEGL MODEL

In a study aimed at investigating the presence of excess synthetic waxes that had been added to beeswax, since the addition of micro crystalline wax raises the melting point of beeswax. If all the beeswax had the same melting point, its determination would be a reasonable way to detect dilutions. For a set of 512 beeswax, whose melting point is expected to be 63°C, the melting point of each of them was determined and the absolute differences from 63°C were tabulated. A probability density function that will fit the data is needed. The histogram of the data shows that the data is negatively skewed, hence we decided to fit the NSEGL distributions to the data. The estimates of the parameters of the distribution as obtained from optimization computer programme are ($\hat{\mu} = 1.100,\ \hat{\sigma} = 0.216,\ \hat{\lambda} = 0.699, \hat{p} = 3.182$). The distribution of the data and the analysis are shown in the table below. The adequacy of the model is tested using the method of chi-square test. From the table below, $\mathcal{X}^2_{calculated}$ for the NSEGL model is 0.3903 while from the table of chi-square, $\mathcal{X}^2_{(6,0.05)} = 12.5916$ which implies that the NSEGL model is very adequate for describing this data.

Class limit of Differences	Class mid-point x	Class frequency observed	NSEGL estimated frequency	$\mathcal{X}^2$ NSEGL model
0.01 – 0.15	0.08	7	7.14	0.0028
0.16 – 0.30	0.23	15	13.91	0.0859
0.31 – 0.45	0.38	27	26.35	0.0160
0.46 – 0.60	0.53	47	47.35	0.0026
0.61 – 0.75	0.68	77	77.06	0.0005
0.76 – 0.90	0.83	107	105.25	0.0292
0.91 – 1.05	0.98	110	108.44	0.0223
1.06 – 1.20	1.13	78	75.28	0.0981
1.21 – 1.35	1.28	34	32.81	0.0432
1.36 – 1.50	1.43	10	9.09	0.0907
Total		512		0.3903

Table. Distribution of absolute differences of melting point of beeswax from the minimum melting point of 63°C

CONCLUSION

We have established some properties of the NSEGL distribution. The moments μ_1, μ_2, μ_3 and μ_4 have been theoretically obtained and the mean μ, variance σ^2, skewness β_1 and kurtosis β_2 have been established. The median, the $100k$-percentage point and the mode of the distribution were presented. The distributions of the r^{th} order statistics$X_{r:n}$, the maximum order statistics $X_{n:n}$ and the minimum order statistics $X_{1:n}$ of the distribution are also established. Estimation of the parameters of the distribution and an application in modeling a chemical data conclude the paper. Further generalization of the NSEGL distribution which contains four parameters to a five-parameter NSEGL distribution is currently under study.

REFERENCES

1. Balakrishnan, N. and Leung M. Y. Order statistics from the Type I generalized Logistic Distribution; Communications in Statistics - Simulation and Computation; Vol. 17(1), 1988, pp. 25-50.
2. Copson, E. T. An Introduction to the theory of functions of a complex variable; Oxford University Press, 1962.
3. David, H. A. Order Statistics; John Wiley, New York, 1970.
4. George, E. O. and Ojo, M. O. On a generalization of the logistic distribution; Annals of Statistical Mathematics; 32, 2, A, 1980, pp. 161-169.
5. Ojo, M. O. Some Relationships between the generalized logistic and other distributions; Statistica; LVII, No. 4, 1997, pp. 573-579.
6. Ojo, M. O. Approximations to the distribution of the sum of the generalized logistic random variables; Kragujevac Journal of Mathematics; Vol. 24, 2002, pp. 135-145.
7. Olapade, A. K. Some relationships between the Type II generalized logistic and other distributions; Journal of Statistical Research; Vol. 36, No. 2, 2002, pp. 213-218.
8. Olapade, A. K. On negatively skewed extended generalized logistic distribution; Kragujevac J. Math.; 27 (2005), 2005, pp. 175-182.