On the common mean of several inverse Gaussian distributions based on a higher order likelihood method

S.H. Lin a,*, I.M. Wu b

a Department of Insurance and Finance, National Taichung Institute of Technology, 129, Sec. 3, Sanmin road, Taichung 404, Taiwan
b Institute of Statistics, National ChiaoTung University, Hsinchu, Taiwan

A R T I C L E   I N F O

Keywords:
Coverage probability
Expected length
Inverse Gaussian
Signed log-likelihood ratio statistic
Type I errors

A B S T R A C T

An interval estimation method for the common mean of several heterogeneous inverse Gaussian (IG) populations is discussed. The proposed method is based on a higher order likelihood-based procedure. The merits of the proposed method are numerically compared with the signed log-likelihood ratio statistic, two generalized pivot quantities and the simple t-test method with respect to their expected lengths, coverage probabilities and type I errors. Numerical studies show that the coverage probabilities of the proposed method are very accurate and type I errors are close to the nominal level.05 even for very small samples. The methods are also illustrated with two examples.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

In many application areas, such as demography, management science, hydrology, finance, etc., data are frequently positive and right-skewed. In the past four decades, the inverse Gaussian (IG) distribution has drawn much attention and the inferences concerned with the IG distribution have also grown rapidly because IG is an ideal candidate for modeling and analyzing the right-skewed and positive data. For instance, Wise [1,2] and Wise et al. [3] developed the IG population as a possible model to describe cycle time distribution for particles in the blood and Lancaster [4] made use of the IG distribution in describing strike duration data. Furthermore, IG distribution can also serve as a convenient prior for scale in Bayesian approaches to estimation with assumed Gaussian data [5]. The IG distribution can not only accommodate a variety of shapes, from highly skewed to almost normal, but also shares many elegant and convenient properties with Gaussian models; e.g., the associated inference methods are based on the well-known t, χ², and F distributions as for the normal case. See Chhikara and Folks [6], Seshadri [7,8] and Mudholkar and Tian [9] more details of Gaussian and IG analogies.

The probability density function (pdf) of IG distribution, IG (μ, λ), is defined as

\[ f(x; \mu, \lambda) = \left(\frac{\lambda}{2\pi \mu^3}\right)^{1/2} \exp\left\{ -\frac{\lambda}{2\mu^2} (x - \mu)^2 \right\}, \quad x > 0, \quad \mu > 0, \quad \lambda > 0 \tag{1.1} \]

where \(\mu\) is the mean parameter and \(\lambda\) is the scale parameter. The inference methods of the IG model are closely analogous to those of the Gaussian model; for example, a very common problem in applied fields is to compare the means of several Gaussian populations, i.e.

\[ H_0 : \mu_1 = \mu_2 = \cdots = \mu_i \quad \text{vs.} \quad H_1 : \text{not all } \mu_i \text{'s are equal} \]
If the variance of each population is homogeneous, the analysis of variance (ANOVA) can be used to perform the test. Similarly, the analysis of reciprocals (ANORE) can be used to test the equality of means of several IG samples if all scale parameters among groups are assumed to be equal [6]. When the scale parameters are non-homogeneous, the ANORE fails to solve the problem. Tian [10] proposed a method to test the equality of IG means under heterogeneity, based on a generalized test variable. However, when the null hypothesis is not rejected, the inferences for the common mean remain unsolved. Recently, Ye et al. [11] proposed a mixture method for the common mean problem based on generalized inference and the large sample theory. However, as the author has mentioned, if the sample sizes $n_i$ are not large and/or the scale parameter $\lambda_i$ is not large compared to $\mu$, the approximate distributions don’t fit well. Therefore, an alternate method which can be applied to general cases deserves further research.

In this paper, we will estimate and construct the 100(1 – $\alpha$)% confidence interval for the common mean of several non-homogeneous IG populations based on a higher order likelihood-based method. This method, in theory, has a higher order accuracy, $O(n^{-3/2})$, even when the sample size is small. Reid [12] provided a review and annotated the development of his method. The method has been applied to solve many practical problems involving interval estimation for a skewed distribution, e.g., Wu et al. [13] presented a confidence interval for a log-normal mean based on this method; Wu and Wong [14] used the method to improve the interval estimation for the two-parameter Birnbaum-Saunders distribution; and Tian and Wilding [15] used the method to construct confidence interval for the ratio of means of two independent IG distributions. In our case, the likelihood-based method also gives a satisfactory result for the problem of interval estimation for the common mean of several IG distributions.

The remainder of this article is organized as follows. In Section 2, we will briefly introduce the properties of IG distribution and the concepts of the signed log-likelihood ratio statistic and a higher order asymptotic method. The method is then applied to construct a confidence interval for the common mean of several independent IG populations in Section 3. We present several simulation studies and two numerical examples in Section 4 to illustrate the merits of our proposed method. Some concluding remarks are given in Section 5.

2. A general review

2.1. Some properties of IG distribution

For a random sample of $n$ observations $x_1, x_2, \ldots, x_n$ from IG ($\mu, \lambda$), the uniformly minimum variance unbiased estimators (UMVUES) of $\mu$ and $\lambda^{-1}$ are $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ and $W = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})$, respectively, and a minimum sufficient statistic of $(\mu, \lambda)$ is $\left(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i^2 \frac{1}{\bar{x}}\right)$. It is easy to verify that

$$\bar{x} \sim IG(\mu, n\bar{x}) \quad \text{and} \quad (n-1)W \sim \frac{1}{\lambda} \chi^2_{n-1}.$$  \hfill (2.1)

and that these two statistics are independently distributed.

Remark 1. Let $x \sim IG(\mu, \lambda)$ and $A \sim \frac{1}{2} \chi^2{n}$ be two independent random variables, then $\frac{\bar{x} - \mu}{\sqrt{V(x)}} \sim \chi^2_1$ and its distribution is independent of $\lambda A \sim \chi^2_n$. Let $M = \frac{\bar{x} - \mu}{\sqrt{V(x)}}$, then the distribution of $|M|$ is the truncated Student’s $t$ variable with $n$ degrees of freedom and $M^2$ has the F distribution with 1 and $n$ degrees of freedom [6].

From (2.1) and Remark 1, we know that $\frac{\bar{x} - \mu}{\sqrt{V(x)}} \sim \chi^2_1$ which is independent of $(n-1)W \sim \frac{1}{\lambda} \chi^2_{n-1}$. Let $U = \frac{\bar{x} - \mu}{\sqrt{V(x)}}$, then the distribution of $|U|$ is the truncated Student’s $t$ with $n-1$ degrees of freedom and $U^2 \sim F_{1, n-1}$.

2.2. The likelihood-based inference

Let $x = (x_1, x_2, \ldots, x_n)$ be an independent sample from some distribution and $l(\theta) = l(\theta; x)$ be the log-likelihood function based on the sample data. Suppose $\theta$ is the $p$-dimensional vector of parameters that can be partitioned into $(\mu, \hat{\lambda})$ with $\mu$ being the parameter of interest with dimension 1, and $\hat{\lambda}$ being the nuisance parameters with dimensions $p - 1$. The signed log-likelihood ratio statistic $r(\mu)$ for inference on $\mu$ is defined as

$$r(\mu) = \text{sgn}(\hat{\mu} - \mu) \left\{ \frac{1}{2} \left[ l(\hat{\theta}) - l(\hat{\theta}_{\mu}) \right] \right\}^{1/2},$$ \hfill (2.2)

where $\hat{\theta} = (\hat{\mu}, \hat{\lambda})$ is the overall maximum likelihood estimator (MLE) of $\theta$ and $\hat{\theta}_\mu = (\hat{\mu}, \hat{\lambda}_\mu)$ is the constrained MLE of $\theta$ for a given $\mu$. Cox and Hinkley [16] verified that $r(\mu)$ is asymptotically distributed as the standard normal distribution with first-order accuracy $O(n^{-1/2})$. A 100(1 – $\alpha$)% confidence interval for $\mu$ based on $r(\mu)$ can be obtained by

$$\{ \mu : |r(\mu)| \leq z_{\alpha/2} \},$$ \hfill (2.3)

where $z_{\alpha/2}$ is the 100(1 – $\alpha/2$)th percentile of a standard normal distribution. Since the log-likelihood ratio statistic is quite inaccurate when the sample size is small, Barndorff-Nielsen [17,18] proposed a higher order likelihood-based method which is known as the modified signed log-likelihood ratio,
where \( r(\mu) \) is the sign log-likelihood ratio statistic and \( q(\mu) \) is a statistic which can be expressed in various forms depending on the information available. A widely applicable formula for \( q(\mu) \) that ensures the \( O(n^{-3/2}) \) accuracy provided by Fraser et al. [19] is defined as

\[
q(\mu) = \left( \frac{[L(\hat{\theta}) - L(\hat{\theta}_0)]}{L(\hat{\theta})} \right) \left( \frac{[j_{\mu}(\hat{\theta})]}{j_{\mu}(\hat{\theta}_0)} \right)^{1/2},
\]

(2.5)

where \( j_{\mu}(\hat{\theta}) \) is the \( p \times p \) observed information matrix and \( j_{\mu}(\hat{\theta}_0) \) is the \((p-1) \times (p-1) \) observed nuisance information matrix. The vector array \( V = (v_1, \ldots, v_p) \) in (2.5) is obtained from a vector pivot quantity \( R(x; \theta) = (R_1(x; \theta), \ldots, R_n(x; \theta)) \) by

\[
V = \left( \frac{\partial R(x; \theta)}{\partial \theta} \right)^{-1} \left( \frac{\partial R(x; \theta)}{\partial \theta} \right)_L,
\]

(2.6)

where the distribution of \( R(x; \theta) \) is free of the nuisance parameters \( \lambda \). The quantity \( l_{\lambda}(\theta) \) is the likelihood gradient with

\[
l_{\lambda}(\theta) = \left\{ \frac{d}{d\theta_1} l(\theta; x), \ldots, \frac{d}{d\theta_p} l(\theta; x) \right\} = \left\{ \sum_{j=1}^n l_{\lambda}(\theta) v_{1j}, \ldots, \sum_{j=1}^n l_{\lambda}(\theta) v_{pj} \right\},
\]

(2.7)

where \( \frac{d}{d\theta} l(\theta; x) \) is the directional derivative of the log-likelihood function along \( \nu_i = (v_{i1}, \ldots, v_{ip}), \; i = 1, \ldots, p \), and \( l_{\lambda}(\theta) = \frac{\partial \nu}{\partial \lambda} \). Moreover,

\[
l_{\lambda}(\hat{\theta}_0) = \left. \frac{\partial \nu}{\partial \lambda} \right|_{\hat{\theta}_0}, \quad l_{\lambda}(\hat{\theta}) = \left. \frac{\partial l_{\lambda}(\theta)}{\partial \theta} \right|_{\hat{\theta}}, \quad \text{and} \quad l_{\lambda}(\hat{\theta}) = \left. \frac{\partial l_{\lambda}(\theta)}{\partial \theta} \right|_{\hat{\theta}}.
\]

Note that \( r^+(\mu) \) achieves third-order accuracy to a standard normal distribution [19]. Hence a 100(1 - \( \alpha \))% confidence interval for \( \mu \) based on \( r^+(\mu) \) is

\[
\{ \mu : |r^+(\mu)| \leq z_{\alpha/2} \}.
\]

(2.8)

### 3. Inferences for the common mean of several independent IG populations

#### 3.1. The likelihood-based confidence interval in the general case

Suppose \( x_i = (x_{i1}, x_{i2}, \ldots, x_{in}) \), \( i = 1, 2, \ldots, l \), are \( l \) random samples from IG \((\mu, \lambda_i)\) populations. The parameters, \( \theta = (\mu, \lambda_1, \ldots, \lambda_i) \), contain \( \mu \) being the parameter of interest and \( (\lambda_1, \ldots, \lambda_i) \) being the nuisance parameters. The log-likelihood function is

\[
l(\theta; x_{11}, x_{22}, \ldots, x_{ll}) = \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^n n_i \log \frac{\lambda_i}{2\pi} + \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^n \frac{\lambda_i x_{ij}^2}{2\mu^2} + \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^n \lambda_i x_{ij} - \frac{1}{2\mu} \sum_{i=1}^l \sum_{j=1}^n n_i \lambda_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^n \lambda_i.
\]

(3.1)

Differentiating the log-likelihood function (3.1) with respect to \( \theta \) for the first order yields the following results:

\[
\frac{\partial \nu}{\partial \mu} = -\frac{1}{\mu^2} \sum_{i=1}^l \sum_{j=1}^n n_i \lambda_i + \frac{1}{\mu^2} \sum_{i=1}^l \sum_{j=1}^n \lambda_i x_{ij},
\]

\[
\frac{\partial \nu}{\partial \lambda_i} = \frac{n_i}{2 \lambda_i} + \frac{1}{2} \sum_{j=1}^n \frac{x_{ij}}{\lambda_i} - \frac{1}{2\mu^2} \sum_{j=1}^n x_{ij}, \quad i = 1, \ldots, l.
\]

(3.2)

The overall MLEs \( \hat{\theta} = (\hat{\mu}, \hat{\lambda}_1, \ldots, \hat{\lambda}_i) \) can be uniquely obtained by solving the non-linear system (3.2) simultaneously. Furthermore, the constrained MLEs \( \hat{\theta}_\mu = (\mu, \hat{\lambda}_1, \ldots, \hat{\lambda}_i, \mu) \) for a given \( \mu \) are

\[
\hat{\lambda}_{i,\mu} = \frac{-n_i \mu^2}{2n_\mu - \sum_{j=1}^l n_j X_j - \mu^2 \sum_{j=1}^l \frac{1}{n_j}}, \quad i = 1, \ldots, l.
\]

(3.3)

Choosing a vector pivot quantity \( R = (R_{11}, \ldots, R_{kn}) \) with \( R_{ij} = \frac{\lambda_{i}(\delta_{ij} - \mu^2)}{\mu X_j - \mu^3}, \; i = 1, \ldots, l; \; j = 1, \ldots, n_j \), then \( R_{ij} \sim \chi^2_k \) with the distribution free of any unknown parameters. Differentiating \( R_{ij} \) with respect to \( x \) and \( \theta \), we have
\[
\frac{\partial R_{ij}}{\partial \lambda_k} = \left( \frac{\mu^2 x_{ij}^2}{\lambda_k (x_{ij}^2 - \mu^2)} \right)^{-1}, \quad \text{if } j = k; \quad \text{else } \frac{\partial R_{ij}}{\partial \lambda_k} = 0;
\]
\[
\frac{\partial R_{ij}}{\partial \mu} = -2\lambda_i (x_{ij} - \mu); \quad \frac{\partial R_{ij}}{\partial x_{ij}} = (x_{ij} - \mu)^2, \quad \text{if } j = k; \quad \text{else } \frac{\partial R_{ij}}{\partial x_{ij}} = 0.
\]

Furthermore, \( V = (v_1, \ldots, v_{I+1}) = -\frac{\partial}{\partial \theta} \left( \frac{\partial}{\partial \theta} \right)^{-1} \left( \frac{\partial}{\partial \theta} \right) \bigg|_{\theta} \) with

\[
v_1 = \left( \frac{2x_i^2}{\mu(x_i + \mu)}, \ldots, \frac{2x_i^2}{\mu(x_i + \mu)} \right), \quad v_{i+1} = \left( \frac{x_i(x_i - \mu)}{\lambda_i(x_i + \mu)}, \ldots, \frac{x_i(x_i - \mu)}{\lambda_i(x_i + \mu)} \right), \quad i = 1, \ldots, I.
\]

The likelihood gradients, \( l_v(\theta) = \left\{ \frac{d}{d\theta_1} l(\theta; x), \ldots, \frac{d}{d\theta_{I+1}} l(\theta; x) \right\} \), \( l_{vV}(\theta) \) and \( l_{v:V}(\theta) \) are

\[
l_{vV}(\theta) = \left[ \begin{array}{c}
\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} l_{x_1}(\theta) v_{1,j+1}(i-1) x_{i,j} - n_{i,j}}{n_{i,j}} \\
\vdots \\
\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} l_{x_I}(\theta) v_{I+1,j+1}(i-1) x_{i,j} - n_{i,j}}{n_{i,j}}
\end{array} \right],
\]

\[
l_{v:V}(\theta) = \left[ \begin{array}{c}
\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} l_{x_1}(\theta) v_{1,j+1}(i-1) x_{i,j} - n_{i,j}}{n_{i,j}} \\
\vdots \\
\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} l_{x_I}(\theta) v_{I+1,j+1}(i-1) x_{i,j} - n_{i,j}}{n_{i,j}}
\end{array} \right],
\]

respectively. The observed Fisher information matrix and the observed nuisance information matrix are

\[
j_{vV}(\theta) = \left[ \begin{array}{cccccc}
\frac{\sum_{i=1}^{n} l_{x_1}(\theta) x_{i,j}}{n_{i,j}^2} & \frac{\sum_{i=1}^{n} l_{x_1}(\theta) x_{i,j}}{n_{i,j}^2} & \cdots & \frac{\sum_{i=1}^{n} l_{x_1}(\theta) x_{i,j}}{n_{i,j}^2} \\
\frac{\sum_{i=1}^{n} l_{x_1}(\theta) x_{i,j}}{n_{i,j}^2} & \frac{\sum_{i=1}^{n} l_{x_1}(\theta) x_{i,j}}{n_{i,j}^2} & \cdots & \frac{\sum_{i=1}^{n} l_{x_1}(\theta) x_{i,j}}{n_{i,j}^2} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\sum_{i=1}^{n} l_{x_1}(\theta) x_{i,j}}{n_{i,j}^2} & \frac{\sum_{i=1}^{n} l_{x_1}(\theta) x_{i,j}}{n_{i,j}^2} & \cdots & \frac{\sum_{i=1}^{n} l_{x_1}(\theta) x_{i,j}}{n_{i,j}^2}
\end{array} \right],
\]

and

\[
j_{v:V}(\theta) = \left[ \begin{array}{cccccc}
\frac{n_{i,j}}{n_{i,j}^2} & \frac{n_{i,j}}{n_{i,j}^2} & \cdots & \frac{n_{i,j}}{n_{i,j}^2} \\
\frac{n_{i,j}}{n_{i,j}^2} & \frac{n_{i,j}}{n_{i,j}^2} & \cdots & \frac{n_{i,j}}{n_{i,j}^2} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{n_{i,j}}{n_{i,j}^2} & \frac{n_{i,j}}{n_{i,j}^2} & \cdots & \frac{n_{i,j}}{n_{i,j}^2}
\end{array} \right].
\]

Apply the above quantities to (2.5) and (2.4), \( q(\mu) \) and then \( r'(\mu) \) can be obtained.
Although the confidence intervals based on \( r(\mu) \) and \( r'(\mu) \) can be obtained here, in general, some simple numerical iteration procedure is needed to solve the upper bound limit and lower bound limit. In this paper we use the so-called secent method; the algorithm is summarized as follows:

Step 1: Give the tolerance \( \varepsilon \) for the purpose of accuracy;
Step 2: Select \( \delta \) for the purpose of numerical differentiation;
Step 3: Give the initial estimate \( \mu_0 \) to start the iteration;
Step 4: Compute

\[
\mu_i = \mu_0 + \frac{[Z_{z/2} - r(\mu_0)]}{r(\mu_0 + \delta) - r(\mu_0 - \delta)} \quad (3.5)
\]

Step 5: If \( |\mu_1 - \mu_0| > \varepsilon \), replace \( \mu_0 \) with \( \mu_1 \) and return to Step 4 again, otherwise take the latest \( \mu_1 \) as the lower bound limit of the 100(1 - \( \alpha \))\% confidence interval.

Replacing \( Z_{z/2} \) with \( Z_{1-z/2} \) in (3.5), we can obtain the upper bound limit for the 100(1 - \( \alpha \))\% confidence interval of the common mean \( \mu \). Similarly, the confidence interval based on \( r'(\mu) \) can be obtained by substituting \( r'(\mu) \) for \( r(\mu) \) in (3.5).

3.1.1. The likelihood-based confidence interval when \( I = 2 \)

In order to express the proposed method in details, we present the derivation of the confidence interval for the common mean \( \mu \). The log-likelihood function based on the observations is:

\[
\begin{align*}
I(\theta; x_1, x_2) &= \frac{n_1}{2} \log \frac{\lambda_1}{2\pi} - \frac{3}{2} \sum_{i=1}^{n_1} \log x_{1i} - \frac{\lambda_1}{2\mu} \sum_{i=1}^{n_1} x_{1i} + \frac{\lambda_1}{\mu} - \frac{\lambda_2}{2\mu} \sum_{i=1}^{n_1} \log x_{2i} - \frac{\lambda_2}{2\mu} \sum_{i=1}^{n_2} x_{2i} \\
&\quad + \frac{\lambda_1}{\mu} - \frac{\lambda_2}{2\mu} \sum_{i=1}^{n_2} \frac{1}{x_{2i}}.
\end{align*}
\]  

(3.6)

Take \( R_{ij} = \lambda_{ij} - \lambda^2 \) to be the pivot quantity as we mentioned earlier and differentiate \( R_{ij} \) with respect to \( x \) and \( \theta \), we then have

\[
\begin{align*}
R_1 &= \begin{pmatrix} -2\lambda_1 (x_1 - \mu) \\
\mu^2 x_1n_1 \\
\vdots \\
\mu^2 x_{2n_2} \\
\end{pmatrix}, \\
R_2 &= \begin{pmatrix} -2\lambda_2 (x_2 - \mu) \\
\mu^2 x_2n_2 \\
\vdots \\
\mu^2 x_{2n_2} \\
\end{pmatrix}, \\
R_3 &= \begin{pmatrix} 0, \ldots, 0, \mu^2 x_{2n_2} \\
\mu^2 x_2n_2, \ldots, \mu^2 x_{2n_2} \\
\vdots \\
\mu^2 x_{2n_2}, \ldots, \mu^2 x_{2n_2} \\
\end{pmatrix}.
\end{align*}
\]

(3.7)

Moreover, the likelihood gradients are

\[
I_\nu(\theta) = \begin{bmatrix}
\sum_{i=1}^{n_1} \frac{\lambda_1}{\mu(x_{1i} + \mu)} \left( \frac{1}{x_{1i}^2} - \frac{1}{x_{1i}^2} - \frac{3}{x_{1i}^2} \right) + \sum_{i=1}^{n_2} \frac{\lambda_2}{\mu(x_{2i} + \mu)} \left( \frac{1}{x_{2i}^2} - \frac{1}{x_{2i}^2} - \frac{3}{x_{2i}^2} \right) \\
\sum_{i=1}^{n_1} \frac{\lambda_1}{\mu(x_{1i} + \mu)} \left( \frac{1}{x_{1i}^2} - \frac{1}{x_{1i}^2} - \frac{3}{x_{1i}^2} \right) + \sum_{i=1}^{n_2} \frac{\lambda_2}{\mu(x_{2i} + \mu)} \left( \frac{1}{x_{2i}^2} - \frac{1}{x_{2i}^2} - \frac{3}{x_{2i}^2} \right) \\
\sum_{i=1}^{n_1} \frac{\lambda_1}{\mu(x_{1i} + \mu)} \left( \frac{1}{x_{1i}^2} - \frac{1}{x_{1i}^2} - \frac{3}{x_{1i}^2} \right) + \sum_{i=1}^{n_2} \frac{\lambda_2}{\mu(x_{2i} + \mu)} \left( \frac{1}{x_{2i}^2} - \frac{1}{x_{2i}^2} - \frac{3}{x_{2i}^2} \right)
\end{bmatrix},
\]  

Furthermore, if the scale parameters 

\[ s_i = \frac{2x_i}{\mu} + \frac{2(\bar{x} - x_i)}{\mu} \]  

and  

\[ s_i = \frac{2(\bar{x} - x_i)}{\mu} \]


\[ l_{SR}(\theta) = \begin{bmatrix} \sum_{i=1}^{n_1} \frac{2x_i^2}{\mu s_i} (\frac{1}{2\mu^2} - \frac{1}{2^{2\mu}}) & \frac{n_2}{\sqrt{n_2}} \frac{2x_i^2}{\mu s_i} (\frac{1}{2\mu^2} - \frac{1}{2^{2\mu}}) \\ \frac{n_2}{\sqrt{n_2}} \frac{2x_i^2}{\mu s_i} (\frac{1}{2\mu^2} - \frac{1}{2^{2\mu}}) & 0 \\ 0 & \sum_{i=1}^{n_2} (\bar{x}_i - x_i) (\frac{1}{2\mu^2} - \frac{1}{2^{2\mu}}) \end{bmatrix} \]

Furthermore, the observed Fisher information matrix and the observed nuisance information matrix are

\[ J_{SG}(\theta) = \begin{bmatrix} \frac{2x_i^2}{\mu} + \frac{2x_i^2}{\mu} + \frac{2x_i^2}{\mu} + \frac{2x_i^2}{\mu} & \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} \\ \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} & 0 \\ \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} & \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} \end{bmatrix} \]

and  

\[ J_{SN}(\theta) = \begin{bmatrix} \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} \\ \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} & 0 \\ \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} & \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} + \frac{n_1}{\sqrt{n_1}} \frac{2x_i^2}{\mu} \end{bmatrix} \]


Ye et al. [11] proposed a mixture of generalized inference method and the large sample theory. Their procedures for deriving two generalized pivot quantities, \( \tilde{T}_1 \) and \( \tilde{T}_2 \), are briefly introduced below.

Suppose \( x_i = (x_{i1}, x_{i2}, \ldots, x_{in_i}) \), \( i = 1, 2, \ldots, I \), are \( I \) independent populations with parameters \( (\mu, \lambda_i) \) for each population, from (2.1) we know that \( x_i \sim IG(\mu, \lambda_i) \) and \( n_i x_i \sim \chi^2_{n_i} \), where \( V_i = \frac{1}{n} \sum_{i=1}^{n_i} (\lambda_i - \frac{1}{\lambda_i}) \). Let \( x_i^* \) and \( v_i^* \) denote the observed values of \( x_i \) and \( V_i \) respectively. The author defined the first generalized pivot quantity (GP1) for common mean \( \mu \) as

\[ \tilde{T}_1 = \frac{\sum_{i=1}^{n_i} R_i T_i}{\sum_{i=1}^{n_i} R_i} \]

where

\[ R_i = \frac{n_i \lambda_i V_i}{n_i x_i^*} \sim \chi^2_{n_i - 1} \text{ and } T_i = \frac{x_i^*}{1 + \sqrt{\frac{n_i \lambda_i (x_i^* - \mu)}{\mu^2}} \sqrt{\frac{n_i \lambda_i}{\mu^2}}} \]

(3.8)

are the generalized pivot quantities for \( \lambda_i \) and \( \mu \), respectively, based on the \( i \)th sample. It is noted that \( \chi^2 \) denotes “approximately distributed” and \( \sqrt{\frac{n_i \lambda_i (x_i^* - \mu)}{\mu^2}} \sqrt{\frac{n_i \lambda_i}{\mu^2}} \). When the sample sizes \( n_i \) are large and/or the scale parameter \( \lambda_i \) is large compared to \( \mu \), let \( \tilde{T}_1(\alpha) \) be the 100\( \alpha \)th percentile of \( \tilde{T}_1 \), the 100\( (1-\alpha) \% \) generalized confidence interval for \( \mu \) based on \( \tilde{T}_1 \) is

\[ \{ \tilde{T}_1(\alpha/2), \tilde{T}_1(1-\alpha/2) \} \]

(3.10)

The generalized p-value for testing \( H_0 : \mu = \mu_0 \) vs. \( H_1 : \mu \neq \mu_0 \) can be computed as

\[ p_1 = 2 \min \{ Pr(\tilde{T}_1 \leq \mu_0), Pr(\tilde{T}_1 \geq \mu_0) \} \]

(3.11)

Furthermore, if the scale parameters \( \lambda_i \)'s are known, then \( \hat{\mu} = \frac{\sum_{i=1}^{n_i} n_i \lambda_i x_i}{n_i} \) is the MLE of \( \mu \) based on \( I \) independent IG populations and \( \hat{\mu} \sim IG(\mu, \sum_{i=1}^{n_i} n_i \lambda_i) \) [6]. The author provided a second generalized pivot quantity (GP2) for \( \mu \) as

\[ \tilde{T}_2 = \frac{R}{1 + \sqrt{\frac{\sum_{i=1}^{n_i} n_i \lambda_i (\hat{\mu} - \mu)}{\mu^2}} \sqrt{\frac{R}{\sum_{i=1}^{n_i} R_i}}} \]

(3.12)

where

\[ R = \frac{\sum_{i=1}^{n_i} n_i \lambda_i \sqrt{\mu}}{\sum_{i=1}^{n_i} R_i} \]  

\( Z \sim N(0, 1) \) and \( R_i \) is defined in (3.9). The 100\( (1-\alpha) \% \) generalized confidence interval for \( \mu \) and the generalized p-value for testing \( H_0 : \mu = \mu_0 \) vs. \( H_1 : \mu \neq \mu_0 \) based on \( \tilde{T}_2 \) can be respectively obtained by substituting \( \tilde{T}_2 \) in place of \( \tilde{T}_1 \) in (3.10) and (3.11).
3.3. Simple t-test confidence interval

For the purpose of comparison, we calculate a simple t-test confidence interval that is inspired from the analysis of reciprocals (ANORE). This method can provide an exact confidence interval when the scale parameters are homogeneous. Suppose \( x_i = (x_{j1}, x_{j2}, \ldots, x_{jn}) \), \( i = 1, 2, \ldots, J \) are independent populations with parameters \((\mu, \lambda_i)\) for each population. It can be shown that \( N^2 \sum_{j=1}^N x_{ij} \sim IG(\mu, N \lambda_i) \) and \((N - 1) W \sim \chi^2_{N-1} \) are independent distributed, where \( N = \sum_{i=1}^N n_i, \ x_i = \frac{1}{n} \sum_{j=1}^{n_i} x_{ij} \) and \( W = \frac{1}{N-1} \sum_{i=1}^{N} (x_{ij} \lambda_i^{-1}) \). Moreover, from Remark 1, we know \( \frac{|\bar{x} \mu - \mu|}{\sqrt{N x^2}} \) is the truncated student’s t distribution with \( N - 1 \) degrees of freedom. Therefore, the two-sided 100(1 – \( x \))% for \( \mu \) is

\[
\left\{ \begin{array}{ll}
\bar{x} \left( 1 + t_1 \frac{\sqrt{x^2}}{N} \right)^{-1}, & \text{if } 1 - t_1 \frac{\sqrt{x^2}}{N} > 0 \\
\bar{x} \left( 1 + t_1 \frac{\sqrt{x^2}}{N} \right)^{-1}, & \text{otherwise.}
\end{array} \right.
\]

(3.13)

4. Simulation studies and numerical examples

4.1. Simulation studies

To illustrate the merits of the proposed method, we present simulation studies of the confidence intervals and type I errors applied to a variety of scale parameter configurations and different combinations of small sample sizes for two and three populations. In the simulation, we exhibit the coverage probabilities, the average lengths of the 95% confidence intervals and \( \tilde{x} \) to these coverage probabilities decrease when the heterogeneity increases. Moreover, when the scale parameter is small relative to \( \overline{\mu} \), the interval lengths constructed by the simple t-test are unbounded (i.e., a one-sided interval). In these cases, the confidence intervals based on the simple t-test method have good coverage probabilities but compared to the other three methods, the coverage probabilities are too liberal to attain the proposed coverage probabilities. The results given in Tables 1–4 below are based on 10,000 simulation runs for each combination.

From Tables 1 and 2, we see that although the confidence intervals based on \( r(\mu) \) and GP2 have shorter average lengths comparing to the other three methods, the coverage probabilities are too liberal to attain the proposed coverage probabilities.95 for each combination. The confidence intervals based on the simple t-test method have good coverage probabilities but these coverage probabilities decrease when the heterogeneity increases. Moreover, when the scale parameter is small relative to \( \overline{\mu} \), the interval lengths constructed by the simple t-test are unbounded (i.e., a one-sided interval). In these cases, the simple t-test method gives less information about the target value than those based on the other methods. The GP1 method performs well in the coverage probabilities when the scale parameters are large compared to \( \overline{\mu} \), but the performance grows worse when the scale parameters decrease. On the other hand, the confidence intervals based on \( r^*(\mu) \) not only have almost exact coverage probabilities in each combination (except for few exceptions), but the average lengths are also quite decent and acceptable. Therefore, in terms of the overall comparisons, the higher order likelihood-based method outperforms the other four methods.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation results of 95% confidence interval of ( \mu = 1 ) for two populations.</td>
</tr>
<tr>
<td>( (n_1, n_2) )</td>
</tr>
<tr>
<td>( (5, 10) )</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>( (10, 5) )</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>( (10, 10) )</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

CP: coverage probability; length: average length.
Simulation results of 95% confidence interval of Table 2 type I errors based on GP2 and errors based on the GP1 method are getting worse compared to the nominal level.\(0.05\) as the mean parameters increase. The type I errors decrease as the mean parameter under the null hypothesis increases. Similarly, the type I errors obtained by the compared to the nominal level \(0.05\). By contrast, the type I errors based on GP1 0.0587 0.0467 0.0472 0.0404 0.0355 0.0602 0.0562 0.0502 0.0395 0.0347

Furthermore, from Tables 3 and 4, we can see the type I errors based on the simple t-test method are not stable since the type I errors decrease as the mean parameter under the null hypothesis increases. Similarly, the type I errors obtained by the GP1 methods do not perform well under those small sample sizes and small scale parameters configurations. The type I errors based on the GP1 method are getting worse compared to the nominal level \(0.05\) as the mean parameters increase. The type I errors based on GP2 and \(r(\mu)\) are respectively around 0.6 to 0.7 and 0.6 to 1.0 for each combination which are too large compared to the nominal level \(0.05\). By contrast, the type I errors based on \(r(\mu)\) are not only stable, but the values are also very close to the nominal level \(0.05\). Thus, we can say that the proposed procedure can well tolerate heterogeneity among populations and give robust and reliable results under different scenarios.

### Table 2
Simulation results of 95% confidence interval of \(\mu - 1\) for three populations.

<table>
<thead>
<tr>
<th>((n_1, n_2, n_3))</th>
<th>(l_1)</th>
<th>(l_2)</th>
<th>(n_1)</th>
<th>(n_2)</th>
<th>(n_3)</th>
<th>(r(\mu))</th>
<th>(\mu = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>((5,8,10))</td>
<td>0.1</td>
<td>0.1</td>
<td>1</td>
<td>0.949</td>
<td>4.885</td>
<td>0.972</td>
<td>6.162</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.5</td>
<td>1</td>
<td>0.949</td>
<td>3.382</td>
<td>0.966</td>
<td>21.096</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>5</td>
<td>0.948</td>
<td>0.671</td>
<td>0.952</td>
<td>0.807</td>
<td>0.934</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>10</td>
<td>0.949</td>
<td>0.467</td>
<td>0.958</td>
<td>0.606</td>
<td>0.939</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>5</td>
<td>0.948</td>
<td>0.393</td>
<td>0.944</td>
<td>0.497</td>
<td>0.931</td>
</tr>
</tbody>
</table>

### Table 3
Type I errors for \(H_0: \mu = \mu_0\) vs. \(H_1: \mu \neq \mu_0\) at \(l = 2\) and \(x = 0.05\).

<table>
<thead>
<tr>
<th>((\lambda_1, \lambda_2))</th>
<th>Tests (n_1 = 5, n_2 = 10)</th>
<th>(n_1 = 10, n_2 = 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r(\mu)) (\mu_0)</td>
<td>(0.2)</td>
<td>(0.8)</td>
</tr>
<tr>
<td>GP1</td>
<td>0.0522</td>
<td>0.0528</td>
</tr>
<tr>
<td>GP2</td>
<td>0.0439</td>
<td>0.0432</td>
</tr>
<tr>
<td>S.T.</td>
<td>0.0615</td>
<td>0.0430</td>
</tr>
<tr>
<td>(r(\mu)) (\mu_0)</td>
<td>0.0485</td>
<td>0.0562</td>
</tr>
<tr>
<td>GP1</td>
<td>0.0590</td>
<td>0.0363</td>
</tr>
<tr>
<td>GP2</td>
<td>0.0685</td>
<td>0.0632</td>
</tr>
<tr>
<td>S.T.</td>
<td>0.0832</td>
<td>0.0862</td>
</tr>
<tr>
<td>(r(\mu)) (\mu_0)</td>
<td>0.0555</td>
<td>0.0507</td>
</tr>
<tr>
<td>GP1</td>
<td>0.0574</td>
<td>0.0514</td>
</tr>
<tr>
<td>GP2</td>
<td>0.0619</td>
<td>0.0660</td>
</tr>
<tr>
<td>S.T.</td>
<td>0.0783</td>
<td>0.0783</td>
</tr>
<tr>
<td>(r(\mu)) (\mu_0)</td>
<td>0.0583</td>
<td>0.0545</td>
</tr>
<tr>
<td>GP1</td>
<td>0.0507</td>
<td>0.0566</td>
</tr>
<tr>
<td>GP2</td>
<td>0.0587</td>
<td>0.0467</td>
</tr>
<tr>
<td>S.T.</td>
<td>0.0788</td>
<td>0.0803</td>
</tr>
<tr>
<td>(r(\mu)) (\mu_0)</td>
<td>0.0634</td>
<td>0.0616</td>
</tr>
<tr>
<td>GP1</td>
<td>0.0528</td>
<td>0.0550</td>
</tr>
<tr>
<td>GP2</td>
<td>0.0559</td>
<td>0.0497</td>
</tr>
<tr>
<td>S.T.</td>
<td>0.0751</td>
<td>0.0759</td>
</tr>
</tbody>
</table>

CP: coverage probability; length: average length.
Table 4
Type I errors for $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$ at $I = 3$ and $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>$(k_1, k_2, k_3)$</th>
<th>Tests</th>
<th>$(n_1, n_2, n_3) = (5, 8, 10)$</th>
<th>$(n_1, n_2, n_3) = (5, 10, 8)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r'(\mu)$</td>
<td>$\mu_0$</td>
<td>$r'(\mu)$</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td>(0.1, 0.1, 1)</td>
<td>GP1</td>
<td>0.0481</td>
<td>0.0564</td>
</tr>
<tr>
<td></td>
<td>GP2</td>
<td>0.0656</td>
<td>0.0600</td>
</tr>
<tr>
<td></td>
<td>S.T.</td>
<td>0.0483</td>
<td>0.0391</td>
</tr>
<tr>
<td>(0.1, 0.5, 1)</td>
<td>GP1</td>
<td>0.0539</td>
<td>0.0572</td>
</tr>
<tr>
<td></td>
<td>GP2</td>
<td>0.0884</td>
<td>0.0666</td>
</tr>
<tr>
<td></td>
<td>S.T.</td>
<td>0.0594</td>
<td>0.0462</td>
</tr>
<tr>
<td>(1.1, 5)</td>
<td>GP1</td>
<td>0.0520</td>
<td>0.0663</td>
</tr>
<tr>
<td></td>
<td>GP2</td>
<td>0.0812</td>
<td>0.0832</td>
</tr>
<tr>
<td></td>
<td>S.T.</td>
<td>0.0572</td>
<td>0.0594</td>
</tr>
<tr>
<td>(1.1, 10)</td>
<td>GP1</td>
<td>0.0485</td>
<td>0.0516</td>
</tr>
<tr>
<td></td>
<td>GP2</td>
<td>0.0570</td>
<td>0.0478</td>
</tr>
<tr>
<td></td>
<td>S.T.</td>
<td>0.0579</td>
<td>0.0534</td>
</tr>
<tr>
<td>(1.5, 10)</td>
<td>GP1</td>
<td>0.0517</td>
<td>0.0514</td>
</tr>
<tr>
<td></td>
<td>GP2</td>
<td>0.0642</td>
<td>0.0540</td>
</tr>
<tr>
<td></td>
<td>S.T.</td>
<td>0.0638</td>
<td>0.0659</td>
</tr>
</tbody>
</table>

Table 5
Data for Example 1.

<table>
<thead>
<tr>
<th>Population $i$</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_i$</td>
<td>0.7312</td>
<td>1.3932</td>
<td>1.6999</td>
</tr>
<tr>
<td>$w_i$</td>
<td>1.7314</td>
<td>0.5394</td>
<td>1.2698</td>
</tr>
<tr>
<td>$\bar{x}_i$</td>
<td>0.7109</td>
<td>1.6046</td>
<td>0.7887</td>
</tr>
<tr>
<td>$w_i$</td>
<td>0.0303</td>
<td>2.0649</td>
<td>1.0535</td>
</tr>
<tr>
<td>$\bar{x}_i$</td>
<td>0.7044</td>
<td>0.7929</td>
<td>0.7973</td>
</tr>
<tr>
<td>$w_i$</td>
<td>0.0579</td>
<td>1.2238</td>
<td>1.4988</td>
</tr>
<tr>
<td>$\bar{x}_i$</td>
<td>0.0538</td>
<td>31.3779</td>
<td>1.4685</td>
</tr>
<tr>
<td>$w_i$</td>
<td>0.7816</td>
<td>1.2252</td>
<td>0.4620</td>
</tr>
</tbody>
</table>

$w_i = \sum_{j=1}^{n_i} (x_i^j - \bar{x}_i^1)$.

Table 6
The 95% confidence intervals for the common mean.

<table>
<thead>
<tr>
<th>Method</th>
<th>Point estimate $\mu$</th>
<th>Interval estimate</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r'(\mu)$</td>
<td>1.221</td>
<td>(0.961, 1.728)</td>
<td>0.767</td>
</tr>
<tr>
<td>GP1</td>
<td>1.817</td>
<td>(0.972, 2.089)</td>
<td>1.117</td>
</tr>
<tr>
<td>GP2</td>
<td>1.258</td>
<td>(0.952, 1.691)</td>
<td>0.739</td>
</tr>
<tr>
<td>$r'(\mu)$</td>
<td>1.221</td>
<td>(0.980, 1.605)</td>
<td>0.625</td>
</tr>
<tr>
<td>S.T.</td>
<td>1.078</td>
<td>(0.553, 2.071)</td>
<td>20.158</td>
</tr>
</tbody>
</table>

Table 7
Data for Example 2 ($n_i = 10, i = 1, 2, 3, 4$).

| $x_i$ | 2.635 | 2.055 | 1.748 | 2.023 |
| $w_i$ | 4.5147 | 107.7351 | 3.1558 | 68.4330 |

$w_i = \sum_{j=1}^{n_i} (x_i^j - \bar{x}_i^1)$.
4.2. Two numerical examples

**Example 1.** We first present a three population IG simulated data with \((n_1, n_2, n_3) = (5, 6, 7)\) and \((\mu_1, \mu_2, \mu_3) = (1, 0.2, 1, 10)\) as illustrative example. The original data and the summary data are depicted in Table 5. The interval estimations based on five methods are given in Table 6. Four confidence intervals based on \(r^*(\mu)\), \(r(\mu)\), GP1 and GP2 give satisfactory result under the heterogeneous data set when compared with that based on the simple t-test method. Although the one based on \(r^*(\mu)\) is a little wider than those of GP2 and \(r(\mu)\), in general, it gives better coverage probabilities compared with them.

**Example 2.** The data is available in p.462, Nelson [20]. In this data, there are 60 “times-to-breakdown” in minutes of an insulating fluid subjected to high voltage stress. Since IG distribution is widely applied as a lifetime model in reliability analysis, here we consider the failure time of the insulating fluid for each group as an IG distributed random variable. If the experiment was under control, the mean of each group should be the same. For illustrative purpose, we pick the first four groups as demonstration and apply the procedure induced by Tian [10] to test the equality of the means for the first four groups. The resulting \(p\)-value is 0.8693; we can follow up by constructing the confidence interval for the common mean parameter. The summary data and the results are given in Tables 7 and 8, respectively.

From Table 7, we see \(w_i, i = 1, 2, 3, 4\) the estimators of the reciprocal of the scale parameters are quite different among groups implying the existence of heterogeneity.

In Table 8, all five intervals cover the corresponding point estimates and those based on \(r^*(\mu)\), \(r(\mu)\) and GP2 give satisfactory interval lengths compared to GP1 and the simple t-test. In this case, the simple t-test only provides a one-sided interval.

5. Conclusions

In this paper, we presented a higher order likelihood-based procedure to construct the confidence interval of the common mean of several independent IG populations. In our simulation, we compared this procedure with four alternative methods. The numerical examples showed that the proposed method gives nearly exact coverage probabilities and the type I errors calculated are close to the nominal level 0.05 even for small sample sizes and small scale parameters. The method is able to integrate the information of several heterogeneous IG populations, and therefore is useful for a variety of practical applications.

Acknowledgement

The author cordially thanks the Editor-in-Chief and anonymous reviewers for their valuable comments which led to the improvement of this paper. This research has been supported by grant NSC 98-2118-M-025-001 from the National Science Council of Taiwan.

References