Communications of the Korean Statistical Society Vol. 15, No. 5, 2008, pp. 655–666

Likelihood Based Inference for the Shape Parameter of the Inverse Gaussian Distribution

Woo Dong Lee[1), Sang Gil Kang[2), Dongseok Kim[3)

Abstract

Small sample likelihood based inference for the shape parameter of the inverse Gaussian distribution is the purpose of this paper. When shape parameter is of interest, the signed log-likelihood ratio statistic and the modified signed log-likelihood ratio statistic are derived. Hsieh (1990) gave a statistical inference for the shape parameter based on an exact method. Throughout simulation, we will compare the statistical properties of the proposed statistics to the statistic given by Hsieh (1990) in term of confidence interval and power of test. We also discuss a real data example.

Keywords: Inverse Gaussian distribution; shape parameter; signed log-likelihood statistics; modified signed log-likelihood ratio statistics.

1. Introduction

The inverse Gaussian distribution has its origin in the Wiener process as the first passage time distribution. It is also an approximation to the sample size distribution in a sequential probability ratio test. For these reasons, the inverse Gaussian distribution is also known as the first passage time or the Wald distribution.

This distribution has potentially useful applications in a wide variety of fields such as biology, economics, reliability theory, life testing and social sciences as discussed in Folks and Chhikara (1978), Chhikara and Folks (1989), Whitmore (1979), Seshadri (1999) and Mudholkar and Natarajan (2002). Tweedie (1957a, 1957b) studied many important statistical properties of this distribution and discussed the similarity between statistical methods based on the inverse Gaussian distribution and those based on the normal distribution. A comprehensive discussion on the inverse Gaussian distributions can be found in books of Chhikara and Folks (1989) and Seshadri (1999).

From a reliability point of view, Chhikara and Folks (1977) showed that if the lifetime of a machine has the inverse Gaussian distribution and the shape parameter is less than 2 and given that the machine has survived up to time t_0(a known value), then the mean residual time will eventually exceed the original mean lifetime. In practice, the shape

1) Department of Asset Management, Daegu Haany University, Kyungsan 712-240, Korea.
E-mail: wdlee@dhu.ac.kr
2) Department of Computer and Data Information, Sangji University, Wonju 220-702, Korea.
Correspondence: sangkg@sangji.ac.kr
3) Department of Mathematics, Kyonggi University, Suwon 443-760, Korea.
E-mail: dongseok@kgu.ac.kr

parameter is typically unknown. So, one may want to estimate it or perform a statistical test for the shape parameter.

Related with this problem, Kang *et al.* (2004) developed a Bayesian inference for the shape parameter using noninformative priors. Hsieh (1990) derived an exact test for the shape parameter of the inverse Gaussian distribution. He showed that, through simulation, the confidence intervals proposed by Folks and Chhikara (1978) did not achieve nominal levels. He also gave some useful quantiles of test statistics for small to moderate sample sizes of odd numbers with respect to preassigned values of the shape parameter. However, one may be interested in testing shape parameters with some other values given in his Table or even sample size. In these cases, one must compute along with the method given in Hsieh (1990). However, it is not a convenient way for statistical inferences in practice.

Instead of using his Table, if there exists a highly accurate statistic which distributes as a well-known distribution such as the standard normal, it is convenient to perform a test or to construct a confidence interval. There are two famous such statistics based on likelihood-based methods, the signed log-likelihood ratio test and the modified signed log-likelihood ratio test, found in Barndorff-Nielsen and Cox (1994). These two statistics distribute as asymptotically standard normal distributions. Specially, the modified signed log-likelihood statistic can be applied for a small sample and moderate sample sizes.

In this paper, we propose the signed log-likelihood ratio test and the modified signed log-likelihood ratio test for testing the shape parameter of the inverse Gaussian distribution. Using these statistics, we will compare exact test and two likelihood-based tests in terms of the coverage probability and the expected length of confidence intervals, Type I error and the power of test statistics.

This article is organized as follows. In Section 2, we introduce the likelihood-based inference methods, which will employ to test the hypothesis of the shape parameter and to construct the confidence intervals for the shape parameters. In Section 3, we show some simulation results to demonstrate the accuracy of the proposed methods and Hsieh (1990). In Section 4, a real data example is examined and simulation results are provided. At last, some final remarks are recorded.

2. Likelihood-Based Methods

The probability density function(p.d.f.) of the two parameter inverse Gaussian distribution is given by

$$f(x \mid \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi}} x^{-\frac{3}{2}} \exp\left\{ -\frac{\lambda(x-\mu)^2}{2\mu^2 x} \right\}, \quad \lambda,\ \mu > 0,\ x > 0. \tag{2.1}$$

Let the notation $X \sim \mathrm{IG}(\mu, \lambda)$ denote that the random variable X has an inverse Gaussian distribution with *p.d.f.* given (2.1). As a note, the mean and variance of X are μ and μ^3/λ, respectively. Let

$$\theta_1 = \frac{\lambda}{\mu}$$

be the parameter of interest. This parameter θ_1 determines the shape of the distribution and the density function is highly skewed for moderate values of θ_1. As θ_1 increases the inverse Gaussian tends towards the normal law. The coefficient of variation, the skewness and the kurtosis of X are closely related with θ_1. These values are $\theta_1^{-1/2}$, $3\theta_1^{-1/2}$ and $15\theta_1^{-1}$, respectively. So, once we construct the confidence interval for θ_1, one can easily construct the confidence interval for the coefficient of variation, the skewness and the kurtosis, since these functions are monotone functions of θ_1.

Let $X_1, X_2, \ldots, X_n$ be the random sample from *p.d.f.* (2.1). Then the likelihood function of μ and λ given observations $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is given by

$$L(\mu, \lambda \,|\, \mathbf{x}) = (2\pi)^{-\frac{n}{2}} \lambda^{\frac{n}{2}} \left(\prod_{i=1}^{n} x_i^{-\frac{3}{2}} \right) \exp\left(-\frac{\lambda}{2\mu^2} t_1 + \frac{n\lambda}{\mu} - \frac{\lambda}{2} t_2 \right), \tag{2.2}$$

where $t_1 = \sum_{i=1}^{n} x_i$ and $t_2 = \sum_{i=1}^{n} x_i^{-1}$. It can be easily verified that the maximum likelihood estimates(m.l.e.) of μ and λ are given by

$$\widehat{\mu} = \frac{t_1}{n} \quad \text{and} \quad \widehat{\lambda} = \frac{n}{t_2 - \dfrac{n^2}{t_1}},$$

respectively. Let

$$\theta_1 = \frac{\lambda}{\mu} \quad \text{and} \quad \theta_2 = \frac{2}{\mu} + \frac{1}{\lambda}.$$

Here θ_1 is the parameter of interest and θ_2 is nuisance parameter. This transformation is actually an orthogonal transformation, *i.e.*, the information matrix is diagonal, in the sense of Cox and Reid (1987). Though there may exist other transformations, we prefer to use it, because the constraint *m.l.e.* for nuisance parameter can be obtained explicitly. Under this transformation, the log-likelihood function of $\theta = (\theta_1, \theta_2)$ given $\mathbf{x}$ is

$$\begin{aligned} l(\theta_1, \theta_2) = &-\frac{n}{2}\log(2\pi) + \frac{n}{2}\log(2\theta_1 + 1) - \frac{n}{2}\log(\theta_2) - \frac{3}{2}\sum_{i=1}^{n}\log(x_i) \\ &- \frac{\theta_1^2\theta_2}{2(2\theta_1 + 1)} t_1 + n\theta_1 - \frac{2\theta_1 + 1}{2\theta_2} t_2. \end{aligned} \tag{2.3}$$

Now, one can make an inference about θ_1 based on the signed log-likelihood ratio statistic using (2.3).

$$r \equiv r(\theta_1) = \operatorname{sgn}\left(\widehat{\theta_1} - \theta_1\right) \left[2 \left\{ l\left(\widehat{\theta_1}, \widehat{\theta_2}\right) - l\left(\theta_1, \widehat{\theta}_{2(\theta_1)}\right) \right\} \right]^{\frac{1}{2}}, \tag{2.4}$$

where $\widehat{\theta} = (\widehat{\theta_1}, \widehat{\theta_2})$ denote *m.l.e.* of $\theta = (\theta_1, \theta_2)$ and $\widehat{\theta}_{2(\theta_1)}$ denotes the constrained *m.l.e.* of θ_2 for a fixed θ_1. This constrained *m.l.e.* of θ_2 can be obtained by solving the following equation.

$$\frac{\partial l(\theta_1, \theta_2)}{\partial \theta_2} = -\frac{n}{2\theta_2} - \frac{\theta_1^2 t_1}{2(2\theta_1 + 1)} + \frac{(2\theta_1 + 1) t_2}{2\theta_2^2} = 0.$$

The solution of the above equation for θ_2, the constrained *m.l.e.* of θ_2 is given by

$$\widehat{\theta}_{2(\theta_1)} = \frac{n(2\theta_1 + 1)\left(\sqrt{1 + 4\theta_1^2 \dfrac{t_1 t_2}{n^2}} - 1 \right)}{2\theta_1^2 t_1}. \tag{2.5}$$

It is well known that r is approximately distributed as a standard normal to the first order, $O(n^{-1/2})$. For testing the null hypothesis $H_0 : \theta_1 = \theta_{10}$, a two-sided p-value can be obtained from

$$p\text{-value} = 2P(r > |r_0|) \approx 2\{1 - \Phi(|r_0|)\}, \tag{2.6}$$

where $\Phi(\cdot)$ is the distribution function(d.f.) of the standard normal distribution and $r_0 = r(\theta_{10})$ is the observed value of r under H_0. Furthermore, the approximate $100(1-\alpha)\%$ confidence interval for θ_1 can be obtained from

$$\left\{\theta_1 : |r(\theta_1)| \le z_{\frac{\alpha}{2}}\right\}, \tag{2.7}$$

where $z_{\alpha/2}$ is the $100(1-\alpha)\%$ percentile of the standard normal distribution.

In general, the first order approximation is not accurate especially when the sample size is small. There are several ways to improve the accuracy of the approximation by adjusting the signed log-likelihood statistic r. Among the others approximations, the modified signed log-likelihood ratio statistics, r^*, developed by Barndorff-Nielsen (1986) and Barndorff-Nielsen (1991), is a highly accurate approximation. In fact, this is the third order approximation and is given by

$$r^* \equiv r^*(\theta_1) = r(\theta_1) + r(\theta_1)^{-1} \log\left\{\frac{u(\theta_1)}{r(\theta_1)}\right\}, \tag{2.8}$$

where $u(\theta_1)$ is a statistic given by

$$u(\theta_1) = \frac{\left| l_{;\widehat{\theta}}\left(\widehat{\theta}_1, \widehat{\theta}_2\right) - l_{;\widehat{\theta}}\left(\theta_1, \widehat{\theta}_{2(\theta_1)}\right) l_{\theta_2;\widehat{\theta}}\left(\theta_1, \widehat{\theta}_{2(\theta_1)}\right)\right|}{\left\{\left|j_{\theta\theta}\left(\widehat{\theta}_1, \widehat{\theta}_2\right)\right| \left|j_{\theta_2\theta_2}\left(\theta_1, \widehat{\theta}_{2(\theta_1)}\right)\right|\right\}^{\frac{1}{2}}}, \tag{2.9}$$

where the sample-space derivatives in (2.9) are defined by

$$l_{;\widehat{\theta}}(\theta_1, \theta_2) = \frac{\partial}{\partial\widehat{\theta}} l\left(\theta_1, \theta_2; \widehat{\theta}\right),$$

the mixed derivatives as

$$l_{\theta_2;\widehat{\theta}}(\theta_1, \theta_2) = \frac{\partial}{\partial\theta_2} l_{;\widehat{\theta}}(\theta_1, \theta_2)$$

and $j_{\theta_1\theta_2}(\widehat{\theta}_1, \widehat{\theta}_2)$ is the observed information matrix and $j_{\theta_2\theta_2}(\theta_1, \widehat{\theta}_{2(\theta_1)})$ is the observed nuisance information matrix (Barndorff-Nielsen, 1991).

This statistic r^* also distributes as approximately standard normal distribution with error $O(n^{-3/2})$. So, one can obtain the p-value of testing $H_0 : \theta_1 = \theta_{10}$ based on r^* as

$$p\text{-value} = 2P(r^* > |r_0^*|) \approx 2\{1 - \Phi(|r_0^*|)\}, \tag{2.10}$$

where $r_0^* = r^*(\theta_{10})$ and the approximate $100(1-\alpha)\%$ confidence interval can be derived from

$$\left\{\theta_1 : |r^*(\theta_1)| \le z_{\frac{\alpha}{2}}\right\}, \tag{2.11}$$

which will be more accurate than (2.6) and (2.7).

Since the inverse Gaussian distribution is a full rank exponential model, the log-likelihood based on sample data is only related to a minimum sufficient statistics $t = (t_1, t_2)$ given in (2.2), and there is an one-to-one correspondence between the *m.l.e.* $\widehat{\theta}$ and t, the sample space derivatives with respect to $\widehat{\theta}$ in (2.9) can be derived based on the sample space derivatives with respect to t. In this transformation, the Jacobian matrix is $\partial\widehat{\theta}/\partial t$. Using the identity $j_{\theta\theta}(\widehat{\theta}) = l_{\theta;\widehat{\theta}}(\widehat{\theta})$ and canceling the determinant of the transformation Jacobian matrix, one can show that u can be rewritten as

$$u(\theta_1) = \frac{\left| l_{;t}\left(\widehat{\theta}_1, \widehat{\theta}_2\right) - l_{;t}\left(\theta_1, \widehat{\theta}_{2(\theta_1)}\right) \; l_{\theta_2;t}\left(\theta_2, \widehat{\theta}_{2(\theta_1)}\right) \right|}{\left| l_{\theta;t}\left(\widehat{\theta}_1, \widehat{\theta}_2\right) \right|} \left\{ \frac{\left| j_{\theta\theta}\left(\widehat{\theta}_1, \widehat{\theta}_2\right) \right|}{\left| j_{\theta_2\theta_2}\left(\theta_1, \widehat{\theta}_{2(\theta_1)}\right) \right|} \right\}^{\frac{1}{2}},$$

where the sample space derivatives $l_{;t}(\theta)$ and the mixed derivatives $l_{\theta_2;t}(\theta) = \partial^2 l(\theta)/(\partial\theta_2 \partial t)$ are given by

$$l_{;t}(\theta_1, \theta_2) = \left(-\frac{\theta_1^2\theta_2}{2(2\theta_1 + 1)} \quad -\frac{2\theta_1 + 1}{2\theta_2} \right)'$$

and

$$l_{\theta_2;t}(\theta_1, \theta_2) = \left(-\frac{\theta_1^2}{2(2\theta_1 + 1)} \quad -\frac{2\theta_1 + 1}{2\theta_2^2} \right)',$$

respectively. The observed information matrix is given by

$$j_{\theta\theta}(\theta_1, \theta_2) = \begin{bmatrix} -\dfrac{2n(2\theta_1 + 1) + \theta_2 t_1}{(2\theta_1 + 1)^3} & -\dfrac{\theta_1(\theta_1 + 1)\theta_2^1 t_1 - (2\theta_1 + 1)^2 t_2}{(2\theta_1 + 1)^2\theta_2^2} \\ \text{sym} & \dfrac{n\theta_2 - 2(2\theta_1 + 1)t_2}{2\theta_2^3} \end{bmatrix},$$

and the mixed derivative matrix $l_{\theta;t}(\theta_1,\, \theta_2)$ is given by

$$l_{\theta;t}(\theta_1,\, \theta_2) = \begin{bmatrix} -\dfrac{\theta_1\theta_2(\theta_1 + 1)}{(2\theta_1 + 1)^2} & -\dfrac{1}{\theta_2} \\ -\dfrac{\theta_1^2}{2(2\theta_1 + 1)} & \dfrac{2\theta_1 + 1}{2\theta_2^2} \end{bmatrix}.$$

Finally, the observed nuisance information matrix is given by

$$j_{\theta_2\theta_2}(\theta_1,\, \theta_2) = \frac{n\theta_2 - 2(2\theta_1 + 1)t_2}{2\theta_2^3}.$$

Using the above results, we can calculate the following determinants.

$$\left| l_{;t}\left(\widehat{\theta}_1,\, \widehat{\theta}_2\right) - l_{;t}\left(\theta_1,\, \widehat{\theta}_{2(\theta_1)}\right) \; l_{\theta_2;t}\left(\theta_2,\, \widehat{\theta}_{2(\theta_1)}\right) \right| = \frac{2\theta_1^2}{4\widehat{\theta}_{2(\theta_1)}} - \frac{(2\theta_1 + 1)\widehat{\theta}_1^{\,2}\widehat{\theta}_2}{4\left[\widehat{\theta}_{2(\theta_1)}\right]^2 \left(2\widehat{\theta}_1 + 1\right)} - \frac{\theta_1^2\left(2\widehat{\theta}_1 + 1\right)}{4\widehat{\theta}_2(2\theta_1 + 1)},$$

$$\left| l_{\theta;t}\left(\widehat{\theta}_1,\, \widehat{\theta}_2\right) \right| = -\frac{\widehat{\theta}_1}{2\widehat{\theta}_2},$$

$$\left|j_{\theta\theta}\left(\widehat{\theta}_1,\widehat{\theta}_2\right)\right| = -\frac{\left(2\widehat{\theta}_1+1\right)\widehat{\theta}_2\left\{2n\left(2\widehat{\theta}_1+1\right)+\widehat{\theta}_2 t_1\right\}\left\{n\widehat{\theta}_2-2\left(2\widehat{\theta}_1+1\right)t_2\right\}}{2\left(2\widehat{\theta}_1+1\right)^4\widehat{\theta}_2^{\,4}}$$
$$-\frac{2\left\{\widehat{\theta}_1\left(\widehat{\theta}_1+1\right)\widehat{\theta}_2^2 t_1-\left(2\widehat{\theta}_1+1\right)^2 t_2\right\}^2}{2\left(2\widehat{\theta}_1+1\right)^4\widehat{\theta}_2^{\,4}},$$

and

$$\left|j_{\theta_2\theta_2}\left(\theta_1,\widehat{\theta}_{2(\theta_1)}\right)\right| = \frac{n\widehat{\theta}_{2(\theta_1)}-2(2\theta_1+1)t_2}{2\left[\widehat{\theta}_{2(\theta_1)}\right]^3}.$$

Hence, we are ready to make a statistical inference of θ_1 based on r and r^*.

Hsieh (1990) calculated the distribution function of $W = (\overline{X}V)^{-1}$, where $\overline{X} = \sum_{i=1}^n X_i/n$ and $V = \sum_{i=1}^n X_i^{-1}/n - n/\sum_{i=1}^n X_i$, using the negative moments of the inverse Gaussian distribution. The distribution function F of W is given by

$$F(w)=\frac{e^{n\theta_1(1-\sqrt{1+1/w})}}{\sqrt{1+1/w}}\sum_{j=0}^{k-1}\left\{\frac{(n\theta_1)^j}{j!\left[2w\sqrt{1+1/w}\right]^j}\sum_{s=0}^{j}\frac{(j+s)!}{s!\,(j-s)!\left[2(n\theta_1)\sqrt{1+1/w}\right]^s}\right\}, \quad (2.12)$$

where $k=(n-1)/2$.

He reported the quantiles of W from Table I to Table IV with several values of θ_1 and odd sample size. When the sample size is even, he suggested the use of average of neighboring odd numbers. Indeed, one can make an exact statistical inference using these Tables. But it is still inconvenient for statistical inferences with another values of parameter and sample size.

3. Simulations and Real Example

In this Section, we want to compare the coverage probabilities, expected lengths of the confidence intervals and power of the test based on two likelihood-based methods and exact method by Hsieh (1990).

We take 10,000 independent random samples from inverse Gaussian distribution. We assume $\theta_1 = 0.5, 1, 2, 5$ and $\theta_2 = 0.5, 1, 2, 3$ and construct confidence intervals with $\alpha = 0.1$. We also assume the sample size as 5, 7, 9, 11, 15, 25, 35 on the purpose of comparisons with Hsieh (1990).

In simulation results, the values of nuisance parameter do not affect the coverage probabilities and expected lengths of two likelihood-based methods. Table 3.1 reports estimated coverage probabilities and expected length of three methods. In Table 3.1, r_L and r_U, r_L^* and r_U^* and F_L and F_U are lower and upper coverage probabilities of r, r^* and exact method, respectively. length_r, length_r^* and length_F are estimated expected lengths of r, r^* and exact method, respectively.

From Table 3.1, we can observe that

Table 3.1: Estimated coverage probabilities and expected length

$\theta_1 = 0.5$

	n	r_L	r_U	r_L^*	r_U^*	F_L	F_U	length$_r$	length$_r^*$	length$_F$
$\theta_2 = 0.5$, $\mu = 8$, $\lambda = 4$	5	.7646	.0000	.2232	.0000	.2052	.0000	9.53855	7.91468	7.81395
	9	.2526	.0000	.1089	.0000	.1050	.0000	2.57988	2.33488	2.32659
	11	.1726	.0000	.0786	.0000	.0764	.0000	1.82367	1.68326	1.67901
	15	.1193	.0000	.0579	.0000	.0570	.0000	1.21871	1.15116	1.14929
	20	.0993	.0000	.0532	.0108	.0398	.0249	.92372	.88579	.86098
	25	.0864	.0232	.0495	.0465	.0495	.0466	.77315	.74789	.74737
	35	.0846	.0279	.0522	.0516	.0521	.0517	.62471	.61012	.60990
$\theta_2 = 1$, $\mu = 4$, $\lambda = 2$	5	.7624	.0000	.2177	.0000	.1995	.0000	9.45937	7.84896	7.74912
	9	.2556	.0000	.1088	.0000	.1058	.0000	2.58117	2.33602	2.32774
	11	.1802	.0000	.0778	.0000	.0765	.0000	1.83366	1.69246	1.68818
	15	.1218	.0000	.0610	.0000	.0605	.0000	1.22299	1.15518	1.15331
	20	.1011	.0000	.0597	.0093	.0443	.0213	.92576	.88775	.86288
	25	.0890	.0243	.0540	.0499	.0540	.0501	.77506	.74973	.74920
	35	.0788	.0290	.0508	.0470	.0507	.0470	.62401	.60943	.60921
$\theta_2 = 2$, $\mu = 2$, $\lambda = 1$	5	.7712	.0000	.2176	.0000	.1978	.0000	9.48158	7.86738	7.76729
	9	.2505	.0000	.1062	.0000	.1041	.0000	2.59064	2.34454	2.33624
	11	.1810	.0000	.0862	.0000	.0847	.0000	1.85309	1.71033	1.70601
	15	.1235	.0000	.0589	.0000	.0581	.0000	1.21978	1.15218	1.15030
	20	.0999	.0000	.0548	.0098	.0406	.0228	.92351	.88561	.86080
	25	.0867	.0269	.0500	.0508	.0498	.0508	.76932	.74421	.74368
	35	.0831	.0297	.0521	.0496	.0518	.0497	.62399	.60941	.60919
$\theta_2 = 3$, $\mu = 1.33$, $\lambda = 0.67$	5	.7599	.0000	.2181	.0000	.1956	.0000	9.38378	7.78623	7.68722
	9	.2486	.0000	.1007	.0000	.0976	.0000	2.55572	2.31305	2.30485
	11	.1794	.0000	.0821	.0000	.0802	.0000	1.84514	1.70300	1.69870
	15	.1214	.0000	.0585	.0000	.0578	.0000	1.21730	1.14984	1.14796
	20	.0977	.0000	.0529	.0117	.0398	.0252	.92348	.88557	.86077
	25	.0890	.0251	.0514	.0491	.0511	.0492	.77545	.75010	.74957
	35	.0806	.0311	.0500	.0535	.0495	.0535	.62035	.60587	.60566

$\theta_1 = 1$

	n	r_L	r_U	r_L^*	r_U^*	F_L	F_U	length$_r$	length$_r^*$	length$_F$
$\theta_2 = 0.5$, $\mu = 6$, $\lambda = 6$	5	.3649	.0000	.0829	.0000	.0717	.0000	9.94939	8.25565	8.15029
	9	.1409	.0000	.0584	.0000	.0569	.0000	3.28010	2.96656	2.95620
	11	.1131	.0000	.0502	.0035	.0497	.0045	2.51062	2.31440	2.30900
	15	.0982	.0231	.0527	.0504	.0525	.0507	1.89642	1.78764	1.78540
	20	.0926	.0297	.0503	.0540	.0365	.0670	1.53642	1.47020	1.42919
	25	.0871	.0295	.0526	.0491	.0524	.0492	1.32413	1.27842	1.27785
	35	.0783	.0324	.0507	.0517	.0505	.0518	1.06725	1.04090	1.04067
$\theta_2 = 1$, $\mu = 3$, $\lambda = 3$	5	.3654	.0000	.0826	.0000	.0730	.0000	9.93596	8.24451	8.13930
	9	.1353	.0000	.0550	.0000	.0537	.0000	3.21985	2.91219	2.90201
	11	.1150	.0000	.0517	.0023	.0501	.0028	2.52531	2.32794	2.32249
	15	.1033	.0261	.0522	.0521	.0514	.0525	1.90419	1.79495	1.79270
	20	.0910	.0221	.0523	.0457	.0367	.0610	1.53893	1.47259	1.43152
	25	.0846	.0290	.0467	.0496	.0465	.0499	1.31405	1.26870	1.26814
	35	.0839	.0316	.0521	.0519	.0520	.0519	1.07083	1.04440	1.04416
$\theta_2 = 2$, $\mu = 1.5$, $\lambda = 1.5$	5	.3728	.0000	.0863	.0000	.0759	.0000	10.05795	8.34576	8.23918
	9	.1430	.0000	.0593	.0000	.0575	.0000	3.28032	2.96675	2.95639
	11	.1177	.0000	.0578	.0028	.0564	.0033	2.52868	2.33104	2.32559
	15	.1008	.0222	.0482	.0500	.0476	.0504	1.89971	1.79072	1.78847
	20	.0907	.0270	.0505	.0500	.0329	.0614	1.53069	1.46472	1.42387
	25	.0869	.0292	.0509	.0518	.0504	.0520	1.31735	1.27188	1.27131
	35	.0790	.0318	.0481	.0493	.0480	.0493	1.06660	1.04028	1.04004

Table 3.1 (continued)

	5	.3751	.0000	.0869	.0000	.0790	.0000	10.11743	8.39512	8.28788
$\theta_2 = 3$	9	.1400	.0000	.0581	.0000	.0562	.0000	3.27180	2.95905	2.94872
	11	.1099	.0000	.0481	.0020	.0474	.0029	2.50833	2.31229	2.30689
$\mu = 1$	15	.0974	.0214	.0469	.0460	.0462	.0462	1.88206	1.77412	1.77189
$\lambda = 1$	20	.0923	.0268	.0501	.0496	.0362	.0633	1.53829	1.47198	1.43093
	25	.0808	.0263	.0493	.0484	.0492	.0490	1.31533	1.26993	1.26936
	35	.0789	.0319	.0490	.0498	.0488	.0498	1.06896	1.04257	1.04233

$\theta_1 = 2$

	n	r_L	r_U	r_L^*	r_U^*	F_L	F_U	length$_r$	length$_r^*$	length$_F$
	5	.1740	.0000	.0216	.0000	.0177	.0000	11.58937	9.61679	9.49299
$\theta_2 = 0.5$	9	.1171	.0158	.0477	.0450	.0468	.0459	5.29643	4.78694	4.77010
	11	.1125	.0222	.0517	.0482	.0509	.0491	4.42435	4.07491	4.06553
$\mu = 5$	15	.0932	.0277	.0490	.0518	.0487	.0523	3.41040	3.21208	3.20820
$\lambda = 10$	20	.0884	.0297	.0511	.0543	.0355	.0710	2.77945	2.65789	2.58359
	25	.0839	.0336	.0512	.0547	.0507	.0548	2.39427	2.31036	2.30938
	35	.0756	.0315	.0470	.0499	.0468	.0500	1.94083	1.89215	1.89175
	5	.1621	.0000	.0216	.0000	.0176	.0000	11.43124	9.48555	9.36353
$\theta_2 = 1$	9	.1193	.0165	.0473	.0457	.0460	.0464	5.28452	4.77618	4.75938
	11	.1082	.0222	.0494	.0465	.0487	.0475	4.41000	4.06171	4.05235
$\mu = 2.5$	15	.0983	.0275	.0495	.0529	.0491	.0534	3.41512	3.21652	3.21264
$\lambda = 5$	20	.0841	.0278	.0493	.0476	.0339	.0621	2.76773	2.64669	2.57270
	25	.0851	.0296	.0514	.0499	.0512	.0502	2.39917	2.31509	2.31410
	35	.0754	.0347	.0493	.0509	.0489	.0510	1.93526	1.88672	1.88632
	5	.1633	.0000	.0221	.0000	.0180	.0000	11.42261	9.47838	9.35646
$\theta_2 = 2$	9	.1166	.0144	.0480	.0421	.0464	.0431	5.29608	4.78662	4.76978
	11	.1131	.0234	.0519	.0513	.0511	.0521	4.39953	4.05206	4.04273
$\mu = 1.25$	15	.1014	.0239	.0518	.0467	.0512	.0471	3.44462	3.24429	3.24037
$\lambda = 2.5$	20	.0928	.0280	.0523	.0512	.0368	.0682	2.78350	2.66176	2.58736
	25	.0863	.0308	.0516	.0496	.0509	.0498	2.39455	2.31063	2.30965
	35	.0742	.0305	.0467	.0480	.0466	.0481	1.93204	1.88359	1.88318
	5	.1667	.0000	.0201	.0000	.0162	.0000	11.51323	9.55361	9.43065
$\theta_2 = 3$	9	.1123	.0171	.0450	.0447	.0437	.0455	5.22968	4.72664	4.71002
	11	.1142	.0229	.0489	.0526	.0479	.0534	4.41260	4.06410	4.05474
$\mu = 0.83$	15	.0986	.0287	.0526	.0510	.0521	.0513	3.41590	3.21726	3.21338
$\lambda = 1.67$	20	.0879	.0308	.0495	.0536	.0328	.0708	2.76140	2.64064	2.56682
	25	.0839	.0292	.0526	.0502	.0516	.0505	2.40232	2.31813	2.31715
	35	.0774	.0311	.0508	.0481	.0505	.0482	1.94076	1.89209	1.89169

$\theta_1 = 5$

	n	r_L	r_U	r_L^*	r_U^*	F_L	F_U	length$_r$	length$_r^*$	length$_F$
	5	.0236	.0000	.0000	.0346	.0000	.0395	16.34822	13.56656	13.38938
$\theta_2 = 0.5$	9	.0725	.0233	.0259	.0546	.0250	.0560	11.22873	10.14524	10.10910
	11	.0750	.0219	.0294	.0472	.0283	.0483	9.70142	8.93221	8.91153
$\mu = 4.4$	15	.0821	.0272	.0378	.0512	.0374	.0515	7.80603	7.34990	7.34108
$\lambda = 22$	20	.0827	.0292	.0468	.0491	.0293	.0668	6.43074	6.14799	5.97597
	25	.0858	.0312	.0506	.0513	.0504	.0516	5.58621	5.38934	5.38710
	35	.0718	.0321	.0488	.0491	.0487	.0493	4.52973	4.41546	4.41454
	5	.0258	.0000	.0000	.0358	.0000	.0397	16.39849	13.60829	13.43054
$\theta_2 = 1$	9	.0755	.0228	.0260	.0499	.0253	.0511	11.24394	10.15898	10.12279
	11	.0730	.0215	.0311	.0468	.0302	.0475	9.70115	8.93196	8.91128
$\mu = 2.2$	15	.0866	.0286	.0426	.0509	.0418	.0517	7.83855	7.38051	7.37165
$\lambda = 11$	20	.0773	.0289	.0422	.0525	.0270	.0688	6.38316	6.10250	5.93176
	25	.0804	.0324	.0476	.0516	.0474	.0517	5.56561	5.36947	5.36724
	35	.0740	.0344	.0498	.0505	.0498	.0511	4.51381	4.39994	4.39903

Table 3.1 (continued)

$\theta_2 = 2$ $\mu = 1.1$ $\lambda = 5.5$	5	.0261	.0000	.0000	.0358	.0000	.0391	16.42718	13.63211	13.45403
	9	.0712	.0243	.0246	.0560	.0238	.0569	11.10716	10.03542	9.99967
	11	.0698	.0249	.0280	.0489	.0271	.0496	9.63882	8.87458	8.85404
	15	.0832	.0264	.0370	.0475	.0361	.0485	7.84741	7.38885	7.37999
	20	.0807	.0290	.0455	.0474	.0296	.0658	6.42462	6.14213	5.97028
	25	.0823	.0317	.0484	.0513	.0475	.0516	5.57657	5.38004	5.37781
	35	.0784	.0297	.0500	.0477	.0499	.0479	4.54801	4.43328	4.43235
$\theta_2 = 3$ $\mu = 0.73$ $\lambda = 3.67$	5	.0252	.0000	.0000	.0391	.0000	.0421	16.43485	13.63847	13.46032
	9	.0722	.0203	.0247	.0509	.0232	.0521	11.17478	10.09651	10.06054
	11	.0755	.0241	.0339	.0515	.0330	.0526	9.76215	8.98812	8.96731
	15	.0836	.0287	.0391	.0512	.0385	.0515	7.81748	7.36068	7.35184
	20	.0861	.0280	.0476	.0508	.0291	.0684	6.44740	6.16392	5.99145
	25	.0798	.0284	.0464	.0489	.0461	.0490	5.55471	5.35895	5.35673
	35	.0757	.0321	.0492	.0494	.0490	.0494	4.54271	4.42811	4.42718

1. the result based on r is quite inaccurate and not symmetric. Also, the expected length is the longest of them all.
2. r^* gives relatively accurate coverage probabilities for sample size greater than 25. But it is unsatisfactory for a small sample size less than 25.
3. The results based on exact method are the best in coverage probabilities and expected lengths. And the coverage probabilities and expected length based on r^* are the second. But the results based on r are poor.

Though the result about small sample is unsatisfactory, we obtain the fact that the proposed likelihood-based method r^* has almost the same properties as exact method.

Next, we want to compare three test statistics in terms of the probability of Type I error and power. We assume that the null hypothesis is $H_0 : \theta_1 \geq 2$ and the alternative hypothesis is $H_1 : \theta_1 < 2$. As mentioned above, the inverse Gaussian distribution has special meaning when the shape parameter $\theta_1 < 2$. We obtain the probability of Type I error and power of r and r^* with critical value -1.645, which is 5% quantile of the standard normal distribution. For the test given by Hsieh (1990), we use critical values given in Hsieh (1990) with respect to the values of θ_1 and sample sizes. When the sample size is even, the critical value is obtained by the average of neighboring odd values. These values according to the sample size are given below:

n	5	10	15	20	25	30
critical value	0.978	1.106	1.196	1.260	1.309	1.349

Table 3.2 shows the results. In this Table, we also replicate 10,000 to obtain the probability of Type I error and power. We assume $\theta_1 = 0.5$, 1, 1.5, 2 and $\theta_2 = 0.5$, 1, 2, 5. In Table 3.2, the probability of Type I error corresponds to the value when $\theta_1 = 2$. And the power of tests corresponds to the values when $\theta_1 = 0.5$, 1, 1.5.

The results do not make substantial differences with the change of nuisance parameter θ_2. The probabilities of Type I error of r^* and exact method are almost same. In fact, r^* has almost the same Type I error rate as the exact method. From a viewpoint of the power, r^* is as powerful as the exact method. As we have seen in this Table, the Type I error rate and the power of r were poor.

Table 3.2: Estimated probabilities of Type I error and power

θ_1	n	$\theta_2 = 0.5$			$\theta_2 = 1$			$\theta_2 = 2$			$\theta_2 = 5$		
		r	r^*	F	r	r^*	F	r	r^*	F	r	r^*	F
0.5	5	.394	.528	.521	.387	.528	.520	.390	.530	.523	.395	.530	.523
	10	.695	.777	.775	.687	.769	.768	.689	.766	.765	.692	.777	.775
	15	.849	.892	.891	.852	.894	.894	.851	.894	.893	.852	.892	.891
	20	.929	.950	.949	.927	.948	.947	.931	.952	.951	.931	.951	.951
	25	.968	.978	.978	.967	.977	.977	.968	.977	.977	.969	.979	.979
	30	.985	.989	.989	.984	.989	.989	.985	.990	.990	.986	.991	.990
1.0	5	.142	.257	.252	.144	.262	.256	.136	.253	.246	.143	.258	.252
	10	.295	.403	.400	.290	.404	.402	.294	.403	.401	.292	.401	.399
	15	.415	.513	.512	.417	.519	.518	.415	.516	.514	.419	.519	.518
	20	.526	.609	.608	.528	.610	.609	.524	.609	.608	.527	.609	.608
	25	.618	.688	.687	.617	.687	.686	.619	.692	.691	.616	.687	.686
	30	.691	.752	.751	.689	.747	.747	.689	.750	.749	.696	.755	.755
1.5	5	.044	.119	.114	.048	.121	.116	.044	.114	.109	.047	.117	.112
	10	.089	.158	.157	.086	.156	.154	.088	.159	.157	.085	.155	.154
	15	.120	.189	.188	.121	.188	.187	.122	.189	.188	.123	.190	.188
	20	.151	.218	.217	.151	.216	.215	.155	.219	.218	.156	.222	.221
	25	.186	.252	.251	.184	.247	.247	.183	.244	.244	.189	.254	.254
	30	.217	.284	.283	.213	.275	.274	.213	.274	.274	.219	.280	.280
2.0	5	.014	.049	.047	.014	.050	.048	.014	.051	.048	.012	.047	.044
	10	.022	.050	.050	.019	.048	.047	.020	.049	.048	.021	.049	.049
	15	.024	.050	.050	.025	.048	.048	.024	.049	.049	.024	.051	.050
	20	.026	.048	.047	.028	.050	.050	.028	.049	.049	.027	.048	.048
	25	.031	.053	.052	.028	.048	.048	.029	.049	.049	.028	.049	.048
	30	.030	.049	.049	.028	.046	.046	.029	.049	.049	.031	.050	.050

Table 3.3: Confidence intervals for Jug Bridge

Method	90%	95%	99%
Exact	(0.875, 2.706)	(0.756, 2.945)	(0.628, 3.237)
r	(0.986, 2.885)	(0.861, 3.129)	(0.726, 3.429)
r^*	(0.876, 2.708)	(0.757, 2.946)	(0.628, 3.239)

Example 3.1 Hsieh (1990) considered a set of data representing runoff amounts of Jug Bridge, Maryland :

0.17	0.19	0.23	0.33	0.39	0.39	0.40	0.45	0.52	0.56	0.59	0.64	0.66
0.70	0.76	0.77	0.78	0.95	0.97	1.02	1.12	1.24	1.59	1.74	2.92	

Folks and Chhikara (1978) mentioned the data given above to be very well described by the inverse Gaussian distribution. Hsieh (1990) used the data for testing $H_0 : \theta_1 \geq 2$ versus $H_1 : \theta_1 < 2$. He reported statistics related to this data as $n = 25$, $\overline{X} = 0.803$, $V = 0.695$ and $W = 1.792$. He rejected H_0 at 1% significance level with the critical value 1.078. But this conclusion is wrong. In his Table, 5% quantile of W is 1.309, so he must have accepted $H_0 : \theta_1 \geq 2$ at 5% significance level. In our proposed methods, $r_0 = -0.350$(p-value = 0.363) and $r_0^* = -0.599$(p-value = 0.275), so we accept the null hypothesis.

From the above data, we found that $\widehat{\theta_1} = 1.792$, $\widehat{\theta_2} = 3.185$. We construct the 90%, 95% and 99% confidence interval for θ_1 based on three statistics in Table 3.3.

From Table 3.3, the confidence interval based on r^* is almost same as exact method.

4. Conclusions

We have suggested two likelihood-based methods. Though there exists an exact method, it is inconvenient to perform a test because one must possess tables with respect to specific values of parameters and sample sizes. The performances of the third order normal approximation, r^*, are as good as the exact method proposed by Hsieh (1990). However, we performed a long calculation to gain a highly exact approximation.

Our proposal has a merit that a well-known normal table can be used to test about the shape parameter of the inverse Gaussian distribution even in small sample.

References

Barndorff-Nielsen, O. E. (1986). Inference on full or partial parameters based on the standardized signed log likelihood ratio, *Biometrika*, **73**, 307–322.

Barndorff-Nielsen, O. E. (1991). Modified signed log likelihood ratio, *Biometrika*, **78**, 557–563.

Barndorff-Nielsen, O. E. and Cox, D. R. (1994). *Inference and Asymptotics*, Chapman & Hall/CRC, London.

Chhikara, R. S. and Folks, J. L. (1977). The inverse Gaussian distribution as a lifetime model, *Technometrics,* **19**, 461–468.

Chhikara, R. S. and Folks, J. L. (1989). *The Inverse Gaussian Distribution: Theory, Methodology and Applications*, Marcel Dekker, New York.

Cox, D. R. and Reid, N. (1987). Orthogonal parameters and approximate conditional inference (with discussion), *Journal of the Royal Statistical Society, Series B,* **49**, 1–39.

Folks, J. L. and Chhikara, R. S. (1978). The inverse Gaussian distribution and its statistical application–A review, *Journal of the Royal Statistical Society, Series B,* **40**, 263–289.

Hsieh, H. K. (1990). Inferences on the coefficient of variation of an inverse Gaussian distribution, *Communications in Statistics-Theory and Methods,* **19**, 1589–1605.

Kang, S. G., Kim, D. H. and Lee, W. D. (2004). Noninformative priors for the ratio of parameters in inverse Gaussian distribution, *The Korean Journal of Applied Statistics,* **17**, 49–60.

Mudholkar, G. S. and Natarajan, R. (2002). The inverse Gaussian models: Analogues of symmetry, skewness and kurtosis, *Annals of the Institute of Statistical Mathematics,* **54**, 138–154.

Seshadri, V. (1999). *The Inverse Gaussian Distribution: Statistical Theory and Applications*, Springer, New York.

Tweedie, M. C. K. (1957a). Statistical properties of inverse Gaussian distributions I, *The Annals of Mathematical Statistics,* **28**, 362–377.

Tweedie, M. C. K. (1957b). Statistical properties of inverse Gaussian distributions II, *The Annals of Mathematical Statistics,* **28**, 696–705.

Whitmore, G. A. (1979). An inverse Gaussian model for labour turnover, *Journal of the Royal Statistical Society, Serise A,* **142**, 468–478.

[Received May 2008, Accepted July 2008]