Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 12, Number 5 (2016), pp. 4535-4552

http://www.ripublication.com/gjpam.htm

A Comparative Analysis on the Performance of Lehman type II Inverse Gaussian Model and Standard Inverse Gaussian Model in Terms of Flexibility

Ajitoni S. Amusan*[1] and Zarina M. Khalid[2]

[1, 2] Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 Skudai, Johor Bahru Malaysia
**Corresponding author (Ajitoni S. Amusan)*

Abstract

Searching for flexible parametric models is often the concern of statisticians in data analysis. In order to provide significantly skewed, flexible and heavy-tails models, more parameters are introduced into the existing models. In this paper, a novel univariate model called Lehmann type II inverse Gaussian model is constructed from standard inverse Gaussian distribution. The performance of the new distribution in terms of flexibility is compared with that of the baseline distribution. The construction of the new distribution is achieved by adding a shape parameter to the standard inverse Gaussian through dual transformation of the exponentiated generalized class of distributions. The idea is to verify if the Lehman type II inverse Gaussian model would perform better than the inverse Gaussian distribution in modeling real life situations. Various basic statistical properties of the proposed distribution, including the likelihood function are derived. Parameter estimates are obtained via maximum likelihood estimation method. The ratio of maximized likelihoods and Akaike Information Criteria (AIC) are employed to select the best model. In conclusion, when the two models are applied to two known real lifetime data sets, Lehmann inverse Gaussian shows more flexibility in modeling skew data than inverse Gaussian distribution as evident in the negative sign of likelihoods ratio and lower value of AIC

Keywords: Lehmann type II inverse Gaussian, Univariate distribution, Shape parameter, Skew data & Ratio of maximized likelihoods

1. INTRODUCTION

The concept of comparative performance of competing models in terms of their flexibility plays significant roles in the theory of statistics. Against this backdrop, several authors have constructed various parametric models and compared their performance in terms of flexibility. For instance, the presence of 3 parameters in 3-parameter gamma and 3-parameter Weibull (location, scale and shape parameters) gives each of them more flexibility when compared with their two-parameter baseline distributions in analyzing skewed data [1]. In the same vein, Mudholkar introduced generalized Weibull family with a range of hazard shapes [2]. A three-parameter survival distribution which has as special sub-models, the Weibull, Negative Binomial, Log-logistic was presented by [3] and [4] also introduced Lehmann type II weighted Weibull by extending weighted Weibull model. Pareto Type II distribution is a tail-weighted probability distribution obtained fundamentally from Pareto distribution [5]. For more on different distributions constructed and compared in terms of their performance with their respective baseline models, see the following: [6], [7], [8] and [9, 10]. This paper is designed to compare a newly constructed 3-parameter Inverse Gaussian distribution called Lehmann type II Inverse Gaussian distribution with the standard inverse Gaussian distribution. Here, Lehmann Type II inverse Gaussian is an extension of two-parameter Inverse Gaussian distribution through dual transformation of the exponentiated generalized class of distributions, which is interpreted as a double construction of Lehmann alternatives [7, 11].

The process introduces a shape parameter to the two-parameter Inverse Gaussian distribution with a view of obtaining a distribution that is more flexible in terms of estimates of their characteristics in describing lifetime data. After this section, framework of the new distribution is given, its characteristics such as its moments function, quantile function, after which maximum likelihood estimation of parameters is highlighted, then empirical study of the new distribution follows before its application to two real life data sets

2. METHODS

We provide a comprehensive description of mathematical properties of the subject distribution with the hope that it will attract wider applications in survival, engineering and in other area of research. There are various approaches to the construction of flexible distributions. In this paper, we adopt the approach employed by Coidero [7] as described in the following methodology:

If $F(X)$ is the cumulative density function (cdf) of a continuous distribution, also hereafter referred to as baseline distribution, the exponentiated generalized (EG) class of distributions by [7] is defined as thus:

$$G(X) = \left[1-\left(1-F(X)\right)^{\alpha}\right]^{\beta} \tag{2.1}$$

$G(X)$ is the cumulative density function (cdf) of the new distribution, hereafter referred to as the subject distribution. Where $\alpha > 0, \beta > 0$, are two shape parameters introduced. Here, (2.1) is straightforward since there is no any complex function in it as when compared to the beta generalized family proposed by Eugene [12] which also has two additional shape parameters but appears complicated because of the presence of incomplete beta function in it. It is clearly obvious that (2.1) has tractable properties particularly for simulations because of the simple form of its quantile function, given as

$$x = Q_F\left(\left[1-\left(1-u^{\frac{1}{\beta}}\right)^{\frac{1}{\alpha}}\right]\right) \tag{2.2}$$

where $Q_F(u)$ is the baseline quartile function. When $\alpha = 1$, (2.1) becomes the exponentiated type distributions defined by Gupta [1]. Furthermore, exponentiated Gamma $(E\Gamma)$ and exponentiated exponential (EE) distributions are obtained when Gamma and exponential distributions are respectively taken as the baseline distributions. In the same vein, Nadarajah and Kotz [13] defined exponentiated Gumbel (EGu) and exponentiated Frechet (EF) for $\beta = 1$. Clearly, the class of distributions in (2.1) extends both exponentiated type distributions. The probability density function (pdf) of the subject distribution is

$$g(X) = \frac{d}{dx}G(X) = \alpha\beta[1-F(X)]^{\alpha-1}\left[1-(1-G(X))^{\alpha}\right]^{\beta-1} f(X) \tag{2.3}$$

The exponentiated generalized family of densities allows for greater flexibility of its tails and can be widely applied in different areas of research. It is also important here to note that even if $f(x)$ is symmetric the distribution $g(x)$ will not be symmetric. The two extra parameters in (2.2) can control both tails weights and would likely add entropy to the center of the EG density function. Coidero [7] thereafter defined the exponentiated-G (shortened as Exp-F) distributions for an arbitrary baseline distribution F(x), say $X \sim Exp^c F$, if X has cumulative density function (cdf) $H_c(x) = F(x)^c$ and probability density function (pdf) $h_c(x) = cg(x)G(x)^{c-1}$. $Exp^c F$ is called Lehmann type I distribution. For $c > 1$ and larger values of x, the multiplicative factor $cG(x)^{c-1}$ is greater than one. While, if $c < 1$ and for larger values of x, $cG(x)^{c-1}$ is less than one. Lehmann type II distribution is obtained as a dual transformation $Exp^c(1-G)$ corresponding to the cdf $G(x) = [1-F(x)]^c$. Thus (2.1) encompasses both Lehmann type I ($(Exp^{\beta}F)$ for $\alpha = 1$) and Lehmann type II ($Exp^{\alpha}[1-G]$ for $\beta = 1$) distributions [11]. Thus, the double construction $Exp^{\beta}[Exp^{\alpha}(1-G)]$ generates the EG class of distributions. Several properties of EG family can be easily obtained through this double transition. α and β are positive integers and the family of exponentiated distributions have similar attractive physical interpretation.

If we consider a device that is made up of β independent components in a parallel system and each of the components is also made up of a subcomponents in a parallel system in the same manner. The device fails if all the β components fail and we say that each component fails if any subcomponent fails.

Let $X_{j1},.,.,.,X_{j\alpha}$ represent the lifetimes of the subcomponents within the jth components, $j = 1,2,.,.,.,\beta$ with common cdf $F(x)$. If X_j represents the lifetime of the jth component and X represents the lifetime of the device. Thus, the cumulative density function (cdf) of X can be written as

$$P(X \le x) = P(X_1 \le x,.,.,.,X_\beta \le x) = P(X_1 \le x)^\beta = [1 - P(X_1 > x)]^\beta$$
$$= [1 - P(X_{11} > x,.,.,.,X_{1\alpha} > x)^\alpha]^\beta = [1 - P(X_{11} > x)^\alpha]^\beta$$
$$= [1 - \{1 - P(X_{11} \le x)\}^\alpha]^\beta \tag{2.4}$$

Thus, the lifetime of a device obeys the EG class of distributions

In this paper, we propose a new distribution based on exponentiated generalized family called Exponentiated Generalized Standard inverse Gaussian (or Lehmann type II inverse Gaussian) distribution. The baseline distribution in this case is inverse Gaussian whose pdf is

$$f_{IG}(X;\theta) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right] \quad , \; X > 0 \;\; \theta = (\mu,\lambda) \in \Re^1_+ X \Re^1_+ \subset \Re^2 \tag{2.5}$$

Where $\theta = (\mu,\lambda) \in \Re^1_+ X \Re^1_+ \subset \Re^2$ is the parameter space of the distribution, μ and λ are the mean and the scale parameter respectively, and are of the same physical dimensions as X [14] There are other forms of (2.5) obtainable through re-parameterization, see [15].

The cdf is

$$F_{IG}(X;\theta) = \Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) + e^{\frac{2\lambda}{\mu}} \Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right) \tag{2.6}$$

Where $\Phi(.)$ is the standard Gaussian distribution cdf.

The density is uni-modal which makes it a competitor of models like Log-normal, Log-logistic, generalized Weibull for a lifetime model as described by [16, 17] and the mode is given by the formula

$M_o = \mu\left[\left(1+\frac{9}{4\phi^2}\right)-\frac{3}{2\phi}\right]$; where $\phi = \frac{\lambda}{\mu}$; Mean= $E(X) = \mu$; Variance= $Var(X) = \frac{\mu^3}{\lambda}$

and Coefficient of variation $CV(X) = \frac{\sqrt{Var(X)}}{E(X)} = \sqrt{\frac{\mu}{\lambda}}$

Inverse Gaussian distribution provides a suitable choice as lifetime model for phenomena where early occurrences of events are expected. Here, the failure rate is expected to be non-monotone, initially increasing and later decreasing. For example, situations where product failures or repairs are dominant and incidence of pregnancy or birth to a newly married woman living with husband can be well described by this model. The tails of the distribution decrease more slowly than the normal distribution. This feature makes it suitable to model phenomena where numerically large values are more probable than is the case for the normal distribution. The distribution is excellent for its ability to reach top high and admit some extreme values [18].

Figure 1 and figure 2 show the different looks of density curves when each of the parameters is fixed while varying the other.

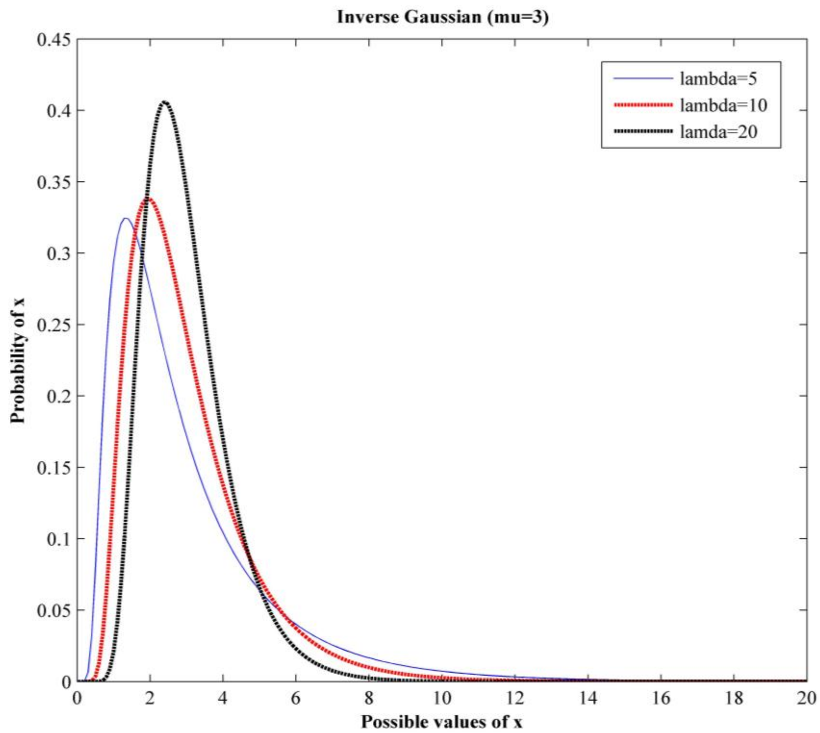

Figure 1: Fixed mu=3 and varying lamda

Figure 1: above shows the density curves when mu parameter is fixed at mu=3 while varying the scale (lamda) parameter.

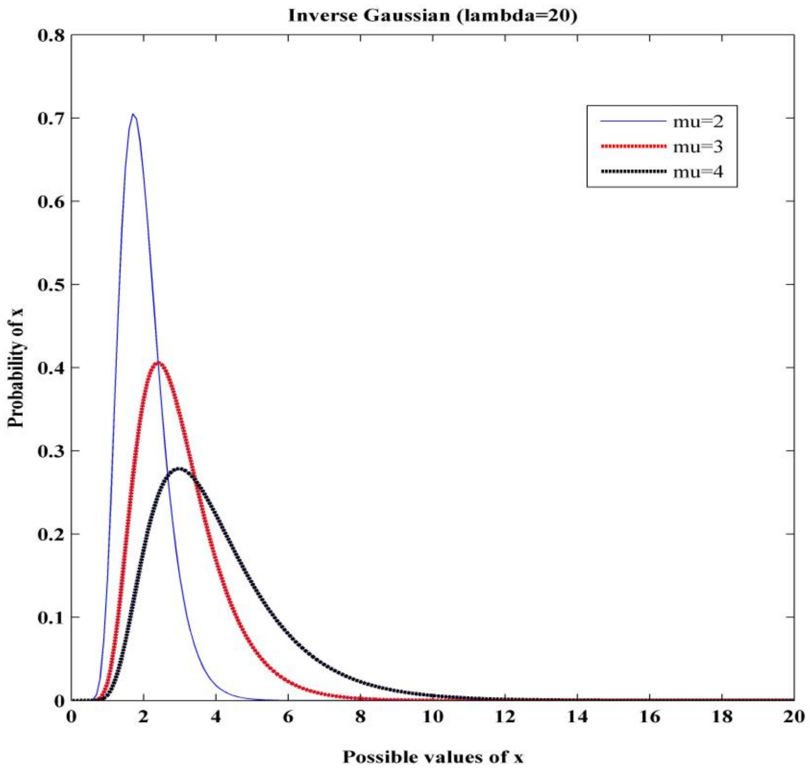

Figure 2: Fixed lamda=20 and varying mu

Figure 2 is showing the curves at fixed value of lamda parameter while varying mu.

2.1 Lehmann type II inverse Gaussian

Recall from equation (2.1) that $G(X)$ is a special case when $\alpha=\beta=1$. That is inverse Gaussian distribution is obtained. Similarly, when $\beta=1$, we obtain exponentiated generalized standard inverse Gaussian referred to as Lehmann type II inverse Gaussian distribution. The cdf and pdf of this new distribution are respectively given as

$$G(X;\mu,\lambda,\alpha)=1-\left\{1-\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right)-e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}+1\right)\right)\right\}^{\alpha} \tag{2.7}$$

$$g(x;\mu,\lambda,\alpha)=\left(\frac{\lambda}{2\pi x^{3}}\right)^{1/2}\exp\left[-\frac{\lambda(x-\mu)^{2}}{2\mu^{2}x}\right]\alpha\left[1-\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right)-e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}+1\right)\right)\right]^{\alpha-1} \tag{2.8}$$

Where μ,λ,α are the location, scale and shape parameters respectively.

The survivorship function of the Lehmann type II inverse Gaussian is defined as following

$$S(x) = P(X > x) = 1 - G(x) = \int_x^{\infty} g(t)dt$$

$$= \left\{1 - \Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) - e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right)\right\}^{\alpha} \tag{2.9}$$

The hazard rate function $h(x)$ of a random variable X of Lehmann type II inverse Gaussian is given as

$$h(x) = \frac{\left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right]\alpha\left[1 - \Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) - e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right)\right]^{\alpha-1}}{\left\{1 - \Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) - e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right)\right\}^{\alpha}} \tag{2.10}$$

2.1.1 Moments

The moments of $X_{i:n}$ can be derived from the probability weighted moments (PWMs), see Greenwood [19], using a result due to Barakat in [20]. If $F(Y)$ is the cdf of random variable Y and $G(X)$ is the cdf of random variable X with pdf as given in equation (2.7) above then moments of the Lehmann type II inverse Gaussian can be found from the $(r, j)th$ probability weighted moment (PWM) of Y defined as

$$\mathrm{T}_{rj} = E(Y^r F(Y)^j) = \int_{-\infty}^{\infty} y^r F(y)^j f(y)dy \tag{2.11}$$

Therefore, rth moment of Lehmann type II inverse Gaussian can be defined as

$$E(X^r) = \alpha \sum_{j=0}^{\infty} t_j \mathrm{T}_{rj} \tag{2.12}$$

That is, the moment has been expressed as an infinite weighted sum of Y, where (2.11) can also be expressed as

$$\mathrm{T}_{rj} = \int_0^{\infty} x^r \left[\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) + e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right)\right]^j \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right]dx$$

Therefore, (2.12) becomes,

$$E(X^r) = \alpha \sum_{j=0}^{\infty} t_j \int_0^{\infty} x^r \left[\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) + e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right)\right]^j \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right]dx$$

$$=\alpha\sum_{j=0}^{\infty}t_j\left[\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right)+e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}+1\right)\right)\right]^j\int_0^{\infty}x^r\left(\frac{\lambda}{2\pi x^3}\right)^{1/2}\exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right]dx$$

When $r=1$ we have

$$E(X)=\alpha\mu\sum_{j=0}^{\infty}t_j I_j \tag{2.13}$$

Where $I_j=\left[\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right)+e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}+1\right)\right)\right]^j$

2.1.2 Quantile Function and Median

The quantile function corresponding to the cdf of Lehmann II inverse Gaussian distribution is given as

$$x=Q_F\left(\left\lfloor 1-(1-u)^{\frac{1}{\alpha}}\right\rfloor\right), \tag{2.14}$$

Where $Q_F(u)$ is the baseline quantile function.

The realization of the quantile can be achieved by implementing the function ‘quantile (LII_sample, prob=seq(0,1,0.25))’in the R-package programming [21]. This is after generating sample from Lehmann II inverse Gaussian distribution which is designated as LII_sample from equation (2.1) for $\beta=1$.

2.2 Estimations and inferences on the Parameters

The paper employs the method of maximum likelihood estimation to find the estimates of all the parameters. Given $X=\{x_1,x_2,.,.,.x_n\}$, a random sample of size n from the Lehmann Type II Inverse Gaussian distribution with parameters μ,λ,α. The likelihood function can be expressed according to [22] as

$$L=\prod_{u}^{n}g(x)\prod_{c}^{n}S(x) \tag{2.15}$$

The log-likelihood is expressed as

$$\log L=l(\theta)=\sum_{u}^{n}\log(g(x))+\sum_{c}^{n}\log(S(x)) \tag{2.16}$$

Where u and c mean sum over uncensored and censored observations respectively.

$$l(\theta) = n\log\alpha + \sum_{i=1}^{n}\log\left(\frac{\lambda}{2\pi x^3}\right)^{1/2}\exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right] + (\alpha-1)\sum_{i=1}^{n}\log\left[\begin{array}{l}1-\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right)- \\ e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}+1\right)\right)\end{array}\right] \quad (2.17)$$

The components of the score vector are

$$U_\alpha(\theta) = \frac{\partial l(\theta)}{\partial\alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n}\log\left[1-\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right) - e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}+1\right)\right)\right] \quad (2.18)$$

$$U_\mu(\theta) = \frac{\partial(\theta)}{\partial\mu} = \frac{-\lambda}{2x}\frac{\partial}{\partial\mu}\left[\frac{(x-\mu)^2}{\mu^2}\right] + \sum_{i=1}^{n}\frac{\partial}{\partial\mu}\left[1-\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right) - e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}+1\right)\right)\right] \quad (2.19)$$

$$U_\lambda(\theta) = \frac{n}{\lambda^{3/2}} - \frac{1}{2\mu^2}\sum_{i=1}^{n}\left(\frac{x-\mu}{x}\right)^2 + \sum_{i=1}^{n}\frac{\partial}{\partial\lambda}\left[1-\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right) - e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}+1\right)\right)\right] \quad (2.20)$$

Solving the above equations simultaneously produce the maximum likelihood estimates of the parameters. For interval estimation and hypothesis testing on the model parameters, a “3 by 3” observed information matrix is required. We therefore proceed to obtained the elements of the observed information matrix $l''(\theta)$ for the parameters $\mu, \lambda, \alpha,$ through second partial derivatives of components of score functions as follows

$$U_{\alpha\alpha}(\theta) = \frac{-n}{\alpha^2} \quad (2.21)$$

$$U_{\mu\mu}(\theta) = \frac{-\lambda}{2x}\frac{\partial^2}{\partial\mu^2}\left[\frac{(x-\mu)^2}{\mu^2}\right] + (\alpha-1)\sum_{i=1}^{n}\frac{\partial^2}{\partial\mu^2}\left[1-\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right) - e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}+1\right)\right)\right] \quad (2.22)$$

$$U_{\lambda\lambda}(\theta) = \frac{-3n}{4\lambda^{5/2}} + \sum_{i=1}^{n}\frac{\partial^2}{\partial\lambda^2}\left[1-\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right) - e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}+1\right)\right)\right] \quad (2.23)$$

2.2.1 Model discrimination

Here, we discriminate between the two models under consideration using their ratio of maximized likelihoods, defined according to [23, 24] , as

$$L = \frac{L_{IG}(\hat{\mu}, \hat{\lambda})}{L_{LIG}(\hat{\mu}, \hat{\lambda}, \hat{\alpha})}; \quad (2.24)$$

Where $L_{IG}(\hat{\mu}, \hat{\lambda})$ and $L_{LIG}(\hat{\mu}, \hat{\lambda}, \hat{\alpha})$ are the respective maximum values of the likelihood functions for inverse Gaussian and Lehmann type II inverse Gaussian models. If $L_{IG}(\hat{\mu},\hat{\lambda}) > L_{LIG}(\hat{\mu},\hat{\lambda},\hat{\alpha})$ then inverse Gaussian model is adjudged better than Lehmann type II inverse Gaussian model. Otherwise, Lehmann type II inverse Gaussian model has a better fit. The natural logarithm of (2.24) is $\ln(L) = T$.

$$T = \ln(L) = \ln\left[\frac{L_{IG}(\hat{\mu},\hat{\lambda})}{L_{LIG}(\hat{\mu},\hat{\lambda},\hat{\alpha})}\right] \tag{2.25}$$

T is either positive or negative depending on the probability of correct selection of a particular model being better than the other. For example, the sign is positive if the model whose likelihood value is in the numerator is preferably and correctly selected as being better than the model whose likelihood value is in the denominator. That is, if $\ln L = T > 0$, data belong to inverse Gaussian model and if $\ln L = T < 0$, then the data set is better fit by Lehmann type II inverse Gaussian model. For more on the distribution of T and computation of probability of correct selection of model, see [25] and [26]

Equation (2.25) can be expressed as difference of the log-likelihoods of the two models, viz;

$$T = L_{IG}(\hat{\mu},\hat{\lambda}) - L_{LIG}(\hat{\mu},\hat{\lambda},\hat{\alpha}) \tag{2.26}$$

$$\begin{aligned} T = {} & \frac{n}{2}\ln\lambda - \frac{n}{2}\ln(2\pi) - \frac{3}{2}\sum_{i=1}^{n}\ln x_i - \sum_{i=1}^{n}\frac{\lambda(x_i-\mu)^2}{2\mu^2 x_i} - n\log\alpha + \\ & \sum_{i=1}^{n}\log\left(\frac{\lambda}{2\pi x^3}\right)^{1/2}\exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right] + (\alpha-1)\sum_{i=1}^{n}\log\left[1-\Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right) - e^{\frac{2\lambda}{\mu}}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}+1\right)\right)\right] \end{aligned} \tag{2.27}$$

All these expressions can be numerically computed in software like R project for statistical computing [21], Matlab [27] and Mathematica [28]

3. RESULTS AND DISCUSSION

3.1 Simulation Study

Tables 1 and 2 present the means and variances respectively of different data sets generated at different fixed values of parameters using the cumulative density function. The data were generated to mimic realization of time (in months) of first birth after marriage of a woman. Assumptions are that the time-to-first birth after marriage lies between 7 and 120 months of marriage. The woman is also assumed to be fecund at the time of marriage and she is exposed to uninterrupted sexual relationship with her husband.

Table 1: Mean of the Lehmann type II inverse Gaussian at different Parameter Values

		$\mu = 25$					$\mu = 30$				
$\lambda =$		40	45	50	55	60	40	45	50	55	60
α	1.0	29.2	32.7	34.4	37.1	38.3	30.1	33.5	37.2	41.6	43.4
	1.1	27.6	31.1	32.7	35.6	36.2	28.8	31.3	33.7	37.2	42.8
	1.2	26.3	28.8	29.0	33.2	34.9	27.2	29.3	32.5	35.6	38.3
	1.3	24.1	27.3	28.3	30.4	33.2	25.5	27.8	28.0	33.4	36.1
	1.4	28.9	33.5	35.2	36.0	38.4	30.3	34.2	36.9	38.1	40.0
	1.5	31.5	37.3	38.6	41.9	42.8	33.0	38.5	41.2	43.7	45.4

As displayed in table 1 above, the mean values of different data sets generated from the Lehmann type II inverse Gaussian model, for varying values of parameters. The mean values increase as the scale and location parameters increase. However, the mean which is initially decreasing, looking at the table vertically downward, as the shape parameter increases tends to rise when the shape parameter exceeds 1.3 (i.e when α >1.3). These impacts of varying parameters on the means of the distributions lay credence on the suitability of the new model to large skew lifetime data with large values.

Table 2: Variance of the Lehmann type II inverse Gaussian at different Parameter Values

		$\mu = 25$					$\mu = 30$				
$\lambda =$		40	45	50	55	60	40	45	50	55	60
	1.0	442.4	448.9	453.2	459.5	466.3	449.1	453.7	464.6	472.3	493.9
	1.1	434.7	418.1	432.5	437.3	443.7	438.9	434.5	442.3	459.7	462.5
α	1.2	425.3	453.8	459.2	461.4	464.1	446.2	452.3	458.5	470.4	484.7
	1.3	441.3	458.3	467.3	472.4	478.2	455.8	463.0	470.0	482.5	492.0
	1.4	448.5	460.2	471.6	479.8	488.4	461.3	469.8	471.4	491.3	510.6
	1.5	453.8	472.3	478.6	486.0	517.1	468.0	473.2	486.1	507.6	519.2

In the same vein, the variances of the data sets behave similarly as the means for various values of the parameters. The variances increase as the location and scale parameters increase.

The following figures depict different shapes of the density curves of Lehmann type II inverse Gaussian model.

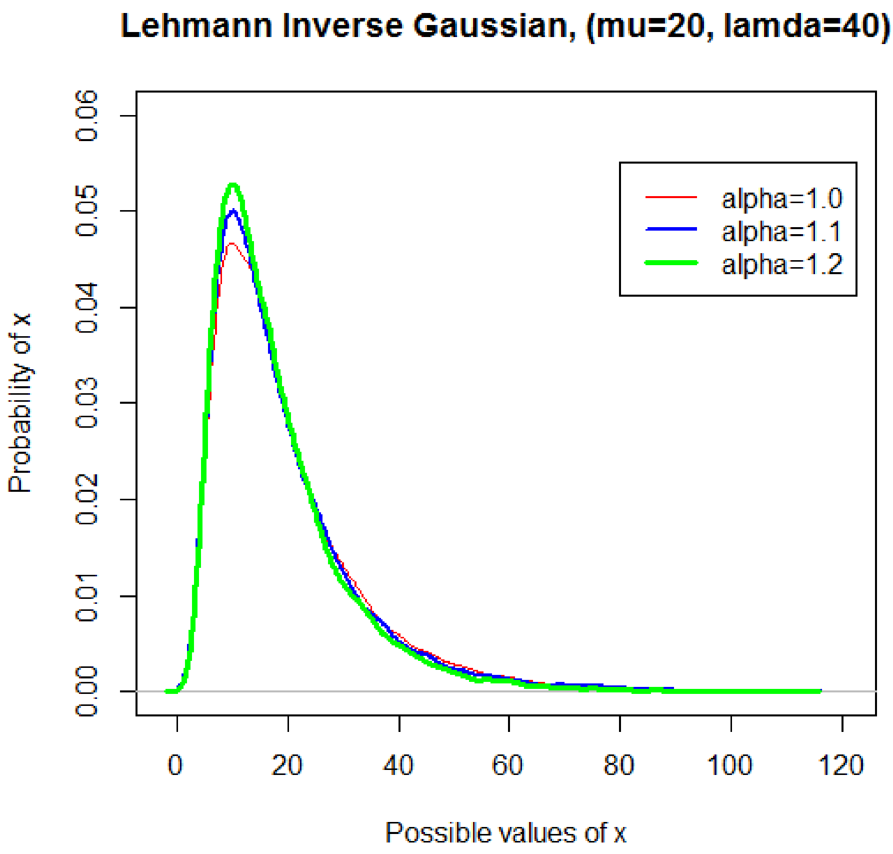

Figure 3: Fixed mu=20, lamda=40 and varying alpha

Figure 3 above corresponds to different shapes of the pdf when the shape parameter is varied at fixed values of location (mu) and scale (lamda) parameters. Here, the shape of the curve that corresponds to alpha=1 is that of inverse Gaussian. The heights of the curves for values of alpha greater than one are higher than that for inverse Gaussian.

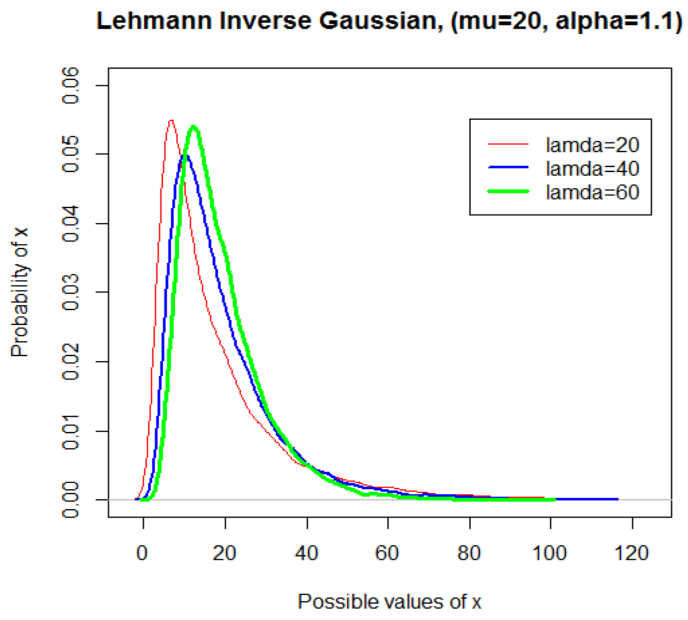

Figure 4: Fixed mu=20, alpha=1.1 and varying lamda

Figure 4 shows the curves of the pdf of the Lehmann type II inverse Gaussian for varying scale (lamda) parameter while other two parameters are fixed.

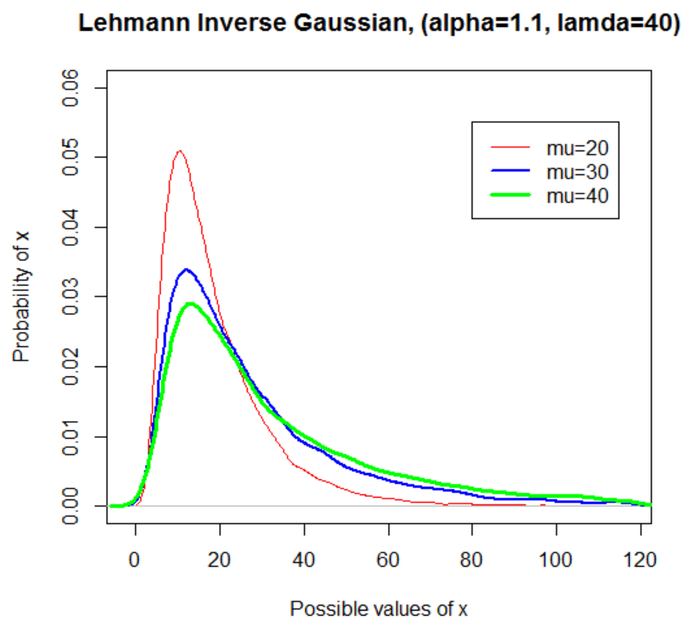

Figure 5: Fixed alpha=1.1, lamda=40 and varying mu

Figure 5 is the figure showing different shapes of pdf curves for the fixed values of alpha and land lamda while varying location (mu) parameter

3.2 Application to Lifetime Data and Results

The Lehmann type II inverse Gaussian distribution and the inverse Gaussian distribution are both applied to two real data sets and the distribution corresponding to the best fit is selected using the Akaike Information Criteria (AIC) and ratio of maximized likelihoods.

Data set 1

The first data set represents the times to first birth after first marriage of women in Nigeria. The data were extracted from the reports of National Demographic and Health Survey (NDHS) [29]. The data is also referred to as first birth interval (FBI). Time to first birth (in months) after first marriage of a woman is one of several variables collected during the survey comprising of 15363 sampled respondents. The data has been studied by Amusan & Zarina [16]. The data is summarized in table 3 and the performance of the two competing distributions is presented in table 4.

Data set II

The second data are the survival times of 72 guinea pigs (in days) infected with virulent tubercle. The data was reported by [30] and has also been studied by [31]. The data are as following

10, 33, 44, 56, 59, 72, 74, 77, 92, 93, 96, 100, 100, 102, 105,107,107, 108, 108, 108, 109, 112, 113, 115, 116, 120, 121, 122, 122, 124, 130, 134, 136, 139, 144, 146, 153, 159, 160, 163, 163, 168, 171, 172, 176, 183, 195, 196, 197, 202, 213, 215, 216, 222, 230, 231, 240, 245, 251, 253, 254, 255, 278, 293, 327, 342, 347, 361, 402, 432, 458, 555.

The data is summarized in table 5 and the performance of the distributions is given in table 6.

Table 3: Summary Statistics of First Birth Interval Data

Min.	Q1	Q2	Q3	Mean	Variance	Skewness	Kurtosis	Max.
7.0	15.3	20.7	33.0	28.7	411	2.7	7.3	120.0

Table 4: Performance of distributions with standard errors in parentheses using FBI Data

Distributions	Parameters	P-value	$l(\theta)$	AIC	T	Chi-square
Inverse Gaussian	$\hat{\lambda} = 45.46(0.5318)$	$< 2\times10^{-16}$	-58306.6	116617.2	-5678.4	11,356.8
	$\hat{\mu} = 28.74(0.1942)$	$< 2\times10^{-16}$				
Lehmann Inverse-Gaussian	$\hat{\lambda} = 46.0(0.4697)$	$< 2\times10^{-16}$	-52628.2	105262.4		
	$\hat{\mu} = 31.0(1.6289)$	$< 2\times10^{-16}$				
	$\hat{\alpha} = 1.1(0.0851)$	$< 2\times10^{-16}$				

Table 5: Summary Statistics of survival times of 72 Guinea pigs

Min.	Q1	Q2	Q3	Mean	Variance	Skewness	Kurtosis	Max.
10.0	105.8	148.9	223.0	177.2	9483.1	1.297	4.739	555.0

Table 6: Performance of distributions with standard errors in parentheses using Guinea pigs Data

Distributions	Parameters	P-value	$l(\theta)$	AIC	T	Chi-square
Inverse Gaussian	$\hat{\lambda} = 261.6(0.117)$	$< 2\times10^{-16}$	-1073.5	2151.0	-54.3	108.6
	$\hat{\mu} = 177.4(0.092)$	$< 2\times10^{-16}$				
Lehmann Inverse-Gaussian	$\hat{\lambda} = 218.3(0.0205)$	$< 2\times10^{-16}$	-1019.2	2044.4		
	$\hat{\mu} = 52.0(1.9204)$	$< 2\times10^{-16}$				
	$\hat{\alpha} = 1.2(0.1130)$	$< 2\times10^{-16}$				

Looking at tables 4 and 6, the model with lowest AIC or the highest log-likelihood or the negative sign of ratios of maximized likelihoods is considered to have a better fit than the other. From these two tables it is clear that Lehmann type II inverse Gaussian model fits the two data sets better than the inverse Gaussian.

CONCLUSION

There are a large number of studies on model construction through a diverse range of methods using various distributions as baseline models. In this paper, a novel univariate model called Lehman type II inverse Gaussian has been successfully constructed and compared with the its baseline, inverse Gaussian model in terms of flexibility and applicability to real life data. It is evident that Lehmann type II inverse Gaussian model outperforms the inverse Gaussian model for the two data sets considered as showed in the magnitude of AIC and log-likelihood including the negative sign of ratio of maximized likelihood. The paper concludes that Lehmann type II inverse Gaussian can be more suitable for modeling situations where large values occur at the tail end and whose probability of some events can be high. Nonetheless, the authors still recognize the usefulness and popularity of inverse Gaussian model in modeling diverse phenomena.

ACKNOWLEDGEMENT:

The authors wish to acknowledge the financial intervention of the Federal Ministry Education, Nigeria, through tertiary education funding program. We appreciate the management of the Universiti Teknologi Malaysia for the opportunity and atmosphere provided.

REFERENCES

[1] Gupta, R.D. and D. Kundu, 1999, "Theory & methods: Generalized exponential distributions″. Australian & New Zealand Journal of Statistics. **41**(2): pp. 173-188.

[2] Mudholkar, G.S., D.K. Srivastava, and G.D. Kollia, 1996, "A generalization of the Weibull distribution with application to the analysis of survival data″. Journal of the American Statistical Association. **91**(436): pp. 1575-1583.

[3] Sparling, Y.H., et al., 2006, "Parametric survival models for interval-censored data with time-dependent covariates″. Biostatistics. **7**(4): pp. 599-614.

[4] Badmus, N., T. Bamiduro, and S. Ogunobi, 2014, "Lehmann Type II weighted Weibull distribution″. International Journal of Physical Sciences. **9**(4): pp. 71-78.

[5] Kim, H.-C., 2016, "A Performance Analysis of Software Reliability Model using Lomax and Gompertz Distribution Property″. Indian Journal of Science and Technology. **9**(20).

6. Alexander, C., et al., 2012, "Generalized beta-generated distributions″. Computational Statistics & Data Analysis. **56**(6): pp. 1880-1897.

[7] Cordeiro, G.M., E.M. Ortega, and D.C. da Cunha, 2013, "The exponentiated generalized class of distributions″. Journal of Data Science. **11**(1): pp. 1-27.

[8] Cordeiro[, G.M., E.M. Ortega, and G.O. Silva, 2014, "The Kumaraswamy modified Weibull distribution: theory and applications″. Journal of Statistical Computation and Simulation. **84**(7): pp. 1387-1411.

[9] Saboor, A., S.B. Provost, and M. Ahmad, 2015, "Pak. J. Statist. 2015 Vol. 31 (1), 21-32 A Confluent Hypergeometric Generalized Inverse Gaussian Distribution″. Pak. J. Statist. **31**(1): pp. 21-32.

[10] Saboor, A., M. Kamal, and M. Ahmad, 2015, "Pak. J. Statist. 2015 Vol. 31 (2), 229-250 The Transmuted Exponential–Weibull Distribution With Applications″. Pak. J. Statist. **31**(2): pp. 229-250.
[11] Lehmann, E.L., 1953, "The power of rank tests″. The Annals of Mathematical Statistics: pp. 23-43.
[12] Eugene, N., C. Lee, and F. Famoye, 2002, "Beta-normal distribution and its applications″. Communications in Statistics-Theory and Methods. **31**(4): pp. 497-512.
[13] Nadarajah, S. and S. Kotz, 2006, "The exponentiated type distributions″. Acta Applicandae Mathematica. **92**(2): pp. 97-111.
[14] Tweedie, M.C., 1957, "Statistical Properties of Inverse Gaussian Distributions. I″. The Annals of Mathematical Statistics: pp. 362-377.
[15] Folks, J. and R. Chhikara, 1978, "The inverse Gaussian distribution and its statistical application--a review″. Journal of the Royal Statistical Society. Series B (Methodological): pp. 263-289.
[16] Amusan, A.S., et al., 2015, "Discrimination between Lifetime Distributions with Ratios of Maximized Likelihoods″. Research Journal of Applied Sciences. **10**(7): pp. 287-293.
[17] Lemeshko, B.Y., et al., 2010, "Inverse Gaussian model and its applications in reliability and survival analysis″, in Mathematical and Statistical Models and Methods in Reliability. Springer. p. 433-453.
[18] Tweedie, M.C., 1957, "Statistical properties of inverse Gaussian distributions. II″. The Annals of Mathematical Statistics. **28**(3): pp. 696-705.
[19] Greenwood, J.A., et al., 1979, "Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form″. Water Resources Research. **15**(5): pp. 1049-1054.
[20] Barakat, H. and Y. Abdelkader, 2004, "Computing the moments of order statistics from nonidentical random variables″. Statistical Methods and applications. **13**(1): pp. 15-26.
[21] Team, R.C., 2013, "R: A language and environment for statistical computing″.
[22] Tableman, M., 2008, "Survival Analysis Using S/R″. Department of Mathematics & Statistics Portland State University Portland, Oregon, USA (http://stat. ethz. ch/wbl/Skript_SurvivalAnalysis. pdf).
[23] Dey, A.K. and D. Kundu, 2009, "Discriminating between the log-normal and log-logistic distributions″. Communications in Statistics-Theory and Methods. **39**(2): pp. 280-292.
[24] Goh, S., H. Kwon, and M. Choi, 2014, "Discriminating between Weibull distributions and log-normal distributions emerging in branching processes″. Journal of Physics A: Mathematical and Theoretical. **47**(22): pp. 225101.
[25] Dumonceaux, R., C.E. Antle, and G. Haas, 1973, "Likelihood Ratio Test for DiscriminaGon Between Two Models with Unknown Location and Scale Parameters″. Technometrics. **15**(1): pp. 19-27.
[26] Dey, A.K. and D. Kundu, 2009, "Discriminating among the log-normal, weibull, and generalized exponential distributions″. Reliability, IEEE Transactions on. **58**(3): pp. 416-424.

[27] MathWorks, I., 2005, "MATLAB: the language of technical computing. Desktop tools and development environment, version 7″. Vol. 9: MathWorks.

[28] Wolfram Research, I., 2015, "Mathematica″. Version 10.3 ed., Champaign, Illinois: Wolfram Research, Inc.

[29] Commission, N.P., 2009, "ICF Macro. Nigeria Demographic and Health Survey 2008″. National Population Commission, ICF Macro, Abuja, Nigeria.

[30] Bjerkedal, T., 1960, "Acquisition of Resistance in Guinea Pies infected with Different Doses of Virulent Tubercle Bacilli″. American Journal of Hygiene. **72**(1): pp. 130-48.

[31] Tahir, M., et al., 2016, "The Logistic-X family of distributions and its applications″. Communications in Statistics-Theory and Methods(just-accepted).