Available online at http://www.ssstj.sci.ssru.ac.th

Suan Sunandha Science and Technology Journal

An Extended Mixture Inverse Gaussian Distribution

Chookait Pudprommarat*

Faculty of Science and Technology, Suan Sunandha Rajabhat University
U-thong Nok Road, Dusit, Bangkok 10300, Thailand

Corresponding author E-mail: *chookait.pu@ssru.ac.th

Abstract: This paper proposes an extend mixture inverse Gaussian (EMIG) distribution which is mixed between the inverse Gaussian distribution and the length biased inverse Gaussian (LBIG) distribution. The Birnbaum-Saunders (BS) distribution and LBIG distribution are presented as special cases of the EMIG distribution. The properties of this distribution are discussed which include the shapes of the probability density functions, distribution functions survival functions and hazard rate functions, mean and variance. The EMIG has two parameters and it is shown that maximum likelihood estimation (MLE) can be obtained by solving equation. An application of the model to a real data set is analyzed using the new distribution, which shows that the EMIG distribution can be used quite effectively in analyzing real data by using Akaike's information criterion (AIC) statistics and goodness of fit tests.

Keywords: Mixture inverse Gaussian distribution, Length biased inverse Gaussian distribution, Birnbaum-Saunders distribution

1. Introduction

The inverse Gaussian (IG) distribution is a continuous non-negative random variable which is a right skewed distribution and it plays an important role in reliability analysis. Jorgensen et al. (1991), Gupta and Akman (1995), and Henze et al. (2002) studied the IG distribution.

Let $X_1 \sim \mathrm{IG}(a,b)$, i.e. X_1 has a IG distribution with the parameters $a > 0$, $b > 0$ and its probability density function (pdf) and distribution function are given by

$$f_{X_1}(x_1;a,b)=\frac{a}{b\sqrt{2\pi}}\left(\frac{a}{x_1}\right)^{3/2}\exp\left[-\frac{1}{2}\left(\sqrt{\frac{x_1}{b}}-a\sqrt{\frac{b}{x_1}}\right)^2\right],$$

$$F_{X_1}(x_1;a,b)=\Phi\left(\sqrt{\frac{x_1}{b}}-a\sqrt{\frac{b}{x_1}}\right)+\Phi\left(-\left(\sqrt{\frac{x_1}{b}}+a\sqrt{\frac{b}{x_1}}\right)\right)\exp(2a),$$

where Φ (.) is the distribution function of the standard normal distribution.

The length biased inverse Gaussian (LBIG) distribution is a weighted distribution by mean of IG distribution which has received considerable attention due to its various applications in different biomedical areas, such as family history of diseases, early detection of diseases, latency periods of AIDS etc. (Akman and Gupta, 1992; Gupta and Akman,1995). The LBIG distribution was studied by Jorgensen et al. (1991), Akman and Gupta (1992) and Gupta and Akman (1995).

Let $X_2 \sim \mathrm{LBIG}(a,b)$, then X_2 has a LBIG distribution with the parameters $a > 0$, $b > 0$ and its pdf and distribution function are given as follows

$$f_{X_2}(x_2;a,b)=\frac{1}{b\sqrt{2\pi}}\left(\frac{b}{x_2}\right)^{1/2}\exp\left[-\frac{1}{2}\left(\sqrt{\frac{x_2}{b}}-a\sqrt{\frac{b}{x_2}}\right)^2\right],$$

$$F_{X_2}(x_2;a,b)=\Phi\left(\sqrt{\frac{x_2}{b}}-a\sqrt{\frac{b}{x_2}}\right)-\exp(2a)\Phi\left(-\left(\sqrt{\frac{x_2}{b}}+a\sqrt{\frac{b}{x_2}}\right)\right).$$

The mixture inverse Gaussian (MIG) distribution, also known as the weighted inverse Gaussian distribution or the three-parameters generalized inverse Gaussian distribution or, which is mixed between the IG distribution and the LBIG distribution which was studied by Jorgensen et al. (1991), and Gupta and Akman, (1995).

Let $X \sim \mathrm{MIG}(a,b,p)$, then X has a MIG distribution with the parameters $a > 0$, $b > 0$ and $0 \le p \le 1$, which pdf is given by

$$f_X(x;a,b,p)=pf_{X_1}(x;a,b)+(1-p)f_{X_2}(x;a,b),$$

where X_1 is random variable of IG distribution and X_2 is random variable of LBIG distribution

Then, the pdf of MIG can be written in the form:

$$f_X(x)=\frac{1}{b\sqrt{2\pi}}\left[pa\left(\frac{b}{x}\right)^{3/2}+(1-p)\left(\frac{b}{x}\right)^{1/2}\right]$$

$$f_X(x)=\frac{1}{b\sqrt{2\pi}}\left[pa\left(\frac{b}{x}\right)^{3/2}+(1-p)\left(\frac{b}{x}\right)^{1/2}\right]$$

$$\times\exp\left[-\frac{1}{2}\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)^2\right],\quad x>0,$$

Suan Sunandha Science and Technology Journal

where $a > 0$, $b > 0$ and $0 \le p \le 1$.

For MIG distribution has several interesting special cases. In particular, the MIG (a,b,p) distribution becomes the LBIG distribution when $p=0$, the IG distribution when $p=1$ and Birnbaum-Saunders (BS) distribution (Gupta, 2011) when $p=0.5$. Although, the MIG distribution has many desirable properties in applications, parameter estimation may still have problems which have mentioned that finding the efficient initial guesses and solving the non-linear equations simultaneously are non-trivial issues (Jorgensen et al.,1991; Gupta and Akman,1995). Therefore, in order to solve such problems, a new weight parameter p is considered. We propose an extended mixture inverse Gaussian distribution which is obtained by adding a new weight parameter p to the mixture inverse Gaussian distribution.

In this paper, we present an extended mixture inverse Gaussian (EMIG) distribution. Several properties of the new distribution including the probability density functions, distribution functions, survival functions, hazard rate functions, cumulants and moments are provided. In addition, we use maximum likelihood estimation for parameter estimation and present the comparison analysis between the extended mixture inverse Gaussian distributions based on a real data set using Akaike's Information Criterion (AIC) statistics and goodness of fit tests.

2. Material and Methods

In this part, we introduce the definition of the an extended mixture inverse Gaussian distribution denoted by $X \sim \text{EMIG}(a,b)$. We begin with a general definition of the EMIG distribution which will consequently reveal its probability density function.

Definition 1. Let X_1 and X_2 be independent random variables such that $X_1 \sim \text{IG}(a,b)$ and $X_2 \sim \text{LBIG}(a,b)$. Then the new random variable X is said to have an EMIG distribution with parameter $a > 0$ and $b > 0$ if the pdf of X is defined by

$$f_X(x;a,b) = \left(\frac{a}{a+1}\right) f_{X_1}(x;a,b) + \left(\frac{1}{a+1}\right) f_{X_2}(x;a,b).$$

Theorem 1. Let X be a random variable of the EMIG distribution with parameters a and b. The pdf of X is given by

$$f(x) = \frac{1}{(a+1)b\sqrt{2\pi}}\left[a^2\left(\frac{b}{x}\right)^{3/2} + \left(\frac{b}{x}\right)^{1/2}\right] \times \exp\left[-\frac{1}{2}\left(\sqrt{\frac{x}{b}} - a\sqrt{\frac{b}{x}}\right)^2\right], \quad x > 0$$

where $a > 0$ and $b > 0$.

Proof: From Definition 1, the pdf of the EMIG distribution can be obtained by

$$f(x) = \left(\frac{a}{a+1}\right)\frac{a}{b\sqrt{2\pi}}\left(\frac{b}{x}\right)^{3/2}\exp\left[-\frac{1}{2}\left(\sqrt{\frac{x}{b}} - a\sqrt{\frac{b}{x}}\right)^2\right]$$

$$+\left(\frac{1}{a+1}\right)\frac{1}{b\sqrt{2\pi}}\left(\frac{b}{x}\right)^{1/2}\exp\left[-\frac{1}{2}\left(\sqrt{\frac{x}{b}} - a\sqrt{\frac{b}{x}}\right)^2\right]$$

$$= \frac{1}{(a+1)b\sqrt{2\pi}}\left[a^2\left(\frac{b}{x}\right)^{3/2} + \left(\frac{b}{x}\right)^{1/2}\right]\exp\left[-\frac{1}{2}\left(\sqrt{\frac{x}{b}} - a\sqrt{\frac{b}{x}}\right)^2\right].$$

Corollary 1. If $a=0$ then EMIG distribution reduces to LBIG distribution with parameter $a = 0$ and $b > 0$ with pdf given by

$$f(x) = \frac{1}{b\sqrt{2\pi}}\left[\left(\frac{b}{x}\right)^{1/2}\right]\exp\left[-\frac{1}{2}\left(\sqrt{\frac{x}{b}}\right)^2\right]$$

Corollary 2. If $a=1$ then EMIG distribution reduces to BS distribution with parameter $a = 1$ and $b > 0$ with pdf given by

$$f(x) = \frac{1}{2b\sqrt{2\pi}}\left[\left(\frac{b}{x}\right)^{3/2} + \left(\frac{b}{x}\right)^{1/2}\right]\exp\left[-\frac{1}{2}\left(\sqrt{\frac{x}{b}} - \sqrt{\frac{b}{x}}\right)^2\right]$$

Some parameters of the EMIG distribution and their probability density functions are provided in Figure 1.

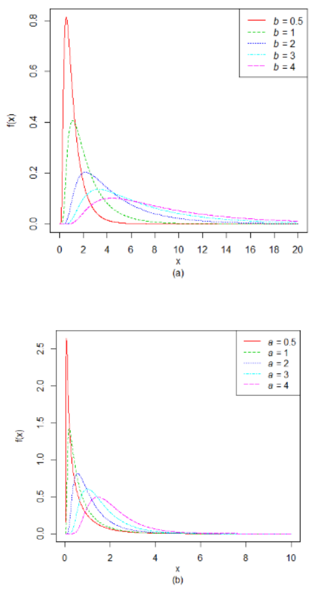

Figure 1. The probability density functions of the EMIG distribution for some values of parameters (a) $a=2$ and (b) $b=0.5$.

Suan Sunandha Science and Technology Journal

Theorem 2. Let X be a random variable of the EMIG distribution with parameters a and b. The distribution function of X is given by

$$F(x)=\Phi\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)-\left\{\frac{1-a}{a+1}\times\exp(2a)\left[1-\Phi\left(\sqrt{\frac{x}{b}}+a\sqrt{\frac{b}{x}}\right)\right]\right\}, x>0$$

where Φ (.) is the distribution function of the standard normal distribution.

Proof: Let X is a continuous non-negative random variable, then the distribution function of X is given by

$$F(x)=\int_0^x f(t)\,dt.$$

If the distribution function of X is EMIG distribution, which is expressed by

$$F(x)=\frac{a}{a+1}\int_0^x \frac{a}{b\sqrt{2\pi}}\left(\frac{b}{t}\right)^{3/2}\exp\left[-\frac{1}{2}\left(\sqrt{\frac{t}{b}}-a\sqrt{\frac{b}{t}}\right)^2\right]dt$$
$$+\frac{1}{a+1}\int_0^x \frac{1}{b\sqrt{2\pi}}\left(\frac{b}{t}\right)^{1/2}\exp\left[-\frac{1}{2}\left(\sqrt{\frac{t}{b}}-a\sqrt{\frac{b}{t}}\right)^2\right]dt.$$

Using the distribution function of IG and LBIG distributions, then distribution function of EMIG becomes

$$F(x)=\frac{a}{a+1}\left[\Phi\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)+\Phi\left(-\left(\sqrt{\frac{x}{b}}+a\sqrt{\frac{b}{x}}\right)\right)\exp(2a)\right]$$
$$+\frac{1}{a+1}\left[\Phi\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)-\exp(2a)\Phi\left(-\left(\sqrt{\frac{x}{b}}+a\sqrt{\frac{b}{x}}\right)\right)\right]$$
$$=\Phi\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)-\frac{1-a}{a+1}\exp(2a)\left[1-\Phi\left(\sqrt{\frac{x}{b}}+a\sqrt{\frac{b}{x}}\right)\right]$$

The distribution functions of the EMIG with some parameter values are shown in Figure 2.

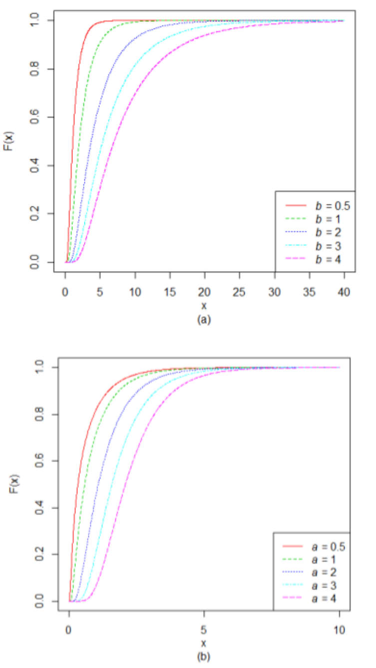

Figure 2. Distribution functions of EMIG for some values of parameters: (a) $a=2$ and (b) $b=0.5$.

Theorem 3. Let X be a random variable of the EMIG distribution with parameters a and b. The survival function of X is given by

$$S(x)=1-\Phi\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)+\left\{\frac{1-a}{a+1}\times\exp(2a)\left[1-\Phi\left(\sqrt{\frac{x}{b}}+a\sqrt{\frac{b}{x}}\right)\right]\right\}.$$

Proof: Let X is a continuous random variable with distribution function $F(x)$ on the interval $[0,\infty)$ then the survival function is defined by

$S(x)=\int_x^\infty f(t)\,dt=1-F(x)$. From the distribution function of EMIG in Theorem 2, then the survival function of X is given by

$$S(x)=1-\left\{\Phi\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)-\frac{1-a}{a+1}\exp(2a)\left[1-\Phi\left(\sqrt{\frac{x}{b}}+a\sqrt{\frac{b}{x}}\right)\right]\right\}$$
$$=1-\Phi\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)+\frac{1-a}{a+1}\exp(2a)\left[1-\Phi\left(\sqrt{\frac{x}{b}}+a\sqrt{\frac{b}{x}}\right)\right].$$

Theorem 4. Let X be a random variable of the EMIG distribution with parameters a and b. The hazard rate function of X can be written as

Suan Sunandha Science and Technology Journal

$$h(x)=\frac{\frac{1}{(a+1)b\sqrt{2\pi}}\left[a^2\left(\frac{b}{x}\right)^{3/2}+\left(\frac{b}{x}\right)^{1/2}\right]\exp\left[-\frac{1}{2}\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)^2\right]}{1-\Phi\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)+\frac{1-a}{a+1}\exp(2a)\left[1-\Phi\left(\sqrt{\frac{x}{b}}+a\sqrt{\frac{b}{x}}\right)\right]}.$$

Proof: Let X is an continuous non-negative random variable with the probability density function and survival function then the hazard rate function can be defined as

$$h(x)=\frac{f(x)}{S(x)},$$

By using $f(x)$ in Theorem 1 and $S(x)$ in Theorem 3, we have

$$h(x)=\frac{\frac{1}{b(a+1)\sqrt{2\pi}}\left[a^2 b\left(\frac{b}{x}\right)^{3/2}+\left(\frac{b}{x}\right)^{1/2}\right]\exp\left[-\frac{1}{2}\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)^2\right]}{1-\Phi\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)+\frac{1-a}{a+1}\exp(2a)\left[1-\Phi\left(\sqrt{\frac{x}{b}}+a\sqrt{\frac{b}{x}}\right)\right]}.$$

Some Survival functions and hazard rate functions plots of the EMIG distribution with some parameter values are displayed in Figure 3 and Figure 4

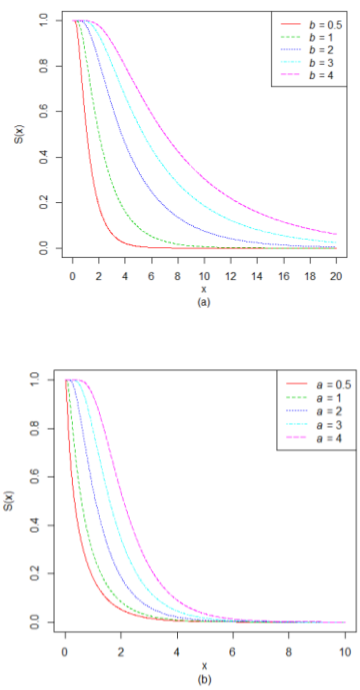

Figure 3. Survival functions of the EMIG distribution for some values of parameters: (a) $a=2$ and (b) $b=0.5$.

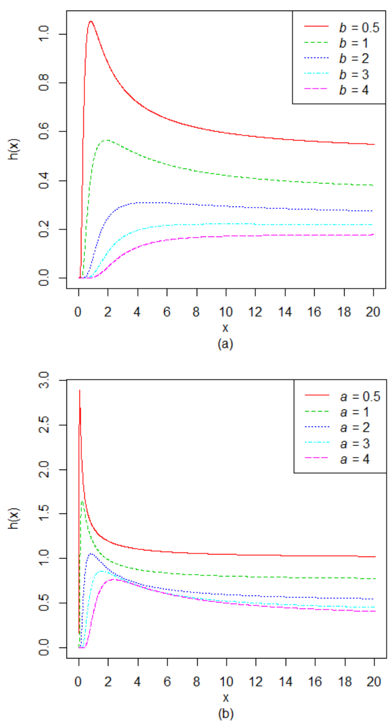

Figure 4. Hazard rate functions of the EMIG distribution for some values of parameters: (a) $a=2$ and (b) $b=0.5$.

3. Results

3.1 Theoretical results

3.1.1 Statistical Properties of the EMIG

The characteristic function, cumulants, moments, mean and variance of EMIG distribution are studied in this section.

Theorem 5. Let X be a random variable of the EMIG distribution with parameters a and b. The characteristic function of X can be written in the form

$$\varphi_X(t)=\frac{\exp\left[a\left(1-\sqrt{1-2bti}\right)\right]}{\sqrt{1-2bti}}\left(\frac{1+a\sqrt{1-2bti}}{a+1}\right).$$

Proof: The characteristic function of a random variable X is defined by

$$\varphi_X(t)=E\left(e^{itx}\right).$$

The distribution of X is a EMIG distribution, the characteristic function takes the form

$$\varphi_X(t)=\int_0^{\infty}\exp(itx)\frac{1}{(a+1)b\sqrt{2\pi}}\left[a^2\left(\frac{b}{x}\right)^{3/2}+\left(\frac{b}{x}\right)^{1/2}\right]\times\exp\left[-\frac{1}{2}\left(\sqrt{\frac{x}{b}}-a\sqrt{\frac{b}{x}}\right)^2\right]dx$$

Suan Sunandha Science and Technology Journal

$$= \frac{1}{(a+1)b\sqrt{2\pi}}\int_0^\infty a^2\left(\frac{b}{x}\right)^{3/2}\exp\left[itx-\frac{1}{2}\left(\frac{x}{b}-2a+\frac{a^2b}{x}\right)\right]dx$$

$$+\frac{1}{(a+1)b\sqrt{2\pi}}\int_0^\infty \left(\frac{b}{x}\right)^{1/2}\exp\left[itx-\frac{1}{2}\left(\frac{x}{b}-2a+\frac{a^2b}{x}\right)\right]dx$$

$$=\frac{a^2b^{1/2}\exp(a)}{(a+1)\sqrt{2\pi}}\int_0^\infty x^{-3/2}\exp\left[-\left(\frac{1}{2b}-ti\right)x-\frac{a^2b/2}{x}\right]dx$$

$$+\frac{\exp(a)}{\sqrt{b}(a+1)\sqrt{2\pi}}\int_0^\infty x^{-1/2}\exp\left[-\left(\frac{1}{2b}-ti\right)x-\frac{a^2b/2}{x}\right]dx.$$

From the Table of integrals, series, and products by Gradshteyn and Ryzhik 2007, pp.369), the formulas are taken from the following form:

$$\int_0^\infty x^{-n-1/2}\exp(-\rho x-q/x)dx=(-1)^n\sqrt{\frac{\pi}{\rho}}\frac{\partial^n}{\partial q^n}\exp\left(-2\sqrt{\rho q}\right),$$

where $\operatorname{Re}\rho>0, \operatorname{Re}q>0.$

The characteristic function can become

$$\varphi_X(t)=\frac{a^2b^{1/2}\exp(a)}{(a+1)\sqrt{2\pi}}\sqrt{\frac{2\pi}{a^2b}}\exp\left(-a\sqrt{1-2bti}\right)$$

$$+\frac{\exp(a)}{\sqrt{b}(a+1)\sqrt{2\pi}}\sqrt{\frac{2b\pi}{1-2bti}}\exp\left(-a\sqrt{1-2bti}\right)$$

$$=\frac{\exp\left[a\left(1-\sqrt{1-2bti}\right)\right]}{\sqrt{1-2bti}}\left(\frac{1+a\sqrt{1-2bti}}{a+1}\right).$$

Theorem 6. Let X be a random variable of the EMIG distribution with parameters a and b. The cumulant generating function of X can be given by

$$K_X(t)=\log\left\{\left[\frac{a}{a+1}+\frac{(1-2bti)^{-1/2}}{a+1}\right]\exp\left[a\left(1-(1-2bti)^{1/2}\right)\right]\right\}.$$

Proof: The cumulant generating function of a random variable X is defined as

$$K_X(t)=\log\varphi_X(t)$$

$$K_X(t)=\log\left\{\frac{\exp\left[a\left(1-\sqrt{1-2bti}\right)\right]}{\sqrt{1-2bti}}\left(\frac{1+a\sqrt{1-2bti}}{a+1}\right)\right\}$$

$$=\log\left\{\left[\frac{1+a\sqrt{1-2bti}}{(a+1)\sqrt{1-2bti}}\right]\exp\left[a\left(1-\sqrt{1-2bti}\right)\right]\right\}$$

$$=\log\left\{\left[\frac{a}{a+1}+\frac{(1-2bti)^{-1/2}}{a+1}\right]\exp\left[a\left(1-\sqrt{1-2bti}\right)\right]\right\}.$$

Recall that a Maclaurin series, is defined as $\sum_{n=0}^{\infty}\frac{f^{(n)}}{n!}x^n$. For $n=4$, we have

$$f(x)=f(0)+\frac{f^{(1)}(0)}{1!}x+\frac{f^{(2)}(0)}{2!}x^2+\frac{f^{(3)}(0)}{3!}x^3+\frac{f^{(4)}(0)}{4!}x^4+O(x^5).$$

Thus, the Maclaurin series of $(1-x)^{-1/2}$ and $(1-x)^{1/2}$, for $|x|<1$, can be written in the following form;

$$(1-x)^{-1/2}=1+\frac{1}{2}x+\frac{1\times3}{2\times4}x^2+\frac{1\times3\times5}{2\times4\times6}x^3+\frac{1\times3\times5\times7}{2\times4\times6\times8}x^4+O(x^5),$$

$$(1-x)^{1/2}=1-\frac{1}{2}x-\frac{1\times1}{2\times4}x^2-\frac{1\times1\times3}{2\times4\times6}x^3-\frac{1\times1\times3\times5}{2\times4\times6\times8}x^4-O(x^5).$$

Next, we consider $(1-2bti)^{-1/2}$ and $(1-2bti)^{1/2}$ in term of $(1-x)^{-1/2}$ and $(1-x)^{1/2}$ respectively, then the cumulant generating function of X becomes

$$\log\varphi_X(t)=\log\left[\frac{a}{a+1}+\frac{1}{a+1}\left(1+\frac{1}{2}(2bti)+\frac{1\times3}{2\times4}(2bti)^2\right.\right.$$

$$\left.\left.+\frac{1\times3\times5}{2\times4\times6}(2bti)^3+\frac{1\times3\times5\times7}{2\times4\times6\times8}(2bti)^4+\cdots\right)\right]$$

$$+a\left[1-\left(1-\frac{1}{2}(2bti)-\frac{1\times1}{2\times4}(2bti)^2\right.\right.$$

$$\left.\left.-\frac{1\times1\times3}{2\times4\times6}(2bti)^3-\frac{1\times1\times3\times5}{2\times4\times6\times8}(2bti)^4-\cdots\right)\right]$$

$$\approx\log\left[1+\frac{1}{a+1}\left((bti)+\frac{3}{2}(bti)^2+\frac{5}{2}(bti)^3+\frac{35}{8}(bti)^4\right)\right]$$

$$+a\left[(bti)+\frac{1}{2}(bti)^2+\frac{1}{2}(bti)^3+\frac{5}{8}(bti)^4\right],$$

and using the expansion

$$\log(1+x)=1+x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+O(x^5),$$ we obtain

$$\log\varphi_X(t)$$

$$\approx\left[1+\frac{1}{a+1}\left((bti)+\frac{3}{2}(bti)^2+\frac{5}{2}(bti)^3+\frac{35}{8}(bti)^4\right)\right.$$

$$-\frac{1}{2}\left(\frac{1}{a+1}\right)^2\left((bti)^2+3(bti)^3+\frac{29}{4}(bti)^4\right)$$

$$\left.+\frac{1}{3}\left(\frac{1}{a+1}\right)^3\left((bti)^3+\frac{9}{2}(bti)^4\right)-\frac{1}{4}\left(\frac{1}{a+1}\right)^4\left((bti)^4+\cdots\right)\right]$$

$$+a\left[(bti)+\frac{1}{2}(bti)^2+\frac{1}{2}(bti)^3+\frac{5}{8}(bti)^4\right]$$

$$\approx1+\left[\left(\frac{1}{a+1}\right)+a\right](bti)+\left[\frac{3}{2}\left(\frac{1}{a+1}\right)-\frac{1}{2}\left(\frac{1}{a+1}\right)^2+\frac{1}{2}a\right](bti)^2$$

Suan Sunandha Science and Technology Journal

$$+\left[\frac{5}{2}\left(\frac{1}{a+1}\right)-\frac{3}{2}\left(\frac{1}{a+1}\right)^2+\frac{1}{3}\left(\frac{1}{a+1}\right)^3+\frac{1}{2}a\right](bti)^3$$
$$+\left[\frac{35}{8}\left(\frac{1}{a+1}\right)-\frac{29}{8}\left(\frac{1}{a+1}\right)^2+\frac{3}{2}\left(\frac{1}{a+1}\right)^3-\frac{1}{4}\left(\frac{1}{a+1}\right)^4+\frac{5}{8}a\right](bti)^4$$
$$\approx\left[\left(\frac{1}{a+1}\right)+a\right]\frac{(bti)}{1!}+\left[3\left(\frac{1}{a+1}\right)-\left(\frac{1}{a+1}\right)^2+a\right]\frac{(bti)^2}{2!}$$
$$+\left[15\left(\frac{1}{a+1}\right)-9\left(\frac{1}{a+1}\right)^2+2\left(\frac{1}{a+1}\right)^3+3a\right]\frac{(bti)^3}{3!}$$
$$+\left[105\left(\frac{1}{a+1}\right)-87\left(\frac{1}{a+1}\right)^2+36\left(\frac{1}{a+1}\right)^3-6\left(\frac{1}{a+1}\right)^4+15a\right]\frac{(bti)^4}{4!}.$$

From $K_X(t)=\log\varphi_X(t)=\sum_{n=1}^{m}\frac{\kappa_n}{n!}(ti)^n+O(t^{m+1})$,

the first four cumulants are given as follows:

$$\kappa_1=\left(\frac{1}{a+1}+a\right)b,$$
$$\kappa_2=\left(\frac{3}{a+1}-\frac{1}{(a+1)^2}+a\right)b^2,$$
$$\kappa_3=\left(\frac{15}{a+1}-\frac{9}{(a+1)^2}+\frac{2}{(a+1)^3}+3a\right)b^3,$$
$$\kappa_4=\left(\frac{105}{a+1}-\frac{87}{(a+1)^2}+\frac{36}{(a+1)^3}-\frac{6}{(a+1)^4}+15a\right)b^4,\quad \text{and}$$

the raw moments are related to the cumulants by the following formula:

$$E(X)=\kappa_1,$$
$$E(X^2)=\kappa_2+\kappa_1^2,$$
$$E(X^3)=\kappa_3+3\kappa_2\kappa_1+\kappa_1^3,$$
$$E(X^4)=\kappa_4+4\kappa_3\kappa_1+3\kappa_2^2+6\kappa_2\kappa_1^2+\kappa_1^4.$$

Substituting the cumulants in the equation above, we can find mean and variance are given by

$$E(X)=\left(\frac{1}{a+1}+a\right)b,$$

$$Var(X)=\kappa_2=\left(\frac{3}{a+1}-\frac{1}{(a+1)^2}+a\right)b^2.$$

Some mean and variance plots of the EMIG distribution with some parameter values are displayed in Figure 5.

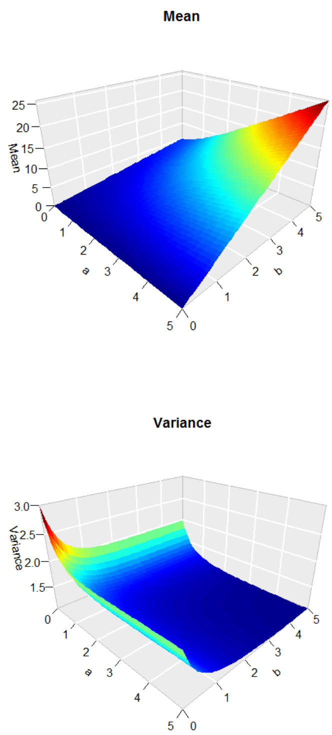

Figure 5. Mean and Variance of the EMIG distribution for difference values of (a,b).

3.1.2 Parameters estimation

The estimation of parameters for the extended mixture inverse Gaussian distribution via the Maximum Likelihood Estimation (MLE) procedure.The likelihood function of the distribution with parameters a and b is given by

$$L(a,b)=\prod_{i=1}^{n}\left\{\frac{1}{(a+1)b\sqrt{2\pi}}\left[a^2\left(\frac{b}{x_i}\right)^{3/2}+\left(\frac{b}{x_i}\right)^{1/2}\right]\times\exp\left[-\frac{1}{2}\left(\sqrt{\frac{x_i}{b}}-a\sqrt{\frac{b}{x_i}}\right)^2\right]\right\}$$

The log-likelihood function can be written as

$$\ell(a,b)=\log L(a,b)$$
$$=-n\log(ab+b)-\frac{n}{2}\log(2\pi)+na$$
$$-\frac{1}{2b}\sum_{i=1}^{n}x_i-\frac{a^2b}{2}\sum_{i=1}^{n}\frac{1}{x_i}$$
$$+\sum_{i=1}^{n}\log\left[a^2\left(\frac{b}{x_i}\right)^{3/2}+\left(\frac{b}{x_i}\right)^{1/2}\right]$$

Suan Sunandha Science and Technology Journal

$$= -n\log\left(ab+b\right) - \frac{n}{2}\log\left(2\pi\right) + na$$
$$-\frac{1}{2b}\sum_{i=1}^{n} x_i - \frac{a^2 b}{2}\sum_{i=1}^{n}\frac{1}{x_i}$$
$$+\sum_{i=1}^{n}\log\left[b^{1/2}x_i^{-3/2}\left(a^2 b + x_i\right)\right]$$
$$= -n\log\left(ab+b\right) - \frac{n}{2}\log\left(2\pi\right) + na$$
$$-\frac{1}{2b}\sum_{i=1}^{n} x_i - \frac{a^2 b}{2}\sum_{i=1}^{n}\frac{1}{x_i} + \frac{n}{2}\log b - \frac{3}{2}\sum_{i=1}^{n}\log x_i$$
$$+\sum_{i=1}^{n}\log\left(a^2 b + x_i\right).$$

By taking first partial derivatives of the log-likelihood function each with respect to a and b, we obtain the equations;

$$\frac{\partial}{\partial a}\ell\left(a,b\right) = n - \frac{nb}{ab+b} - ab\sum_{i=1}^{n}\frac{1}{x_i} + \sum_{i=1}^{n}\frac{2ab}{a^2 b + x_i},$$

$$\frac{\partial}{\partial b}\ell\left(a,b\right) = \frac{-n\left(a+1\right)}{\left(ab+b\right)} + \frac{n}{2b} + \frac{1}{2b^2}\sum_{i=1}^{n} x_i$$
$$-\frac{a^2}{2}\sum_{i=1}^{n}\frac{1}{x_i} + \sum_{i=1}^{n}\frac{a^2}{a^2 b + x_i}.$$

The MLE solutions of $\hat{a}, \hat{b}$ can be obtained by equating the above equations to zero and solving the resulting equations simultaneously using a numerical procedure with the Newton-Raphson method in R (R Development Core Team, 2015).

3.2 Numerical Result

In this section, the EMIG distribution is applied on real data set which is taken from Gupta and Kundu (2009) which represent final examination marks in mathematics of the slow pace students in 2003, The data are given in Table 1.

Table 1. The final examination marks in mathematics of the slow pace students in 2003.

29	25	50	15	13	27	15	18
7	7	8	19	12	18	5	21
15	86	21	15	86	21	15	14
70	44	6	23	58	19	50	23
11	6	34	18	28	34	12	37
4	60	20	23	40	65	19	31

For this data, we have that mean=25.89 and variance= 345. We have fitted the EMIG, MIG, IG, LBIG and BS distributions to this data set by using maximum likelihood estimation. We obtain the estimates parameters and AIC statistics for all distributions are shown in Table 2. Next, we compute Anderson-Daring (AD) test for goodness of fit distribution to these data which are shown in Table 3.

Table 2. MLE of the model parameters for the marks data set.

Distribution	**Estimate parameters**			**AIC**
	a	b	p	
EMIG	1.73	12.34	-	388.74
MIG	5.61	12.03	0.74	918.49
IG	5.60	12.08	-	916.50
LBIG	1.54	10.20	-	396.31
BS	5.84	10.67	-	916.11

Distribution	**Estimate parameters**			**AIC**
	a	b	p	
EMIG	1.73	12.34	-	388.74
MIG	5.61	12.03	0.74	918.49
IG	5.60	12.08	-	916.50
LBIG	1.54	10.20		396.31
BS	5.84	10.67	-	916.11

Table 3. Goodness of fit test for the marks data set by using Anderson-Darling test

Distribution	**AD test**	
	Statistic	**P-value**
EMIG	0.3215	0.9208
MIG	0.1633	0.0107
IG	0.3529	0.8935
LBIG	0.3154	0.9109
BS	0.3279	0.9156

The P-value of Anderson-Darling test is shown that the EMIG distribution performs better than MIG, IG, LBIG and BS distributions and the value of AIC statistic is shown that the EMIG distribution is the best fit for this data. The fitted probability density functions and the observed histograms are given in Figure 6.

Suan Sunandha Science and Technology Journal

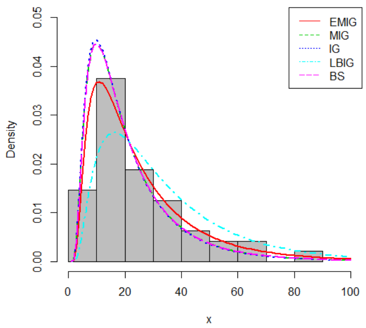

Figure 6. Histogram of the marks data.

4. Conclusion

In this paper, we have presented the EMIG distribution which is given by adding a new weight parameter to the MIG distribution. The BS and LBIG distribution are some special case of EMIG. We mainly studied the statistical properties of EMIG distribution such as pdf, density function, survival function, hazard rate function, cumulants, the first four moments, mean and variance. We discuss the estimation of the parameters by maximum likelihood. Finally, we compare the fit of the EMIG distribution with MIG, IG, LBIG and the BS distributions by using marks data. The AIC statistics indicates that the EMIG is best fit for real data.

Acknowledgement

The author would like to thank the editor and referee for their useful comments and suggestions which considerably improved this article.

References

Akman, O., & Gupta, R.C. (1992). A comparison of various estimators of an inverse Gaussian distribution. *Journal of Statistical Computation and Simulation*,*40*,1-2,71-81. doi:10.1080/0094 9659208811366.

Balakrishnan, N., Leiva, V., Sanhueza, A. & E. Cabrera, (2009). Mixture inverse Gaussian distributions and its transformations, moments and applications. *Statistics*, *43*(1), 91-104, doi: 10.1080/02331880701829948.

Gradshteyn, I.S., & Ryzhik, I.M. (2007). *Table of integrals, series, and products.* (7th ed.). MA, USA: Elsevier Academic Press.

Gupta, R.C., & Akman, H.O. (1995). On the reliability studies of a weighted inverse Gaussian model. *Journal of Statistical Planning and Inference*, *48* (1), 69-83, doi:10. 1016/0378-3758(94)00148-O.

Gupta, R.D., & Kundu, D. (2009). A new class of weighted exponential distributions, *Statistics: A Journal of Theoretical and Applied Statistics*, *43*(6), 621-634, doi:10.1080/023318 80802605346.

Gupta, R.C., & Kundu, D. (2011). Weighted inverse Gaussian a versatile life time model. *Journal of Applied Statistics*, *38*(12), 2695-2708, doi:10.1 080/02664763.2011.567251.

Henze, N., & Klar, B. (2002). Goodness-of-fit tests for the inverse Gaussian distribution based on the empirical laplace transform. *Annals of the Institute of Statistical Mathematics*, *54*(2), 425-444, doi:10.1023/A:1022442506681.

Jorgensen B., Seshadri, V., & Whitmore, G.A. (1991). On the mixture of the inverse Gaussian distribution with its complementary reciprocal. *Scandinavian Journal of Statistics*, *18*(1), 77-89.

R Development Core Team, (2015). *A language and environment for statistical computing. R foundation for statistical computing*, Vienna, Austria.

Von Alven, W.H., (1964). *Reliability Engineering*. New Jersey: prentice-hall Inc.