Asian Journal of Statistical Sciences
Vol. 5, No-1, (2025), pp. 01-12

ARF INDIA
Academic Open Access Publishing
Journal homepage link:
http://arfjournals.com/ajss

Simultaneous fiducial generalized confidence intervals for the successive difference of ordered location parameters for Inverse Gaussian distribution under heteroscedasticity

Mehak Jindal[a] and Narinder Kumar[b]

[a, b] *Department of Statistics, Panjab University, Chandigarh, India (160014).*

ARTICLE HISTORY
Compiled September 24, 2025

Received 07 April 2025; Accepted 21 August 2025

ABSTRACT
In this paper, we propose fiducial simultaneous generalized confidence intervals based on fiducial generalized pivotal quantities for the comparison of the successive differences of ordered location parameters of the Inverse Gaussian populations. The constructed intervals are shown to attain the correct coverage probabilities asymptotically. The simulation study indicates that the Type-I error probabilities are close to the nominal level. The illustration of the proposed procedure is done using real datasets.

Keywords: Inverse Gaussian, Fiducial generalized pivotal quantities, Type-I error, Simultaneous fiducial generalized confidence interval.

1. Introduction

Consider independent samples of size n_i drawn from $k(k > 2)$ populations, where observations from i^{th} population corresponds to a two-parameter Inverse Gaussian distribution having probability density function

$$f(x;\mu_i,\lambda_i) = \left(\frac{\lambda_i}{2\pi x^3}\right)^{(1/2)} exp\left\{-\frac{\lambda_i}{2x\mu_i^2}\right\}(x-\mu_i)^2, \quad x_i > 0 \tag{1}$$

where $\mu_i \in (0,\infty)$ is the location parameter and $\lambda_i \in (0,\infty)$ is the scale parameter,$i = 1, 2, \ldots, k$. The Inverse Gaussian distribution is well-suited for modelling rightly skewed and non-negative data values. The applications of Inverse Gaussian distribution in multiple fields including demography, finance, cardiology, hydrology and linguistics, are discussed in Chhikara (2024), as well as Seshadri (1994). For instance, in pharmacokinetics, it is crucial to evaluate the absorption profile of oral drug formulations at different concentration levels to understand their bioavailability and therapeutic efficacy. In health care, the Inverse Gaussian distribution can model the recovery time for patients after specific surgery or treatment of a particular condition,

CONTACT Narinder Kumar. Email: nkumar67@yahoo.com

categorized by three different treatment methods: standard care, accelerated recovery protocol, and experimental treatment.Here, the goal is to determine whether the mean recovery times of successive treatments differ significantly. In the literature, for the data arising from various Inverse Gaussian populations, based on a generalised test variable, Tian (2006) proposed a procedure for testing the homogeneity of means under heteroscedasticity. The concept of a generalized pivotal quantity was introduced by Weerahandi (1993). A sub-class of Weerahandi's generalized pivotal quantities called fiducial generalised pivotal quantities (FGPQs) was introduced by Hannig, Iyer and Patterson (2006) and Hannig, Lidong, Abdel-Karim and Iyer (2006) proposed simultaneous fiducial generalised confidence intervals (SFGCIs) for the ratio of means of log-normal distributions using the idea of FGPQs. A parametric bootstrap (PB) method to test the equality of the Inverse Gaussian location parameters in the presence of heterogeneity was discussed by Ma and Tian (2009). In one way layout, two types of simultaneous confidence intervals under heteroscedasticity based on FGPQs were proposed by Xiong and Mu (2009). Shi and Lv (2012) proposed a new generalized p-value to test the equality of location parameters. Kharrati-Kopaei et al. (2013) proposed single stage SFGCIs for the comparisons of successive differences of several exponential location parameters under heteroscedasticity. Zhang (2014) proposed simultaneous confidence intervals of all pairwise location parameters of several Inverse Gaussian distributions based on FGPQs. Later, Kharrati-Kopaei (2015) proposed a classical approach and bootstrap approach for simultaneous pairwise confidence intervals for the comparison of Inverse Gaussian location parameters. Kharrati-Kopaei and Eftekhar (2017) formulated simultaneous confidence intervals to test the homogeneity of multiple Inverse Gaussian location parameters under heteroscedasticity using a parametric bootstrap approach. They demonstrated that the coverage probability of the proposed procedure approaches fairly close to the nominal value in contrast to Zhang (2014) procedure. To the best of our knowledge, simultaneous confidence intervals for successive comparisons of the ordered location parameters of Inverse Gaussian distributions under heteroscedasticity have not been yet constructed. This paper focuses on the development of simultaneous confidence intervals based on FGPQs, given the applicability of Inverse Gaussian distribution in numerous fields.

In Section 2, an overview of the Inverse Gaussian distribution and the generalized pivotal quantities is provided. Section 3, presents the construction of SGFCIs for the successive differences of the ordered Inverse Gaussian location parameters when the scale parameters as well as sample sizes are unequal. A simulation study conducted is provided in Section 4, indicating that the probabilities of Type-I error of the proposed SGFCIs are close to the nominal level. Additionally, the estimated power of the procedure is evaluated. A power comparison is done between the proposed method and all pairwise comparison procedure. In Section 5, we illustrate the result using examples. The conclusion of the study is presented in Section 6.

2. Necessary Background and Notations

Let $X_{ij} \sim IG(\mu_i, \lambda_i), i = 1, 2, \ldots, k, j = 1, 2, \ldots, n_i$ be the independent samples of size n_i drawn randomly from k Inverse Gaussian populations. Further let $N_o = \sum_{i=1}^{k} n_i$ and $\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$ be the i^{th} sample mean. The maximum likelihood estimator of the

parameters μ_i and λ_i are as follows:

$$\hat{\mu_i} = \bar{X}_i \tag{2a}$$

$$\hat{\lambda_i} = \left[\frac{1}{n_i}\sum_{j=1}^{n_i}\left(\frac{1}{X_{ij}} - \frac{1}{\bar{X}_i}\right)\right]^{-1} = V_i(say), \quad i = 1, 2, ..., k. \tag{2b}$$

It is well established that $\hat{\mu_i} \sim IG(\mu_i, n_i\lambda_i)$ and $\frac{n_i\lambda_i}{\hat{\lambda_i}} \sim \chi^2_r, i = 1, 2, \ldots, k$, where χ^2_r represents a chi-square variate with r degrees of freedom. Furthermore, $\hat{\mu_i}$ and $\hat{\lambda_i}$ as defined in (2.1a) and (2.1b), are mutually independent as well as complete sufficient statistics for μ_i and λ_i, respectively as given in [**?**]. The estimated vectors of location and scale parameters are denoted as $\hat{\mu_i} = (\hat{\mu_1}, \hat{\mu_2}, ..., \hat{\mu_k})$ and $\hat{\lambda_i} = (\hat{\lambda_1}, \hat{\lambda_2}, ..., \hat{\lambda_k})$, respectively. The reciprocal of location parameters vector and its estimate are respectively denoted as $\theta = (\theta_1, \theta_2, \ldots, \theta_k)$ and $\hat{\theta} = (\hat{\theta_1}, \hat{\theta_2}, ..., \hat{\theta_k})$ where $\theta_i = \frac{1}{\mu_i}$ and $\hat{\theta}_i = \frac{1}{\hat{\mu_i}}$. [**?**] introduced the following generalized pivotal quantity for θ_i,

$$T_{\theta_i} = \hat{\theta}_i \left|1 + Z_i\sqrt{\frac{1}{n_i\hat{\theta}_i T_{\lambda_i}}}\right|, \tag{3}$$

where $Z_i \sim N(0, 1)$ and $T_{\lambda_i} = \frac{U_i\hat{\lambda_i}}{n_i}$ is the generalized pivotal quantity for λ_i introduced by [**?**], where $U_i \sim \chi^2_{(n_i-1)}$.

3. Simultaneous Fiducial Generalized Confidence Intervals

In the present section, simultaneous confidence intervals are proposed based on FGPQs for the successive differences of the ordered location parameters of k Inverse Gaussian populations. The problem of testing is as follows $H_{oi} : \theta_{i+1} = \theta_i$ vs $H^A_{oi} : \theta_i \leq \theta_{i+1}$, $i = 1, 2, \ldots, k-1$.
The FGPQs for $\theta_{i,i+1}$ is

$$T_{\theta_{i,i+1}} = T_{\theta_{i+1}} - T_{\theta_i} = \hat{\theta}_{i+1}\left|1 + Z_{i+1}\sqrt{\frac{1}{n_{i+1}\hat{\theta}_{i+1}T_{\lambda_{i+1}}}}\right| - \hat{\theta}_i\left|1 + Z_i\sqrt{\frac{1}{n_i\hat{\theta}_i T_{\lambda_i}}}\right|. \tag{4}$$

Let $\bar{X} = (\bar{X}_1, \bar{X}_2, ..., \bar{X}_k)'$ and $V = (V_1, V_2, \ldots, V_k)'$ denotes the vector estimators of $(\mu_1, \mu_2, ..., \mu_k)'$ and $(\lambda_1, \lambda_2, ..., \lambda_k)'$ respectively. As given by Xiong and Mu (2009), the conditional expectation of $T_{\theta_{i,i+1}}$ is

$$\delta_{i,i+1} = E(T_{\theta_{i,i+1}}|(\bar{X}, V)) = \hat{\theta}_{i+1} - \hat{\theta}_i = \frac{1}{\bar{X}_{i+1}} - \frac{1}{\bar{X}_i}, \quad 1 \leq i \leq k-1. \tag{5}$$

and variance of $T_{\theta_{i,i+1}}$ is

$$\xi_{i,i+1} = Var\big(T_{\theta_{i,i+1}} \mid (\bar{X}, V)\big) = \frac{\hat{\theta}_{i+1}}{\hat{\lambda}_{i+1}\,(n_{i+1}-3)} + \frac{\hat{\theta}_i}{\hat{\lambda}_i\,(n_i-3)}$$

$$= \frac{V_{i+1}}{\bar{X}_{i+1}\,(n_{i+1}-3)} + \frac{V_i}{\bar{X}_i\,(n_i-3)}. \tag{6}$$

The $100(1-\alpha)\%$ SFGCIs for the successive differences of the ordered location parameters are

$$\theta_{i,i+1} \in \left(\frac{1}{\bar{X}_{i+1}} - \frac{1}{\bar{X}_i}\right) \pm q_\alpha \sqrt{\xi_{i,i+1}}\,, \quad i = 1, 2, \ldots, k-1, \tag{7}$$

where q_α is the upper α^{th} quantile of the conditional distribution of $M^{'}$ where

$$M' = \max_{1 \le i \le k-1} \left| \frac{\theta_{i,i+1} - E(T_{\theta_{i,i+1}} | (\bar{X}, V))}{\sqrt{\xi_{i,i+1}}} \right| \tag{8}$$

The value of q_α is obtained by the algorithm mentioned below.

Algorithm

(1) Generate a random sample X_{ij} of size n_i from an Inverse Gaussian population characterized by parameters μ_i and λ_i, for $i = 1, 2, ..., k$.
(2) Compute the observed values $\bar{x}_i$ and v_i corresponding to $\bar{X}_i$ and $V_i, i = 1, 2, \ldots, k, j = 1, 2, ..., n_i$. Then, initiate the loop for $l = 1, 2, ..., L$.
(3) For $i = 1, 2, \ldots, k$, generate a standard normal random variate $Z_i \sim N(0, 1)$ and a V_i with distribution $\dfrac{\chi^2_{(n_i-1)}}{n_i \lambda_i}$.
(4) Using the equations (3.1) - (3.4), compute, $T_{\theta_{i,i+1}}$ $\delta_{i,i+1}$ $\xi_{i,i+1}$ and $M^{'}$.
(5) End l loop.
(6) Compute $q_\alpha, 100(1-\alpha)\%$ percentile of $M^{'}$.

The following theorem is about development of simultaneous confidence intervals.

Theorem 3.1. *Let $X_{i1}, X_{i2}, ..., X_{in_i}$ be a random sample of size n_i from mutually independent k Inverse Gaussian populations. Assume $N = \sum_{1=i}^{k} n_i$ and $N \to \infty$ such that the ratio $\dfrac{n_i}{N} \to \tau_i \in (0, 1)$, for $i = 1, 2, ..., k$. Then $P\Big[\theta_{i+1} - \theta_i \in \big(\frac{1}{\bar{X}_{i+1}} - \frac{1}{\bar{X}_i}\big) \pm q_\alpha \sqrt{\xi_{i,i+1}}\Big] \xrightarrow{P} (1-\alpha)$.*

Proof. According to the central limit theorem, the vector $\sqrt{N}\big((\delta_{12} - \theta_{12}), (\delta_{23} - \theta_{23}), ..., (\delta_{k-1k} - \theta_{k-1k})\big) \xrightarrow{D} N(0, U)$, where $\xrightarrow{D}$ denotes convergence in distribution and U is a positive definite covariance matrix of order $(k-1)\times(k-1)$ with the elements as $a_{ij} = \dfrac{\theta_i}{\tau_i \lambda_i} + \dfrac{\theta_j}{\tau_j \lambda_j}, i, j = 1, 2, ..., k-1$.

$$\begin{aligned} Var(\delta_{i,i+1}) = \gamma_{i,i+1} &= Var\left\{E\left(T_{\theta_{i,i+1}} | \left(\bar{X}, V\right)\right)\right\} \\ &= Var\left(\frac{1}{\bar{X}_{i+1}} - \frac{1}{\bar{X}_i}\right) \\ &= \frac{\theta_{i+1}}{n_{i+1}\lambda_{i+1}} + \frac{\theta_i}{n_i \lambda_i}. \end{aligned}$$

$$\text{Also, } N\xi_{i,i+1} \to \frac{\theta_i}{\tau_i\lambda_i} + \frac{\theta_{i+1}}{\tau_{i+1}\lambda_{i+1}} \text{ almost surely .} \tag{9}$$

Therefore, $\left(\frac{\delta_{12}-\theta_{12}}{\sqrt{\xi_{12}}}, \frac{\delta_{23}-\theta_{23}}{\sqrt{\xi_{23}}}, ..., \frac{\delta_{k-1,k}-\theta_{k-1,k}}{\sqrt{\xi_{k-1,k}}}\right) \xrightarrow{D} N(0, U^*)$ where the elements of U^* are $\frac{a_{ij}}{\sqrt{a_{ii}}\sqrt{a_{jj}}}$.

Let $\underline{Z} = (Z_1, Z_2, \ldots, Z_{k-1})$ be the random vector where $Z_r = \frac{\delta_{r,r+1}-\theta_{r,r+1}}{\sqrt{\xi_{r,r+1}}}, r = 1, 2, ..., k-1$ which is distributed as $N(0, U^*)$.

By the continuous mapping theorem, $\max\limits_{1\le i\le k-1}\left|\frac{\theta_{i,i+1}-\delta_{i,i+1}}{\sqrt{\xi_{i,i+1}}}\right| \xrightarrow{D} \max\limits_{1\le r\le k-1} Z_r$.

$$\sqrt{N}\left(T_{\theta_{i,i+1}} - \delta_{i,i+1}\right) = Z_{i+1}\sqrt{\frac{\theta_{i+1}}{\tau_{i+1}\lambda_{i+1}}} + Z_i\sqrt{\frac{\theta_i}{\tau_i\lambda_i}} + o_p(1). \tag{10}$$

almost surely given $\left(\bar{X}, V\right)$ for $i = 1, 2, \ldots, k-1$.
(3.6) and (3.7), implies that

$$\max_{1\le i\le k-1}\left|\frac{T_{\theta_{i,i+1}} - \delta_{i,i+1}}{\sqrt{\xi_{i,i+1}}}\right| \xrightarrow{D} \max_{1\le r\le k-1} Z_r, \tag{11}$$

almost surely given $\left(\bar{X}, V\right)$.
Assume that $F(.)$ be the cumulative distribution function of $\max\limits_{1\le r\le k-1}|Z_r|$.
Due to continuity of $F(.)$, the conditional distribution function $F_n\left(x|\left(\bar{X}, V\right)\right)$ of the left-hand side of (3.8) satisfies

$\sup\limits_x \left|F_n(x|\left(\bar{X}, V\right)) - F(x)\right| \to 0$, as $n \to \infty$.

Further, $P\left[\theta_{i,i+1} \in \delta_{i,i+1} \pm q_\alpha\sqrt{\xi_{i,i+1}} \text{ for } i = 1, 2, ..., k-1\right]$

$$= P\left\{F_n\left(\max_{1\le i\le k-1}\left|\frac{\theta_{i,i+1}-\delta_{i,i+1}}{\sqrt{\xi_{i,i+1}}}\right| | \left(\bar{X}, V\right) \le 1-\alpha\right)\right\}$$

$$= P\left\{F_n\left(\max_{1\le i\le k-1}\left|\frac{\theta_{i,i+1}-\delta_{i,i+1}}{\sqrt{\xi_{i,i+1}}}\right| + o_p(1) \le 1-\alpha\right)\right\} \xrightarrow{D} 1-\alpha$$

Corollary 3.2. *Assuming the conditions outlined in the above theorem, we have the following asymptotic statement:*
$P\left(LB_n < \mu_i - \mu_{i+1} < UB_n, i = 1, 2, ..., k-1\right) \xrightarrow{P} 1-\alpha$,
where the lower and upper bound are defined as

$$LB_n = -\bar{X}_i\bar{X}_{i+1}\left(\frac{1}{\bar{X}_i} - \frac{1}{\bar{X}_{i+1}} + q_\alpha\sqrt{\xi_{i,i+1}}\right).$$

$$UB_n = -\bar{X}_i\bar{X}_{i+1}\left(\frac{1}{\bar{X}_i} - \frac{1}{\bar{X}_{i+1}} - q_\alpha\sqrt{\xi_{i,i+1}}\right).$$

□

4. Simultion Study

Under this section, simulation is done to evaluate the probabilities of Type-I error and power of the test under different configurations of the proposed SFGCIs. We use the following procedure for computations:

(1) For a parametric configurations and sample size, generate s vectors $\left(\bar{X_1}^{(s)}, \bar{X_2}^{(s)}, ..., \bar{X_k}^{(s)} V_1^{(s)}, V_2^{(s)}, ..., V_k^{(s)}\right)$ where $\bar{X}_i \sim IG(\mu_i, \lambda_i), V_i \sim \frac{\chi^2_{(n_i-1)}}{n_i \lambda_i}, i = 1, 2, ..., k.$

(2) Calculate

$$q_{i,i+1}^{(s)} = \frac{\left| \frac{1}{\bar{X}_{i+1}^{(s)}} - \frac{1}{\bar{X}_i^{(s)}} \right|}{\sqrt{\xi_{i,i+1}^{(s)}}}, \quad i = 1, 2, ..., k-1.$$

$$q^{(s)} = \max q_{i,i+1}^{(s)}.$$

(3) For each generated vector from previous step, using algorithm, obtain the estimate of conditional upper α^{th} quantile (q_α) from distribution of $M^{'}$ with $s = 5000$.

(4) Count the number of times $q^{(s)} > q_\alpha^{(s)}$. The proportion of count among the total simulations is the simulated Type-I error probability.

The simulation study was conducted by considering parameters with various combinations: Number of groups $k = 3, 5, 7$; population location parameter $\mu^{(k)} = (\mu_1, \mu_2, ..., \mu_k)$; population scale parameter $\lambda^{(k)} = (\lambda_1, \lambda_2, ..., \lambda_k)$ with significance level $\alpha = 0.05$. Some of the specific configurations for group $k = 3$ are: sample sizes $n_1^{(3)} = (10, 20, 30)$, $n_2^{(3)} = (20, 20, 20)$, and
$n_3^{(3)} = (15, 25, 35)$; scale parameters $\lambda_1^{(3)} = (5, 11, 12)$, $\lambda_2^{(3)} = (5, 11, 13)$, and$\lambda_3^{(3)} = (6, 9, 15)$; location parameters $\mu_1^{(3)} = (1, 1, 1)$, $\mu_2^{(3)} = (2, 2, 2)$, and $\mu_3^{(3)} = (3, 3, 3)$. For $k = 5$, samples sizes $n_1^{(5)} = (10, 20, 30, 40, 50)$, $n_2^{(5)} = (20, 20, 20, 20, 20)$, and $n_3^{(5)} = (15, 25, 35, 45, 55)$; scale parameters $\lambda_1^{(5)} = (5, 11, 11, 12, 12)$, $\lambda_2^{(5)} = (5, 11, 11, 13, 13)$, and $\lambda_3^{(5)} = (6, 9, 9, 15, 15)$; location parameters $\mu_1^{(5)} = (1, 1, 1, 1, 1)$, $\mu_2^{(5)} = (2, 2, 2, 2, 2)$, and $\mu_3^{(5)} = (3, 3, 3, 3, 3)$. For $k = 7$, sample sizes $n_1^{(7)} = (10, 20, 30, 40, 50, 60, 70)$, $n_2^{(7)} = (20, 20, 20, 20, 20, 20, 20)$, and $n_3^{(7)} = (15, 20, 25, 30, 35, 40, 45)$; scale parameters $\lambda_1^{(7)} = (5, 5, 11, 11, 11, 12, 12)$,
$\lambda_2^{(7)} = (5, 5, 11, 11, 13, 13, 13)$, and $\lambda_3^{(7)} = (6, 6, 9, 9, 15, 15, 15)$; location parameters $\mu_1^{(7)} = (1, 1, 1, 1, 1, 1, 1)$, $\mu_2^{(7)} = (2, 2, 2, 2, 2, 2, 2)$, and $\mu_3^{(7)} = (3, 3, 3, 3, 3, 3, 3)$. The simulated probability of Type-I error of the proposed procedure are given in Tables 1-3. These simulated probabilities of Type-I error are near the nominal level. Simulated power of the proposed procedure is the estimate of the ratio of rejected null hypothesis when the alternative is true to the total number of simulations. The estimated power is

presented in Table 4. The result summarizes that the power of the proposed procedure improves consistently as both the sample size and scale parameter increases. Since the scale parameter is inversely related to the variability, higher values of scale parameter resulted in greater power. Comparing the estimated power of proposed procedure and all pairwise comparison procedure in Table 4 and Table 5 respectively, it is concluded that the former has higher power than later for small samples. Additionally, both procedures perform well for larger sizes of sample and higher levels of scale parameter.

Table 1. Type-I error simulated probabilities for 3 groups at $\alpha = 0.05$.

(μ, λ)		$n_1^{(3)}$	$n_2^{(3)}$	$n_3^{(3)}$
$\mu_1^{(3)}$	$\lambda_1^{(3)}$	0.0383	0.0379	0.0495
	$\lambda_2^{(3)}$	0.0413	0.0486	0.0493
	$\lambda_3^{(3)}$	0.0463	0.0491	0.0531
$\mu_2^{(3)}$	$\lambda_1^{(3)}$	0.0395	0.0372	0.0381
	$\lambda_2^{(3)}$	0.0462	0.0483	0.0411
	$\lambda_3^{(3)}$	0.0471	0.0434	0.0477
$\mu_3^{(3)}$	$\lambda_1^{(3)}$	0.0375	0.0478	0.0521
	$\lambda_2^{(3)}$	0.0425	0.0493	0.0418
	$\lambda_3^{(3)}$	0.0454	0.0501	0.0565

Table 2. Type-I error simulated probabilities for 5 groups at $\alpha = 0.05$.

(μ, λ)		$n_1^{(5)}$	$n_2^{(5)}$	$n_3^{(5)}$
$\mu_1^{(5)}$	$\lambda_1^{(5)}$	0.0396	0.0383	0.0415
	$\lambda_2^{(5)}$	0.0431	0.0413	0.0435
	$\lambda_3^{(5)}$	0.0451	0.0534	0.0444
$\mu_2^{(5)}$	$\lambda_1^{(5)}$	0.0383	0.0413	0.0398
	$\lambda_2^{(5)}$	0.0391	0.0483	0.0511
	$\lambda_3^{(5)}$	0.0493	0.0561	0.0535
$\mu_3^{(5)}$	$\lambda_1^{(5)}$	0.0378	0.0438	0.0516
	$\lambda_2^{(5)}$	0.0454	0.0485	0.0492
	$\lambda_3^{(5)}$	0.0501	0.0488	0.0577

5. Example 1

To demonstrate the applicability of the proposed test, a real dataset referenced in [**?**] is utilized. This dataset records the stay of patients with respiratory disease of different age groups in the hospital. The dataset is available on Kaggle (2023). The dataset comprises of total 17 variables among which one categorical variable 'age′ and other the metric variable 'time in hospital′ are considered for the study. The random samples of sizes $20, 30, 25, 27, 30$, and 21 are generated randomly from the increasing age groups $[40-50), [50-60), [60-70), [70-80), [80-90)$, and$[90-100)$. The dataset is presented in Table 6. To verify that the data follows Inverse Gaussian distribution, a density plot is provided, demonstrating that the data fits to the Inverse Gaussian

Table 3. Type-I error simulated probabilities for 7 groups at $\alpha = 0.05$.

(μ, λ)		$n_1^{(7)}$	$n_2^{(7)}$	$n_3^{(7)}$
$\mu_1^{(7)}$	$\lambda_1^{(7)}$	0.0461	0.0399	0.0439
	$\lambda_2^{(7)}$	0.0405	0.0458	0.0488
	$\lambda_3^{(7)}$	0.0433	0.0511	0.0581
$\mu_2^{(7)}$	$\lambda_1^{(7)}$	0.0397	0.0387	0.0401
	$\lambda_2^{(7)}$	0.0434	0.0444	0.0413
	$\lambda_3^{(7)}$	0.0490	0.0510	0.0530
$\mu_3^{(7)}$	$\lambda_1^{(7)}$	0.0337	0.0388	0.0437
	$\lambda_2^{(7)}$	0.0466	0.0491	0.0498
	$\lambda_3^{(7)}$	0.0513	0.0508	0.0585

Table 4. Estimated power of the proposed procedure with $n_1^{(3)} = (10, 10, 10)$, $n_2^{(3)} = (10, 15, 20)$, $n_3^{(3)} = (15, 15, 15)$; scale parameter $\lambda_1^{(3)} = (1, 2.5, 3.5)$, $\lambda_2^{(3)} = (1, 3, 5)$, $\lambda_3^{(3)} = (1.5, 4.5, 7.5)$; location parameter $\mu_1^{(3)} = (1, 3, 5)$, $\mu_2^{(3)} = (1, 1.5, 2.5)$, $\mu_3^{(3)} = (1, 5, 9)$ for $\alpha = 0.05$.

(μ, λ)		$n_1^{(3)}$	$n_2^{(3)}$	$n_3^{(3)}$
$\mu_1^{(3)}$	$\lambda_1^{(3)}$	0.6732	0.7155	0.9833
	$\lambda_2^{(3)}$	0.8501	0.8812	0.9932
	$\lambda_3^{(3)}$	0.9912	0.9672	0.9850
$\mu_2^{(3)}$	$\lambda_1^{(3)}$	0.8530	0.8710	0.9531
	$\lambda_2^{(3)}$	0.8233	0.9670	0.9670
	$\lambda_3^{(3)}$	0.9533	0.9881	0.9881
$\mu_3^{(3)}$	$\lambda_1^{(3)}$	0.8353	0.9720	0.9410
	$\lambda_2^{(3)}$	0.8553	0.9631	0.9510
	$\lambda_3^{(3)}$	0.9010	0.9355	0.9953

Table 5. Estimated power of the all pairwise comparison procedure with $n_1^{(3)} = (10, 10, 10)$, $n_2^{(3)} = (10, 15, 20)$, $n_3^{(3)} = (15, 15, 15)$; scale parameter $\lambda_1^{(3)} = (1, 2.5, 3.5)$, $\lambda_2^{(3)} = (1, 3, 5)$, $\lambda_3^{(3)} = (1.5, 4.5, 7.5)$; location parameter $\mu_1^{(3)} = (1, 3, 5)$, $\mu_2^{(3)} = (1, 1.5, 2.5)$, $\mu_3^{(3)} = (1, 5, 9)$ at $\alpha = 0.05$.

(μ, λ)		$n_1^{(3)}$	$n_2^{(3)}$	$n_3^{(3)}$
$\mu_1^{(3)}$	$\lambda_1^{(3)}$	0.3300	0.4750	0.4533
	$\lambda_2^{(3)}$	0.3851	0.4452	0.4232
	$\lambda_3^{(3)}$	0.9155	0.9272	0.8590
$\mu_2^{(3)}$	$\lambda_1^{(3)}$	0.0751	0.1450	0.9531
	$\lambda_2^{(3)}$	0.1153	0.1650	0.0970
	$\lambda_3^{(3)}$	0.4672	0.6155	0.1255
$\mu_3^{(3)}$	$\lambda_1^{(3)}$	0.5753	0.5551	0.5734
	$\lambda_2^{(3)}$	0.5733	0.6341	0.5954
	$\lambda_3^{(3)}$	0.9701	0.9753	0.9652

distribution. Figure 2 presents the box-plot of the sampled observations where on x-axis $g1, g2, g3, g4, g5, \text{and} g6$ corresponds to the ordered age groups and y-axis represents

the stay of patients with respiratory disease of a age group in the hospital. The plot indicates that the data within each age group exhibits a skewed distribution.

Here, the hypothesis is $H_{oi} : \mu_{i+1} = \mu_i$ where μ_i is the average number of days spent

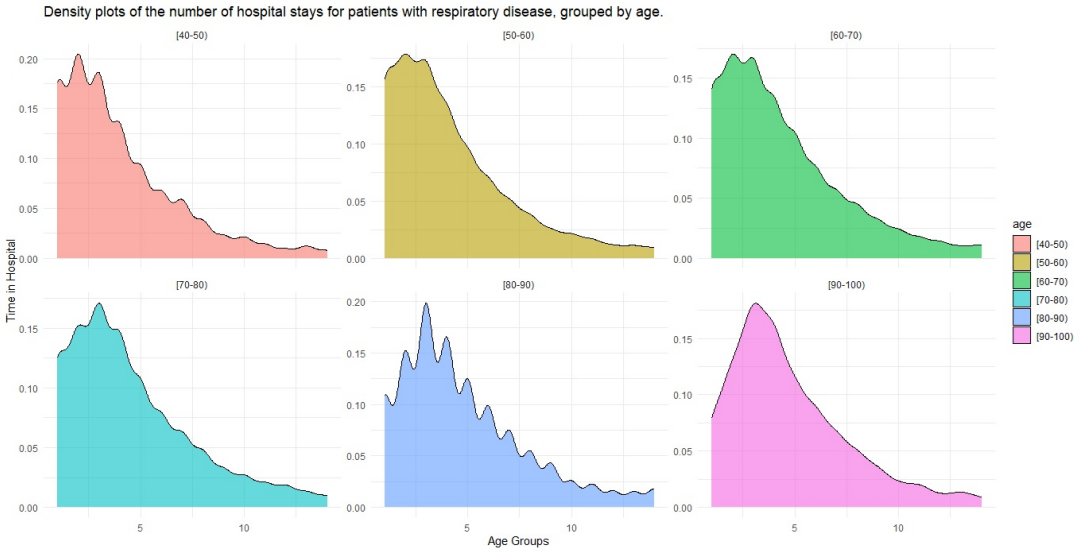

Figure 1. Density plot of population dataset of stays of the patients with respiratory disease in the hospital of different age groups.

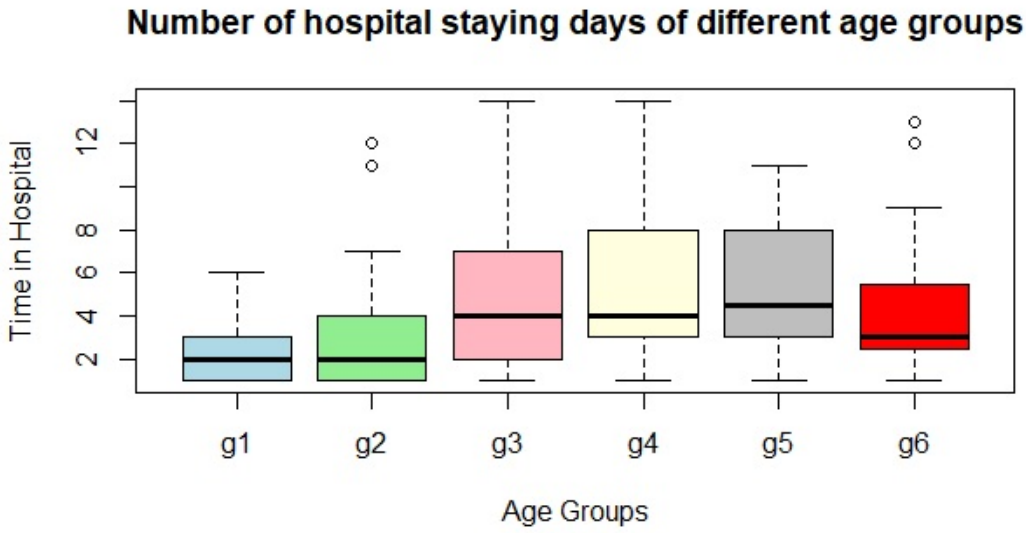

Figure 2. Box and whisker plot of number of hospital staying days of different age groups.

in hospital by i^{th} age group against $H_{1i} : \mu_{i+1} \geq \mu_i, i = 1, 2, \ldots, k-1$. Using the proposed FGPQs, SFGCIs are obtained for the successive differences of the ordered location parameters when scale parameters are unequal and not known.

According to the calculations, the value of the statistics are: $X_1 = 2.4, X_2 = 3.2, X_3 = 4.640, X_4 = 5.821, X_5 = 5.100, X_6 = 4.387 ; V_1 = 2.253, V_2 = 8.234, V_3 = 11.49, V_4 \equiv 16.819, V_5 = 7.9551, V_6 = 9.1118;\ \xi_{12} = 0.1505, \xi_{23} = 0.2079, \xi_{34} = 02281, \xi_{45} =$

Table 6. Records of stay of the patients with respiratory disease in the hospital across different age groups.

Age Groups	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
$[40, 50)$	3	4	3	5	2	1	1	2	3	1	1	3	1	1	1	6	1	3	2	4											
$[50, 60)$	1	2	3	1	6	7	2	2	1	2	4	2	11	2	2	4	2	1	7	1	1	1	4	4	1	4	1	12	4		
$[60, 70)$	4	2	5	9	6	3	10	14	1	1	7	1	2	5	1	1	7	7	3	6	4	4	2	2	9						
$[70, 80)$	2	5	2	14	1	10	7	3	2	2	3	4	4	4	2	8	4	13	14	4	11	5	8	3	6	14	5	3			
$[80, 90)$	8	5	4	2	1	3	7	11	3	3	9	4	5	8	6	3	7	3	8	2	3	6	5	2	8	11	4	8	2	2	
$[90, 100)$	3	3	9	2	3	3	7	1	7	2	6	2	4	1	4	3	9	13	3	4	12	6	9	3	4	5	5	3	1	4	2

$0.1733, \xi_{56} = 0.1319$ based on the sample size $n_1 = 20, n_2 = 29, n_3 = 25, n_4 = 28, n_5 = 31, n_6 = 21$. Using the algorithm, the 95^{th} percentile is obtained as $q_\alpha = 4.0464$. Based on equation (6), the SFGCIs are computed as $\mu_2 - \mu_1 \in (-1.4657, 1.6740), \mu_3 - \mu_2 \in (-1.7478, 1.9418), \mu_4 - \mu_3 \in (-1.8889, 1.9763), \mu_5 - \mu_4 \in (-1.7089, 1.6603), \mu_6 - \mu_5 \in (-1.5017, 1.4379)$. Since each interval includes 0, it is concluded that number of days patients in hospital stays of different age group are not different among the various age groups at 5% level of significance. As no significant differences are found among the increasing age groups, the wide interval may imply that the study may not have sufficient power to detect small but potentially meaningful differences.

6. Example 2

This example refers to the environmental area, focussing on the PM 2.5 concerntartions in various northeastern provinves of thailand. The dataset referenced in Chankham et al. (2024) is considered for this example. The dataset provides the mean PM 2.5

Table 7. Daily PM 2.5 concerntrations of various provinces of Thailand.

Provinces	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	21	22	23
NongKhai	12.4	10.5	12.4	19.3	15.4	10.1	6.4	6	6.5	18.1	18.5	20.7	14.3	23	15.4	22.5	31.2	40.5	44.9	33.5	24.3	27.6
UbonRatchathani	15.9	13.9	9.2	8.8	11.1	8.4	5	6	8.9	13.2	7.4	7.6	6.9	9.9	17.3	28.9	36	38.6	43.1	39.1		
NakhonRatchasima	13.5	15.8	15.2	10.1	11.7	9.9	10.3	11.5	14.3	15	12.3	12.9	14.8	4	19.1	27.6	31.7	36.6				

concerntration from October and December 2023 for three northeastern provinces of Thailand. The provinces serve as a categorical varible ordered by increasing altitude: Nong Khai with altitude 120 mts above sea level, Ubon Ratchathani with altitude 125 mts above sea level, and Nakhon Ratchasima with altitude 150 mts above sea level. To access the fit of Inverse Gaussian distribution, box and whisker plot is presented in Figure 3, showing that the data are positively skewed. The dataset have two variables, Provinces and PM 2.5 concerntrations, given in Table 7. The random sample of sizes $n_1 = 23$, $n_2 = 21$, $n_3 = 18$ are generated randomly from each province. The hypothesis is whether the mean PM 2.5 concerntration is equal for all the provinces at different level of altitude that is, $H_{oi} : \mu_i = \mu_{i+1}$ where μ_i is PM 2.5 concerntration of i^{th} province against $H_1 : \mu_{i+1} \geq \mu_i, i = 1, 2, 3$.

For the SFGCIs, the values of the statistics are: $\bar{X}_1 = 17.0357, \bar{X}_2 = 14.0555, \bar{X}_3 = 16.4611$; $V_1 = 41.9619, V_2 = 36.0576, V_3 = 111.7516$; $\xi_{12} = 0.3403, \xi_{23} = 0.5502$ based on the sample size $n_1 = 23, n_2 = 21, n_3 = 18$. Using the algorithm, the 95^{th} percentile is obtained as $q_\alpha = 37.7$. Based on the equation (3.4), the SFGCIs are computed as $\mu_2 - \mu_1 \in (-24.9747, 19.0144), \mu_3 - \mu_2 \in (-25.5596, 30.3707)$. In both the intervals, zero is included. Therefore, we conclude that the PM 2.5 concerntrations donot differ significantly with increasing altitude across three provinces.

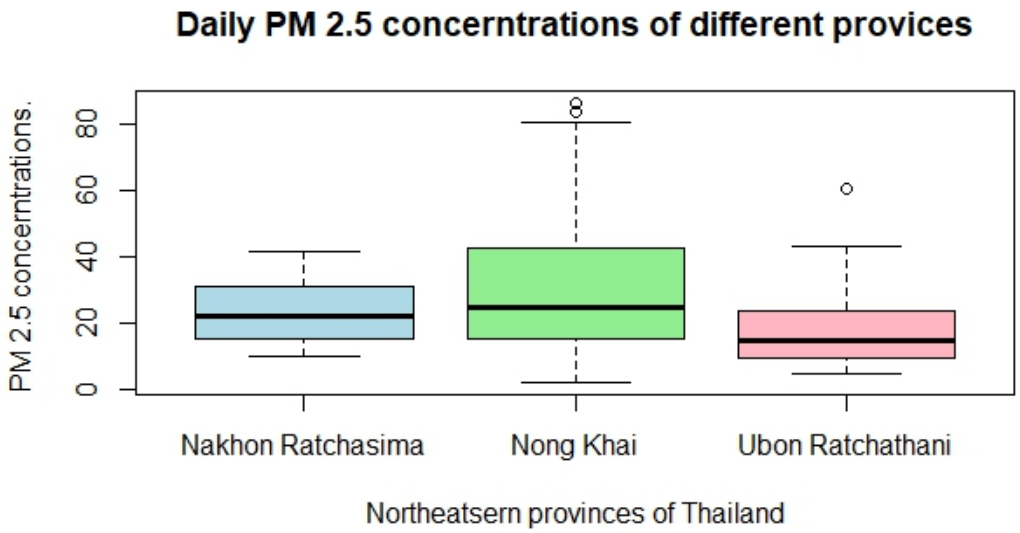

Figure 3. Box and whisker plot for daily PM 2.5 concerntrations in northeastern provinces of Thailand at different sea level.

7. Conclusion

In this paper, we developed a method for the construction of SFGCIs for the successive differences of ordered location parameters of more than two Inverse Gaussian distribution on the basis of FGPQs. The simulation results demonstrate that the proposed framework effectively controls probability of Type-I error and achieves strong power. By the reduced number of comparisons in successive differences, the method ensures probability of Type-I error is closer to the nominal level. Furthermore, the estimated power shows that larger sample sizes and higher scale parameters leads to enhanced statistical power. Also, on comparison with the all pairwise comparison procedure, the proposed procedure shows higher power for small samples. Both procedures perform well for higher values of scale parameters and sample sizes. The practical utility of the procedure is illustrated through a two data set.

Acknowledgement

The authors sincerely appreciate the insightful comments and suggestions of the reviewers, which have greatly enhanced the quality and clarity of the manuscript.

References

[1] Chankham, W., Niwitpong, S.-A. and Niwitpong, S. (2024), 'Estimating the confidence interval for the common coefficient of variation for multiple inverse gaussian distributions", Symmetry 16(7).

[2] Chhikara, R. (2024), The inverse Gaussian distribution: theory: methodology, and applications, CRC Press.

[3] Hannig, J., Iyer, H. and Patterson, P. (2006), 'Fiducial generalized confidence intervals', Journal of the American Statistical Association 101(473), 254-269.
[4] Hannig, J., Lidong, E., Abdel-Karim, A. and Iyer, H. (2006), 'Simultaneous fiducial generalized confidence intervals for ratios of means of lognormal distributions', Austrian Journal of Statistics 35(2&3), 261-269.
[5] Kaggle (2023), 'Predicting hospital readmissions'. URL: https://www:kaggle.com/datasets/dubradave/hospital-readmissions?resource=download
[6] Kharrati-Kopaei, M. (2015), 'A note on the simultaneous confidence intervals for the successive differences of exponential location parameters under heteroscedasticity', Statistical Methodology 22, 1-7.
[7] Kharrati-Kopaei, M. and Eftekhar, S. (2017), 'Simultaneous confidence intervals for comparing several inverse gaussian means under heteroscedasticity', Journal of Statistical Computation and Simulation 87(4), 777-790.
[8] Kharrati-Kopaei, M., Malekzadeh, A. and Sadooghi-Alvandi, M. (2013), 'Simultaneous fiducial generalized confidence intervals for the successive differences of exponential location parameters under heteroscedasticity', Statistics & Probability Letters 83(6), 1547-1552.
[9] Ma, C.-X. and Tian, L. (2009), 'A parametric bootstrap approach for testing equality of inverse gaussian means under heterogeneity', Communications in Statistics-Simulation and Computation 38(6), 1153-1160.
[10] Mondal, A. and Kumar, S. (2024), 'Inference on order restricted means of inverse gaussian populations under heteroscedasticity', Computational Statistics & Data Analysis 194, 107943.
[11] Seshadri, V. (1994), The imerse Gaussian distribution: a case study in exponential families, Oxford University Press.
[12] Shi, J.-H. and Lv, J.-L. (2012), 'A new generalized p-value for testing equality of inverse gaussian means under heterogeneity", Statistics & probability letters 82(1), 96-102.
[13] Tian, L. (2006), "Testing equality of inverse gaussian means under heterogeneity, based on generalized test variable', Computational statistics & data analysis 51(2), 1156-1162.
[14] Tweedie, M. C. (1957), 'Statistical properties of inverse gaussian distributions, i', The Amals of Mathematical Statistics 28(2), 362-377.
[15] Weerahandi, S. (1993), 'Generalized confidence intervals', Journal of the American Statistical Association 88(423), 899-905.
[16] Xiong, S. and Mu. W. (2009), 'Simultancous confidence intervals for one-way layout based on generalized pivotal quantities', Journal of Statistical Computation and Simulation 79(10), 1235-1244.
[17] Ye, R.-D., Ma, T.-F. and Wang. S.-G. (2010), 'Inferences on the common mean of several inverse gaussian populations', Computational Statistics & Data Analysis 54(4), 906-915.
[18] Thang, G. (2014), 'Simultaneous confidence intervals for several inverse gaussian populations', Statistics & Probability Letters 92, 125-131.