SCIREA Journal of Mathematics
ISSN: 2995-5823
http://www.scirea.org/journal/Mathematics
May 7, 2024
Volume 9, Issue 3, June 2024
https://doi.org/10.54647/mathematics110484

Determine Sample Size to Estimate the Average Parameter of a Heavy Tails Distribution Using Bayesian Methodology

Sarmad Abdulkhaleq Salih[1], Omar Ramzi Jasim [1]

[1]University of Al-Hamdaniya , Mosul, Iraq.

Email: sarmadsalih@uohamdaniya.edu.iq (Sarmad Abdulkhaleq Salih) ,
omar.ramzi89@uohamdaniya.edu.iq (Omar Ramzi Jasim)

ABSTRACT:

The generalized modified Bessel distribution is one of the most suitable mixed distributions. It is the result of mixing the normal distribution with the generalized inverse Gaussian distribution.

In this paper, The optimal sample size Analysis has been taken from the generalized modified Bessel population to estimate the mean parameter when the variance and shape parameters are known, using the informative prior information to estimate the mean parameter under the quadratic loss function. Then sampling and non-sampling approaches are used for the estimate of the parameter. Also, it has been noted that the posterior probability distribution for a mean parameter is following a generalized modified Bessel distribution. Through the simulation, we note Bayesian sample size is inversely proportional to the sampling cost (c) per unit.

Keywords: Generalized Modified Bessel Distribution, Quadratic Loss Function, Cost Function, Bayesian Sample Size.

1. INTRODUCTION

The generalized modified Bessel distribution is considered one of the most suitable mixed distributions, as this distribution is more general than the two distributions (normal and T) as special cases, and it is considered that the generalized modified Bessel distribution as a special case of the symmetrical hyperbolic distributions, and this distribution has practical applications in a variety of areas which includes stock market data presentation, quality control data and filtering random-sign analysis.

To conduct any study on data that follows this distribution, its parameters must be estimated to be able to predict and obtain accurate results to solve a particular problem. The characteristics of the community are determined by studying a sample drawn from it, provided that the sample bears the maximum degree of accuracy to represent the community, and this leads to reducing the costs of the field study due to the small size of the sample about the size of the community. There are several sampling methods to determine the sample size to be used in statistical inference, and one of these methods is using the Bayesian sampling method, which in its study depends on taking into account the loss functions with the cost function.

(Thabane and Haq, 2003) was the first to use the Bayesian method to estimate the linear regression model, assuming that the random error distribution is a generalized multivariate modified Bessel distribution when the prior distribution of parameters is a proper distribution, which is the result of mixing two continuous probability distributions. They are the normal distribution and the generalized inverse Gaussian distribution.[14]

(Saiful Islam, 2011)Determine the Bayesian sample size to estimate the mean and the difference between the mean of the normal distribution when the variance is known and estimate the variance when the mean is known, as well as determine the Bayesian sample size to estimate the parameter of the Poisson and Exponential distribution using different loss functions. [11]

The research provided a general introduction to the subject and a description of the generalized modified Bessel distribution, as well as the Bayesian sample size determination to estimate the mean parameter when the prior distribution is proper and using a quadratic loss function.

2. Search Objectives

This paper aims to analyze the optimal Bayesian sample size taken from a population followed by is generalized modified Bessel distribution to estimate the mean parameter using the quadratic loss function.

3. Generalized Modified Bessel distribution

The generalized modified Bessel distribution is expressed in form of a mixed distribution from the normal distribution and the generalized inverse Gaussian distribution as follows: [6][13] [14]

$$f(Y) = \frac{\left(\frac{\lambda}{\psi}\right)^{\frac{1}{4}}}{(2\pi\ \sigma^2)^{\frac{1}{2}} K_v\left(\sqrt{\lambda\psi}\right)} \left[1 + \frac{(Y-\theta)^2}{\psi\sigma^2}\right]^{\frac{2v-1}{4}}$$

$$\times K_{\frac{2v-1}{2}}\left(\sqrt{\lambda\psi\left(1 + \frac{(Y-\theta)^2}{\psi\sigma^2}\right)}\right) \tag{1}$$

Whereas, $K_v(.)$ represents a modified Bessel function of the third type with order (v) [1][9]. This distribution is descriptively expressed as:

$$Y \sim GMBD_1(\theta\ , \sigma^2, \lambda\ , \psi\ , v) \tag{2}$$

4. Quadratic Loss Function

The quadratic loss function is the most commonly used in symmetric loss functions, when the parameter (θ) is univariate, the general formula of the function is as follows:[4][11]

$$L(\hat{\theta}\ , \theta) = k(\hat{\theta} - \theta)^2 \tag{3}$$

Since (k) positive real constant is often equal to an integer one.

Therefore, the Bayesian estimator in the case that information is available for this function represents the average of the posterior probability distribution and the Bayesian estimator in the absence of information represents the average of the prior distribution. [10]

$$\hat{\theta}_{sq} = \hat{\theta}_{Bayes} = E(\theta|\underline{Y}) \tag{4}$$

5. Posterior Risk Function of The Quadratic Loss Function

If $L(\hat{\theta},\theta)$ is a loss function and $P(\theta|\underline{Y})$ is the posterior distribution of (θ), then the expectation of the loss function is called the posterior risk function as follows: [4]

$$PR = \int_{\forall} L(\hat{\theta},\theta) * P(\theta|\underline{Y})\, d\theta \tag{5}$$

(PR) is the posterior risk. The posterior risk function will be calculated for the quadratic loss function as follows:

$$PR_S = \int_{\forall\theta} k(\hat{\theta}_{sq} - \theta)^2 * P(\theta|\underline{Y})\, d\theta = k\, Var\,(\theta|\underline{Y}) \tag{6}$$

This constant (k) does not affect calculating the Bayesian estimator, but it affects determining the Bayesian sample size.

6. The Cost Function

The cost function is one of the important issues in any study that aim to make it as minimal as possible, and there are different formulas for the cost function, and the linear cost function has been used in this research and it's as follows:

$$C_{(n)} = c_0 + cn \tag{7}$$

Where $n > 0$ and $C_{(0)} = 0$

And (c_0) is the cost of preparing the sampling or any cost related to sampling and (c) is the sampling cost per unit and it is one of the important components for obtaining the Bayesian sample size. By using the Bayes method to estimate the sample size, the total cost $(TC_{(n)})$ must be calculated, which can be done by the following formula when the posterior risk function does not depend on the observations Y.[8]

$$TC_{(n)} = C_{(n)} + PR \tag{8}$$

But if the posterior risk function depends on (Y) observations, in this case, we take the expectation of the posterior risk function and add to it the linear cost function to get the average total cost as follows:

$$E(TC_{(n)}/n) = C_{(n)} + E(PR) \tag{9}$$

When two samples are drawn from both groups, the linear cost function will be in the following form:

$$C_{(n_1,n_2)} = c_0 + c_1 n_1 + c_2 n_2 \tag{10}$$

As (c_1) is the sampling cost per unit of (n_1), which represents the sample size for the first population and (c_2) is the sampling cost per unit of (n_2), which represents the sample size for the second population. The costs $(c_1 \& \, c_2)$ can be equal to the two populations, or the sample sizes $(n_1 \& \, n_2)$ can be equal to the two populations. To get the sample size that gives the lowest cost, one must derive the total cost function relative to (*n*) and by equating the derivative to zero, to get the Bayesian sample size, which is a positive integer. [4]

7. Determining the Bayesian sample size to estimate the mean parameter of the generalized modified Bessel distribution for one population using initial information belonging to the conjugate family

Assuming to have a random sample $(Y_1, Y_2, \dots, Y_n)$ conditioned by the random variable (τ) taken from normal population $N(\theta, \sigma^2 \tau)$. And that the probability distribution of (τ) follows the generalized inverse Gaussian distribution, the probability function is defined as follows: [3][7] [9]

$$P(\tau) = \frac{\left(\frac{\lambda}{\psi}\right)^{\frac{v}{2}}}{2K_v(\sqrt{\lambda\psi})} \; \tau^{v-1} \exp\left[-\frac{1}{2}\left\{\left(\frac{\psi}{\tau}\right) + \lambda\tau\right\}\right] \quad , \tau > 0 \tag{11}$$

Whereas:

(λ, ψ): The scale parameters.

v: The shape parameter.

The Bayesian sample size is found to estimate the mean parameter conditioned by the random variable (τ) by calculating the total cost function as follows:

$$TC_{(n)} = C_{(n)} + PR$$

By substituting the linear cost function $C_{(n)}$ defined in equation (7) and the posterior risk function (PR) (expected the loss function), the total cost function becomes as follows:

$$TC_{(n)} = c_0 + cn + E_\theta\left(L(\hat{\theta}, \theta)\right)$$

Substituting for the quadratic loss function defined in equation (3), to get:

$$TC_{(n)} = c_0 + cn + E_\theta\left(k\left(\hat{\theta}_{sq} - \theta\right)^2\right)$$

And by substituting for the subsequent risk function (PR) shown in equation (6), the total cost function becomes as follows:

$$TC_{(n)} = c_0 + cn + k\, Var\left(\theta | \underline{Y}\right) \tag{12}$$

To find $Var\left(\theta|\underline{Y}\right)$ the posterior probability distribution of the mean parameter conditional of the random variable (τ) must be calculated as follows:

The prior distribution of the mean parameter (θ) conditioned by (τ) and (σ^2) is described as follows: [2]

$$(\theta|\sigma^2, \tau) \sim N(\theta_0, \sigma_0^2\, \sigma^2\, \tau)$$

Since the prior probability density function for $(\theta|\sigma^2, \tau)$ is as follows:

$$P\left(\theta|\sigma^2, \tau\right) \propto e^{-\frac{(\theta-\theta_0)^2}{2\sigma_0^2\,\sigma^2\,\tau}} \tag{13}$$

After combining the sample information with the initial information about the parameter (θ), the posterior distribution of (θ) conditional on the variable (τ) and (σ^2) is found using the following equation:

$$P\left(\theta|\underline{Y}, \sigma^2, \tau\right) \propto P\left(\theta|\sigma^2, \tau\right)\; f\left(\underline{Y}|\theta, \sigma^2, \tau\right) \tag{14}$$

After substituting the probability function for the random vector $(\underline{Y})$ conditional by (τ), defined according to the following equation:

$$f\left(\underline{Y}|\theta, \sigma^2, \tau\right) = (2\pi\sigma^2\tau)^{-\frac{n}{2}}\, e^{-\frac{1}{2\sigma^2\tau}\sum(Y_i-\theta)^2} \tag{15}$$

With the prior distribution of the parameter (θ) conditional by the variables (τ) and (σ^2) in the equation (14) above to get the kernel of the posterior distribution of (θ) conditional by (τ) and (σ^2) $\left(P\left(\theta|\underline{Y}, \sigma^2, \tau\right)\right)$ which is defined below:

$$P\left(\theta|\underline{Y}, \sigma^2, \tau\right) \propto e^{-\frac{1}{2\sigma^2\tau}\left[n(\theta-\overline{Y})^2+\frac{(\theta-\theta_0)^2}{\sigma_0{}^2}\right]} \quad (16)$$

The quadratic form in equation (16) above is similar to the following quadratic form: [2]

$$A(z-a)^2 + B(z-b)^2 = (A+B)(Z-m)^2 + \frac{AB}{A+B}(a-b)^2 \quad (17)$$

Whereas

$$m = \frac{aA + bB}{A+B} \quad (18)$$

Therefore, the posterior probability distribution for the mean parameter conditioned by (τ) and (σ^2) is normal and is as follows:

$$P\left(\theta|\underline{Y}, \sigma^2, \tau\right) = \frac{\left(\frac{1}{\sigma_0{}^2}+n\right)^{0.5}}{\sqrt{2\pi\sigma^2\tau}}\ e^{-\frac{\frac{1}{\sigma_0{}^2}+n}{2\sigma^2\tau}[\theta-m]^2} \quad (19)$$

Since:

$$m = \frac{\theta_0 + n\overline{Y}\,\sigma_0{}^2}{1 + n\,\sigma_0{}^2} \quad (20)$$

The distribution is descriptively described by:

$$\left(\theta|\underline{Y}, \sigma^2, \tau\right) \sim N\left(m, \frac{\sigma^2\tau}{\frac{1}{\sigma_0{}^2}+n}\right)$$

Using the concept of mixed distributions, the posterior probability distribution for the mean parameter unconditional of (τ) is as follows:

$$P\left(\theta|\underline{Y}, \sigma^2\right) = \int_0^\infty P\left(\theta|\underline{Y}, \sigma^2, \tau\right)\ \ P(\tau)\ \ d\tau \quad (21)$$

After substituting the components of equation (21) and performing the integration concerning (τ), yielding the following equation:

$$P\left(\theta|\underline{Y},\sigma^2\right)=\frac{\left(\frac{1}{\sigma_0{}^2}+n\right)^{0.5}\left(\frac{\lambda}{\psi}\right)^{0.25}}{\sqrt{2\pi\sigma^2}\;K_v\left(\sqrt{\lambda\psi}\right)}\left[1+\frac{\left(\frac{1}{\sigma_0{}^2}+n\right)(\theta-m)^2}{\psi\sigma^2}\right]^{\frac{2v-1}{4}}$$

$$\times K_{\frac{2v-1}{2}}\left(\sqrt{\lambda\psi\left(1+\frac{\left(\frac{1}{\sigma_0{}^2}+n\right)(\theta-m)^2}{\psi\sigma^2}\right)}\right) \qquad (22)$$

The probability density function of $\left(\theta|\underline{Y},\sigma^2\right)$ defined in equation (22) represents the generalized modified Bessel distribution and is descriptively expressed as:

$$\left(\theta|\underline{Y},\sigma^2\right)\sim GMBD\left(m,\frac{\sigma^2}{\frac{1}{\sigma_0{}^2}+n},\lambda,\psi,v\right)$$

Since

$$m=\frac{\theta_0+n\overline{Y}\,\sigma_0{}^2}{1+n\,\sigma_0{}^2} \qquad (23)$$

The unconditional Bayes estimator for the parameter (θ) is:

$$E\left(\theta|\underline{Y},\sigma^2\right)=E_\tau E\left(\theta|\underline{Y},\sigma^2,\tau\right)$$

$$=m \qquad (24)$$

Since (m) defined in equation (23) represents a Bayes estimator under the quadratic loss function $\widehat{\theta}_{sq}$.

$$\hat{\theta}_{sq}=\frac{\theta_0+n\overline{Y}\,\sigma_0{}^2}{1+n\,\sigma_0{}^2} \qquad (25)$$

Using the properties of mathematical expectation, it is possible to find the variance for the posterior distribution of $\left(\theta|\underline{Y},\sigma^2\right)$ as follows: [5]

$$Var\left(\theta|\underline{Y},\sigma^2\right)=E\left[Var\left(\theta|\underline{Y},\sigma^2,\tau\right)\right]+Var\left[E\left(\theta|\underline{Y},\sigma^2,\tau\right)\right] \qquad (26)$$

$$= \frac{\sigma^2}{\frac{1}{\sigma_0{}^2} + n} \, E(\tau)$$

$$= \frac{\sigma^2}{\frac{1}{\sigma_0{}^2} + n} \, \frac{K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5} \tag{27}$$

By substituting equation (27) above into equation (12), the total cost function is as follows:

$$TC_{(n)} = c_0 + cn + k \;\; \frac{\sigma_0{}^2\sigma^2}{1 + n\sigma_0{}^2} \, \frac{K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5} \tag{28}$$

Now by finding the derivative of equation (28) relative to (n) and equal the derivative to zero, we get the Bayesian sample size under the quadratic loss function:

$$\frac{\partial TC_{(n)}}{\partial n} = c - \;\; \frac{k\,\sigma_0{}^4\sigma^2}{(1 + n\sigma_0{}^2)^2} \, \frac{K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}$$

$$\frac{\partial TC_{(n)}}{\partial n} = 0$$

$$c = \;\; \frac{k\,\sigma_0{}^4\sigma^2}{\left(1 + n^*{}_{sq}\sigma_0{}^2\right)^2} \, \frac{K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}$$

$$1 + n^*{}_{sq}\sigma_0{}^2 = \sigma_0{}^2 \sqrt{\frac{k\,\sigma^2 \;\; K_{v+1}\left(\sqrt{\lambda\psi}\right)}{c \;\; K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}}$$

Where

$$n^*{}_{sq} = \sqrt{\frac{k\;\sigma^2 \;\; K_{v+1}\left(\sqrt{\lambda\psi}\right)}{c \;\; K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \frac{1}{\sigma_0{}^2} \tag{29}$$

Since $(n^*{}_{sq})$ represents the Bayesian sample size under the quadratic loss function.

8. The sampling and non-sampling status of estimating the mean parameter

In the case that is not sample taken ($n^* = 0$), the total cost under the quadratic loss function shown in equation (28) is: [2]

$$TC_{(0)} = k \ \frac{\sigma_0{}^2 \sigma^2 \ \ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5} \tag{30}$$

When taking a sample ($n^* = n^*{}_{sq}$), the total cost will be:

$$TC_{(n)} = c_0 + c\, n^*{}_{sq} + k \ \frac{\sigma_0{}^2 \sigma^2}{1 + n^*{}_{sq}\ \sigma_0{}^2} \frac{K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}$$

Substituting for $(n^*{}_{sq})$ previously defined in equation (29), we get:

$$TC_{(n)} = c_0 + c\left[\sqrt{\frac{k\ \sigma^2 \ \ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{c \ \ K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \frac{1}{\sigma_0{}^2}\right]$$

$$+k \ \frac{\sigma_0{}^2 \sigma^2}{1 + \left[\sqrt{\frac{k\ \sigma^2 \ \ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{c \ \ K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \frac{1}{\sigma_0{}^2}\right] \sigma_0{}^2} \ \frac{K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}$$

Performing some algebraic operations for the above equation, we get:

$$TC_{(n)} = c_0 - \frac{c}{\sigma_0{}^2} + 2\sqrt{\frac{c\ k\ \sigma^2 \ \ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}} \tag{31}$$

Therefore, the optimal Bayes sample size is generally:

$$n^* = max\left[0\ ,\ \sqrt{\frac{k\ \sigma^2 \ \ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{c \ \ K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \frac{1}{\sigma_0{}^2}\right] \tag{32}$$

Therefore, it is decided to take a sample in the case of sampling when the $\left[\sqrt{\frac{k\ \sigma^2 \ \ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{c \ \ K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \frac{1}{\sigma_0{}^2}\right]$ is positive, and the total cost in the case of sampling is less than the total cost in the case of non-sampling.

That is:

$$k\frac{{\sigma_0}^2\sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5} > c_0-\frac{c}{{\sigma_0}^2}+2\sqrt{\frac{c\ k\ \sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}}\ \&\ \sqrt{\frac{k\ \sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{c\ K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}} > \frac{1}{{\sigma_0}^2}$$

And it is decided not to take a sample in the case of sampling when the $\left[\sqrt{\frac{k\ \sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{c\ K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}}-\frac{1}{{\sigma_0}^2}\right]$ is negative, and the total cost in the case of sampling is greater than the total cost in the case of non-sampling.

That is:

$$k\frac{{\sigma_0}^2\sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5} < c_0-\frac{c}{{\sigma_0}^2}+2\sqrt{\frac{c\ k\ \sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}}\ \&\ \sqrt{\frac{k\ \sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{c\ K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}} > \frac{1}{{\sigma_0}^2}$$

By solving the inequality and substituting for each $\sqrt{c}$ of (ϑ) in the sampling case, we get the following formula:

$$k\frac{{\sigma_0}^2\sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5} > c_0-\frac{c}{{\sigma_0}^2}+2\sqrt{\frac{c\ k\ \sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}}\ \&\ \sqrt{\frac{k\ \sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{c\ K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}} > \frac{1}{{\sigma_0}^2}$$

$$\frac{\vartheta^2}{{\sigma_0}^2}-2\vartheta\sqrt{\frac{k\ \sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}}+k\frac{{\sigma_0}^2\sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}-c_0>0$$

$$\&\ \vartheta<\sqrt{\frac{k\ \sigma^2{\sigma_0}^4\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}}$$

$$\frac{\vartheta^2}{{\sigma_0}^2}-2\vartheta\sqrt{\frac{k\ \sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}}+k\frac{{\sigma_0}^2\sigma^2\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}>c_0$$

$$\&\ \vartheta<\sqrt{\frac{k\ \sigma^2{\sigma_0}^4\ K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)}\left(\frac{\lambda}{\psi}\right)^{-0.5}}$$

Completing the square for the left side of the inequality, to get:

$$\left(\frac{\vartheta}{\sqrt{{\sigma_0}^2}} - \sqrt{k\frac{{\sigma_0}^2\sigma^2 \quad K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}}\right)^2 > c_0$$

Taking the square root of both sides of the inequality, to get:

$$\mp\left(\frac{\vartheta}{\sqrt{{\sigma_0}^2}} - \sqrt{k\frac{{\sigma_0}^2\sigma^2 \quad K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}}\right) > \sqrt{c_0}$$

The roots of the above inequality are as follows:

$$\vartheta > \sqrt{k\frac{{\sigma_0}^4\sigma^2 \quad K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}} + \sqrt{c_0{\sigma_0}^2}$$

Or:

$$\vartheta < \sqrt{k\frac{{\sigma_0}^4\sigma^2 \quad K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \sqrt{c_0{\sigma_0}^2}$$

But work interest is in the lowest cost, so taking the second root to show when the sampling takes place or not, as follows:

$$\sqrt{c} < \sqrt{k\frac{{\sigma_0}^4\sigma^2 \quad K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \sqrt{c_0{\sigma_0}^2}$$

So it is sampling when:

$$c < \left[\sqrt{k\frac{{\sigma_0}^4\sigma^2 \quad K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \sqrt{c_0{\sigma_0}^2}\right]^2$$

And non-sampling when:

$$c > \left[\sqrt{k\frac{{\sigma_0}^4\sigma^2 \quad K_{v+1}\left(\sqrt{\lambda\psi}\right)}{K_v\left(\sqrt{\lambda\psi}\right)} \left(\frac{\lambda}{\psi}\right)^{-0.5}} - \sqrt{c_0{\sigma_0}^2}\right]^2$$

If the sampling cost per unit exceeds the $\left[\sqrt{k\,\frac{\sigma_0{}^4\sigma^2}{K_v(\sqrt{\lambda\psi})}\frac{K_{v+1}(\sqrt{\lambda\psi})}{}\left(\frac{\lambda}{\psi}\right)^{-0.5}} - \sqrt{c_0\sigma_0{}^2}\right]$, the sampling is not of value due to the high cost of sampling per unit, and in this case, only the prior information is relied on to estimate the mean parameter for the generalized modified Bessel distribution.

9. Simulation

In this section, random data was generated that follow the generalized modified Bessel distribution through the concept of mixed distribution in the programming language (Matlab R2015a) [12], After generating the observations, it was tested whether the data followed the generalized modified Bessel distribution by Kolmogorov-Smirnov test if it was found that the (p-value) is greater than the level of significance ($\propto = 0.05$) at different values of the shape parameters, and therefore the hypothesis that the data follows the generalized modified Bessel distribution is accepted. And by executing the program, the estimator of the mean parameter is calculated under the quadratic loss function and defined in equation (25) at different initial values and the comparison through (MSE):

$$MSE(\hat{\theta}_{sq}) = \frac{\sum_{i=1}^{R}(\hat{\theta}_{sq} - \theta)^2}{R} \tag{33}$$

Where is the number of iterations $(R = 100)$.

This comparison will be done in two cases. In the first case, the optimal Bayesian sample size was used to calculate the mean parameter defined in the equation (29) and in the second case, different sample sizes were used ($n = 25,75,125$). Table (1) shows the assumed values of generation and calculates estimators:

Table 1. Assumed values for generation and for calculating estimators

	Shape parameters		θ	σ^2	θ_0	${\sigma^2}_0$	c	k
$\lambda = 2$	$\psi = 1.5$	$v = 3$	33.5	0.5	3	1.5	0.01	**2**
$\lambda = 8$	$\psi = 3$	$v = 10$		4	7		0.005	**5**

Table (2) and (3) shows the (MSE) values for an estimate of the mean parameter at different shape parameters and different sample sizes when $(\sigma^2 = 0.5, 4)$.

Table 2. MSE values for different shape parameters (λ, ψ, v) and different sample sizes at $(\sigma^2 = 0.5)$

(λ, ψ, v)	n	θ_0	$\hat{\theta}_{sq}$	MSE(n)
(2,1.5,3)	25	3	3.3001	0.0502
		7	3.0054	0.0681
	75	3	3.3615	0.0453
		7	3.1242	0.0601
	125	3	3.5123	0.0398
		7	3.4458	0.0495
(8,3,10)	25	3	3.3102	0.0428
		7	3.1254	0.0599
	75	3	3.3984	0.0414
		7	3.2451	0.0576
	125	3	3.5012	0.0202
		7	3.4746	0.0399

Table 3. MSE values for different shape parameters (λ, ψ, v) and different sample sizes at $(\sigma^2 = 4)$

(λ, ψ, v)	N	θ_0	$\hat{\theta}_{sq}$	MSE(n)
(2,1.5,3)	25	3	3.9101	0.0880
		7	4.0604	0.0989
	75	3	3.1923	0.0778
		7	3.9781	0.0928
	125	3	3.7825	0.0682
		7	3.1632	0.0789
(8,3,10)	25	3	3.2524	0.0732
		7	3.8942	0.0900

75	3	3.2998	0.0712
	7	3.9321	0.0892
125	3	3.6123	0.0532
	7	3.3549	0.0692

We notice from tables (2) & (3), that the (MSE) value decreases with the increase in the shape parameters and sample size at $(\sigma^2 = 0.5)$ and when the value of (θ_0) is close to (θ).

Table (4) represents the Bayesian sample size values for different shape parameters (λ, ψ, υ) when $(\sigma^2 = 0.5, \theta_0 = 3)$.

Table 4. Bayesian sample size values for different shape parameters (λ, ψ, υ)

$(\lambda, \psi, \upsilon)$	k	c	n^*_{sq}	$\hat{\theta}_{sq}$	MSE(n^*_{sq})
(2,1.5,3)	2	0.01	18	2.9171	0.2345
		0.005	23	3.0014	0.1982
	5	0.01	36	3.2045	0.1156
		0.005	41	3.2204	0.0968
(8,3,10)	2	0.01	27	3.1121	0.1756
		0.005	31	3.1984	0.1536
	5	0.01	55	3.7956	0.0864
		0.005	72	3.6548	0.0645

We notice from table (4) that the Bayesian sample size increases with the increase in the shape parameters and the (MSE) value decreases with the increase in the shape parameters, as well as the Bayesian sample size is inversely proportional to the sampling cost (c) per unit.

10. Conclusions

* The posterior probability distribution for the mean parameter follows the generalized modified Bessel distribution.

* The Bayesian sample size was determined to estimate the mean parameter of the generalized modified Bessel distribution when the parameters ($\lambda, \psi, v, \sigma^2$) are known, and the sampling decision is made under quadratic loss function when the sampling cost per unit (c) is less than the

$$\left[\sqrt{k\,\frac{\sigma_0{}^4\sigma^2\;\;K_{v+1}(\sqrt{\lambda\psi})}{K_v(\sqrt{\lambda\psi})}\left(\frac{\lambda}{\psi}\right)^{-0.5}}-\sqrt{c_0\sigma_0{}^2}\right]^2.$$

* Through simulation, we notice that when the values of the shape parameters are increased, the values of the estimators of the mean parameter using the Bayesian sample size are close to the values of the estimators of the mean parameter using a relatively large sample size.

11. Recommendations

* Using a non-linear cost function to determine the Bayesian sample size to estimate the mean parameter of the generalized modified Bessel distribution under the quadratic loss function.

* Traditional sample size analysis for estimating parameters of the generalized modified Bessel distribution.

* Using other distributions with heavy tails to determine the Bayesian sample size, including the double exponential distribution (Laplace distribution).

* Using other symmetric and asymmetric loss functions to determine the Bayesian sample size to estimate the mean parameter of the generalized modified Bessel distribution.

REFERENCES

[1] Barndorff-Nielsen, O.,"Hyperbolic distributions and distributions on hyperbolae", Scand.J.Statist.5, P.P.151-157. 1978.

[2] Box, G.E.P. and Taio, G.C., “Bayesian inference in statistical Analysis”, Addition-Wesley Publishing Company, California, London, 1973.

[3] Butler, R. W., “Generalized inverse Gaussian distribution and their wishart connection", Colorado state university, Vol.25, P.P.69-75, USA, 1998.

[4] DeGroot, M.H., "Optimal statistical decision”, McGraw-Hill.1970.

[5] Hogg R. et al., "Introduction to Mathematical Statistics (6th edition)", Macmillan publishers company, New york, 2005.

[6] Hu, W., "Calibration of multivariate generalized hyperbolic distribution using the EM algorithm, With applications in risk management, Portfolio optimization and portfolio credit risk", Florida State University, Electronic theses, Treatises and Dissertation, The Graduate school, 2005.

[7] Koudou, A. E. and Ley, C., " Characterizations of GIG laws: a survey complemented with two new results", Proba. Surv. ,vol. 11 , p.p. 161-176, 2014.

[8] Lindley, D.V., "Bayesian statistics, a Review”, Philadelphia Society for Industrial and Applied Mathematics, 1972.

[9] Mora, J. A. N. & Mata, L. M., “Numerical aspects to estimate the generalized hyperbolic probability distribution”, Journal of Finance & Economics, vol. 1(4), p.p. 1-9, 2013.

[10] Press, S. J., "Subjective and objective Bayesian statistics”, A John Wiley and Sons, 2003.

[11] Saiful Islam, A. F. M., “Loss function, utility function and Bayesian sample size determination", PhD Thesis, University of London, 2011.

[12] Salih, S. A. and Aboudi, E. H., " Bayesian Inference of a Non-normal Multivariate Partial Linear Regression Model" Iraqi Journal of Statistical Science(34), pp. 91-115, 2021.

[13] Thabane, L. and Drekic, S., “Hypothesis testing for the generalized multivariate modified Bessel model", Journal of Multivariate Analysis,86, P.P.360-374, 2003.

[14] Thabane, L. & Haq, M. S., " The generalized multivariate modified Bessel distribution and its Bayesian applications", Journal of Applied Statistical Science, vol. 11(3), p.p.225-267, 2003.