Appeared in: Microelectronics Reliability, 1999, 39, 1413-1421.

MAXIMUM LIKELIHOOD PREDICTIVE DENSITIES FOR THE INVERSE GAUSSIAN DISTRIBUTION WITH APPLICATIONS TO RELIABILITY AND LIFETIME PREDICTIONS

Zhenlin Yang
Department of Statistics and Applied Probability
National University of Singapore
Lower Kent Ridge Road
Singapore 119260
Email: ecsyzl@nus.edu.sg

Abstract

Maximum likelihood predictive densities (MLPD) for the inverse Gaussian distribution are derived for the cases of one or both parameters unknown. They are then applied to obtain estimators of the reliability function and prediction or shortest prediction intervals for a future observation. Comparisons with the existing likelihood or frequentist methods show that the MLPD estimators of reliability gives smaller bias and smaller MSE for a wide range of population values, and that the MLPD prediction intervals are shorter while preserving the correct coverage probability. The shortest MLPD prediction intervals further sharpen the above equitailed MLPD intervals in terms of interval lengths.

Keywords: Inverse Gaussian distribution; maximum likelihood predictive density; prediction intervals; reliability; shortest prediction intervals.

Appeared in: Microelectronics Reliability, 1999, 39, 1413-1421.

1 INTRODUCTION

In studying the reliability and life testing of a product or device, one often needs to estimate the reliability at a certain time or construct prediction intervals (PI) or bounds for the lifetime of a new product or device. The inverse Gaussian (IG) distribution has been proven to be a very suitable model for such studies (see, for example,[2, 9, 11]) as it arises as the first passage time distribution of a Brownian motion, which is suitable for describing many physical phenomena such as the time to fatigue of a metal specimen and the time to failure of an electronic component, etc.. It is well known that this distribution is a very useful alternative to the popular lifetime distributions such as Weibull, gamma and lognormal. Its probability density function (pdf), denoted by $IG(\mu, \lambda)$ usually takes the following form:

$$f(y; \mu, \lambda) = (\lambda/2\pi y^3)^{1/2} \exp[-\lambda(y-\mu)^2/2\mu^2 y], y > 0; \mu > 0, \lambda > 0. \tag{1.1}$$

The mean, variance, skewness and kurtosis of this distribution are, respectively, μ, μ^3/λ, $3\sqrt{\mu/\lambda}$, and $15\mu/\lambda$. Thus, λ is a scale parameter, and μ is a location and scale parameter. Let $\overline{Y}$ and $\widetilde{Y}$ be respectively the arithmetic and harmonic means of the past sample. The maximum likelihood estimators (MLE) of μ and λ are, respectively, $\hat{\mu} = \overline{Y}$ and $\hat{\lambda} = 1/(\widetilde{Y}^{-1} - \overline{Y}^{-1})$. The restricted MLE of μ for given λ is $\hat{\mu}(\lambda) = \overline{Y}$ and of λ for given μ is $\hat{\lambda}(\mu) = n\mu^2/Q$ where $Q = \sum_{i=1}^{n}(Y_i - \mu)^2/Y_i$.

The popularity of this distribution can be seen from a large number of research articles already existed in the literature and two special monographs [6, 17]. Chhikara and Folks [6, p156] give a discussion on why and when the inverse Gaussian distribution is better than the other distributions such as lognormal in reliability studies. Many known properties that parallel those of the normal distribution may be another reason why the IG distribution is so attractive. It is known that in many of the reliability and life testing studies, it is necessary to estimate the unknown pdf. There many ways to do so, namely, the maximum likelihood method, nonparametric method, Bayesian method and fiducial method. In this article, we consider the maximum likelihood predictive density (MLPD) [13] as an estimator of the unknown pdf and then apply it to obtain an estimator of the reliability function and to obtain prediction intervals (PI) or shortest prediction intervals (SPI) [8] for a future observation.

Reliability estimation for the IG distribution has been considered by [4, 5] where MLE and MVUE of the reliability were considered and compared, by [14, 18] where confidence bounds on reliability were obtained, and by [1, 15] where Bayes estimation of reliability were considered. Jain and Jain [12] proposed estimating IG reliability function using Weibull distribution as the two distributions are well-known competitors of each other. Since the MLPD estimator is similar in nature to the MLE, we will restrict our comparisons between these two. Monte Carlo simulation results show that the MLPD estimator generally outperforms the MLE as it gives smaller bias and smaller MSE for a wide range of population values.

Exact PIs for a future IG observation were obtained by [7]. These intervals can be very wide and are of two-sided only. The PIs obtained from the MLPD overcome these drawbacks as they can be both one-sided and two-sided, and the interval lengths can be much shorter than the exact frequentist PIs. Furthermore, the MLPD SPIs can readily be obtained, which can further shorten up the equitailed MLPD PIs significantly. Simulations show that the MLPD PIs preserve excellent coverage properties. Numerical examples are given to illustrate the MLPD estimator of the reliability and the MLPD PIs and SPIs.

In Section 2, we derive the MPLDs for a future IG observation based on a past sample from the same population. Section 3 applies these MLPDs to obtain estimators of the reliability function and presents some numerical examples and simulation results for illustrations and comparisons. Section 4 applies the MLPDs to obtain PIs and SPIs for a future observation, and evaluates the performance of these intervals using Monte Carlo simulations. Section 5 gives some discussions.

2 MLPDs FOR THE IG DISTRIBUTION

Speaking about the methods of estimating a pdf, the simplest one may be the maximum likelihood method, in which the MLE of the unknown pdf is obtained by replacing the unknown parameters by their MLEs. Although this approach is simple, it often gives liberal PIs [10]. The MLPD differs from the MLE in that it maximizes the joint pdf of both past and future observations. It was introduced in 1982 by Lejeune and Faulkenberry [13]. It, however, has not received as much attention since then as it deserved, hence it is the purpose of this paper to explore its usefulness in the context of reliability estimation and lifetime prediction for the IG distribution. The formal definition of the MLPD for a future observation is as follows.

Let $\mathbf{Y} = (Y_1, Y_2, ..., Y_n)$ be a sample of past observations and Y_{n+1} be a future observation, from a population with pdf $f(y; \theta)$ indexed by a vector of parameters θ. The MLPD for $T = Y_{n+1}$ based on $\mathbf{Y} = \mathbf{y}$ is defined as

$$\hat{f}(t \mid \mathbf{y}) \; \propto \; \max_{\theta} f(t; \theta) f(\mathbf{y}; \theta), \tag{2.1}$$

where $f(t; \theta)$ is the pdf of T, and $f(\mathbf{y}; \theta)$ is the joint pdf of $\mathbf{Y}$. Clearly, a closed-form expression for an MLPD depends on the existence of explicit solutions for the maximum of (2.1). This is clearly not a problem for the IG distribution. Lejeune and Faulkenberry provided a theorem stating the MLPD possesses the usual large sample property in the sense that it converges as n goes large to the pdf of T. They also gave conditions (though stringent) under which the MLPD and the Bayesian prediction density are equivalent. It should be noted that the definition of the MLPD is not restricted to the case of a single future observation, it can be any function of a future sample. Now, the joint pdf of T and $\mathbf{Y}$ for the case of IG distribution can be written as

$$f(t; \theta) f(\mathbf{y}; \theta) = \left(\frac{\lambda}{2\pi}\right)^{\frac{n+1}{2}} \prod_{i=1}^{n+1} y_i^{-\frac{3}{2}} \exp\left[-\frac{\lambda}{2\mu^2} \sum_{i=1}^{n+1} \frac{(y_i - \mu)^2}{y_i}\right]. \tag{2.2}$$

Maximum of (2.2) with respect to μ or λ or both gives the MLPD of T. The closed-form expressions for the MLEs or the restricted MLEs make it very easy for finding the maximum of (2.2). Now, let $\hat{\mu}^*(\lambda)$ and $\hat{\lambda}^*(\mu)$ be the restricted MLEs, and $\hat{\mu}^*$ and $\hat{\lambda}^*$ be the unrestricted MLEs based on all $n + 1$ observations. When μ is unknown but λ is known, the maximum of (2.2) occurs at $\hat{\mu}^*(\lambda)$. Substituting $\hat{\mu}^*(\lambda)$ into (2.2) for μ and rearranging the terms to make t explicit, one obtains the MLPD as follows,

$$\tilde{f}(t \mid \mathbf{y}; \lambda) = k(\mathbf{y}, \lambda) t^{-\frac{3}{2}} \exp\left[-\frac{\lambda}{2}\left\{\frac{1}{t} + \frac{n}{\tilde{y}} - \frac{(n+1)^2}{t + n\overline{y}}\right\}\right]. \tag{2.3}$$

When λ is unknown but μ is known, (2.2) is maximized at $\hat{\lambda}^*(\mu)$. Similar manipulations give the MLPD for this case as

$$\tilde{f}(t \mid \mathbf{y}; \mu) = k(\mathbf{y}, \mu) t^{-\frac{3}{2}} \left\{1 + \frac{(t-\mu)^2}{qt}\right\}^{-\frac{n+1}{2}}. \tag{2.4}$$

When both μ and λ are unknown, the maximum of (2.2) occurs at $(\hat{\mu}^*,\ \hat{\lambda}^*)$. Substituting $\hat{\mu}^*$ and $\hat{\lambda}^*$ into (2.2) for μ and λ, followed by some tedious algebra, gives the MLPD

$$\tilde{f}(t \mid \mathbf{y}) = k(\mathbf{y}) t^{-\frac{3}{2}} \left\{\frac{n}{\tilde{y}} + \frac{1}{t} - \frac{(n+1)^2}{t + n\overline{y}}\right\}^{-\frac{n+1}{2}}. \tag{2.5}$$

The normalizing constants $k(\mathbf{y}, \lambda)$, $k(\mathbf{y}, \mu)$ and $k(\mathbf{y})$ for the MLPDs can be found through either integration or numerical integration. The MLPD (2.4) is seen to be the same as the Bayesian predictive density of Chhikara and Guttman [7] where the constant $k(\mathbf{y}, \mu)$ was shown to be $\mu/\sqrt{q}\beta(1/2, n/2)$. There are no closed form expressions for $k(\mathbf{y}, \lambda)$ and $k(\mathbf{y})$, hence the use of numerical integration is necessary, which can be done easily using many statistical software such as Mathematica. Now, we summarize the properties of the MLPDs (2.3)-(2.5).

Theorem 2.1. The MLPDs defined in (2.3)-(2.5) are such that, for each fixed $\mathbf{y}$, i) they are all proper probability density functions; ii) they are all unimodal, provided $n > 2$ in (2.4) and (2.5); iii) for (2.3) and (2.5) all $k \geq 1$ moments are infinite; and for (2.4) the kth moment exists provided that $n > 2k - 2$.

Proof: For i), it is easy to see that as $t \to 0$, $\tilde{f}(t \mid \mathbf{y}; \lambda) \to 0$, and $\tilde{f}(t \mid \mathbf{y}; \mu)$ and $\tilde{f}(t \mid \mathbf{y})$ are of order $O(t^{n/2-1})$, hence they are integrable in $[0, 1]$. As $t \to \infty$, all three functions are of order $O(t^{-3/2})$ or smaller, hence are integrable in $[1, \infty)$. Thus, all three functions are integrable and hence proper pdfs.

For ii), when $n > 2$, $\tilde{f}(t \mid \mathbf{y}; \mu)$ and $\tilde{f}(t \mid \mathbf{y})$ also converge to 0 as $t \to 0$. Unimodality follows by further showing that the derivative for each of the (2.3)-(2.5) has one and only one change of sign. We choose the most realistic case (both parameters unknown) to give a detailed proof. The other two cases can be handled in a similar way. We have,

$$\frac{\partial \tilde{f}(t \mid \mathbf{y})}{\partial t} = \frac{\tilde{f}(t \mid \mathbf{y})}{2t} \left\{ \frac{n}{\tilde{y}} + \frac{1}{t} - \frac{(n+1)^2}{(t+n\overline{y})} \right\}^{-1} \left\{ \frac{c(t)}{t(t+n\overline{y})^2} - \frac{3n}{\tilde{y}} \right\},$$

where $c(t) = (n-2)(t+n\overline{y})^2 + 3(n+1)^2 t(t+n\overline{y}) - (n+1)^3 t^2$. Hence the problem reduces to show that the functions $c(t)$ has one and only one change of sign as all the other parts are either positive functions of t or constants. It is easy to show that $d^2c(t)/dt^2 = 2n(4-n^2) < 0$, which means that $c(t)$ is strictly concave. Further $c(0) = (n-2)n^2\overline{y}^2 > 0$ and $c(\infty) = -\infty$, hence $c(t)$ has one and only one change of sign.

For iii), as $t \to \infty$, $\tilde{f}(t \mid \mathbf{y}; \lambda)$ and $\tilde{f}(t \mid \mathbf{y})$ are of order $O(t^{-3/2})$, and $\tilde{f}(t \mid \mathbf{y}; \mu)$ is of order $t^{-(n+4)/2}$. Hence for (2.3) and (2.5) all $k \geq 1$ moments are infinite, whereas for (2.4) the kth moment exists provided $k < (n+2)/2$.

It should be noted that Bayesian predictive densities are available for all the three cases considered above [7]. Thus, one might be interested in considering the Bayesian predictive density as an estimator of the pdf and use it for the reliability estimation, etc..

3 APPLICATION TO RELIABILITY ESTIMATION

The reliability function of a distribution is simply defined as $R(t) = 1 - F(t)$, where $F(t)$ is the cumulative density function of a random variable T. For the IG distribution, the $R(t)$ was shown to be related to the cdf Φ of the standard normal random variable as follows

$$R(t; \mu, \lambda) = \Phi\left(\sqrt{\frac{\lambda}{t}}\left(1 - \frac{t}{\mu}\right)\right) - e^{2\lambda/\mu}\Phi\left(-\sqrt{\frac{\lambda}{t}}\left(1 + \frac{t}{\mu}\right)\right). \tag{3.1}$$

Plots of $R(t; \mu, \lambda)$ and $f(t; \mu, \lambda)$ for different combinations of values of μ and λ are given in Figure 3.1. Clearly, when parameter(s) are unknown, one can easily obtain the MLE of $R(t; \mu, \lambda)$

by replacing the unknown parameter(s) by their MLE(s). This leads to the MLE of $R(t;\mu,\lambda)$ as $\widehat{R}(t;\mu) = R(t;\mu,\hat{\lambda}(\mu))$ when λ is unknown, $\widehat{R}(t;\lambda) = R(t;\hat{\mu}(\lambda),\lambda)$ when μ is unknown, and $\widehat{R}(t)$ $= R(t;\hat{\mu},\hat{\lambda})$ when both μ and λ are unknown, where $\hat{\mu}(\lambda)$ and $\hat{\lambda}(\mu)$ are the restricted MLEs, and $\hat{\mu}$ and $\hat{\lambda}$ are the MLEs, given in the introduction.

(Insert Figure 3.1 near here)

Alternatively, one can easily obtain estimators of $R(t;\mu,\lambda)$ for various situations by considering the MLPD as an estimator of the unknown pdf. Thus, the MLPD estimators of the reliability function are defined as

$$\widetilde{R}(t;\mu) = 1 - \int_0^t \tilde{f}(x \mid \mathbf{y};\mu)dx, \tag{3.2}$$

$$\widetilde{R}(t;\lambda) = 1 - \int_0^t \tilde{f}(x \mid \mathbf{y};\lambda)dx, \tag{3.3}$$

$$\widetilde{R}(t) = 1 - \int_0^t \tilde{f}(x \mid \mathbf{y})dx, \tag{3.4}$$

for the cases of i) μ known but λ unknown, ii) λ known but μ unknown, and iii) both μ and λ unknown. There are other estimators of the reliability function for the IG distribution, such as the MVUE given by Chhikara and Folks [4]. However, our MLPD estimator has close analogy to the MLE, hence we restrict our comparisons between these two estimators.

To see the performance of the MLPD estimators of $R(t)$, we first apply them to a couple of real data sets. The estimated $R(t)$'s are then plotted together with the corresponding MLEs in Figure 3.2. Though the estimation of $R(t)$ from both methods requires only the values of n and the MLEs of the parameters, the original data are provided for completeness. The one parameter known cases can be handled by using the estimated value as the true value for the MLPD estimator. However, it can not be done in the same way for the MLEs, hence comparison can only be done for the case of both parameters unknown. In the first example, the two methods give slightly different estimates of $R(t)$, while in the second example, the two estimates are very similar.

(Insert Figure 3.2 near here)

Example 3.1. Fatigue lives (in hours) for 10 bearings tested on a certain tester [3]: 152.7, 172.0, 172.5, 173.3, 193.3, 204.7, 216.5, 234.9, 262.6, 422.6. The MLEs are $\hat{\mu} = 220.48$ and $\hat{\lambda} = 2708.86$. The data is slightly skewed to the right.

Example 3.2. 46 repair times (in hours) for an airborne communication transceiver [6, p139]: .2, .3, .5, .5, .5, .5, .6, .6, .7, .7, .7, .8, .8, 1.0, 1.0, 1.0, 1.0, 1.1, 1.3, 1.5, 1.5, 1.5, 1.5, 2.0, 2.0, 2.2, 2.5, 2.7, 3.0, 3.0, 3.3, 3.3, 4.0, 4.0, 4.5, 4.7, 5.0, 5.4, 5.4, 7.0, 7.5, 8.8, 9.0, 10.3, 22.0, 24.5. The MLEs for the parameters are $\hat{\mu} = 3.61$ and $\hat{\lambda} = 1.6667$. The data is skewed to the right.

It may be interesting to compare the two estimators when it is known which population the data came from. To this end, one data set is generated from a known IG population by computer simulation so that one can compare directly the estimated $R(t)$'s with the 'true' one. The 'known' $R(t)$ and the estimated ones are plotted together in Figure 3.3.

Example 3.3. *Simulated Data.* Fifteen observations are generated from an IG(1,1) population: 0.9144, 0.2517, 0.6506, 0.9421, 0.9112, 0.2515, 0.5057, 0.9760, 1.5257, 0.5819, 0.4591, 0.6711, 0.3103, 0.3733, 0.3696. The resulted arithmetic and harmonic means are, respectively, $\overline{y} = 0.6463$, and $\widetilde{y} = 0.4936$. Plots of the pdf and $R(t)$, and the estimated pdfs and $R(t)$'s are given in Figure 3.3 for the cases of one parameter unknown and both parameters unknown. It is seen from the plots that the two estimators perform very similarly.

(Insert Figure 3.3 near here)

Numerical examples can not tell the actual performance of the estimators in terms of bias and the mean square error (MSE), etc.. Hence we now turn our attention to the Monte Carlo simulations. Specifically, the bias and MSE of the ML and MLPD estimators are simulated for many different combinations of μ and λ values at several time points. The algorithm described in [6, p52] is used for generating inverse Gaussian random variates. Each row of the simulation results is based on 10,000 random samples. The results are summarized in Tables 3.1. Notice that the selected time points cover a very wide region of the population values (about 80%). The

Table 3.1: Simulated biases and MSEs of the ML and MLPD Estimators of $R(t)$

(μ, λ)	n	t value	$r(t)$	bias(MLE)	bias(MLPD)	MSE(MLE)	MSE(MLPD)
(1.0, 0.25)	10	0.3	0.5446	0.0000	0.0036	0.0166	0.0141
		0.5	0.4001	-0.0132	0.0029	0.0156	0.0133
		1.0	0.2384	-0.0237	0.0048	0.0104	0.0088
		1.5	0.1657	-0.0247	0.0079	0.0073	0.0063
		2.0	0.1237	-0.0204	0.0137	0.0051	0.0048
	20	0.3	0.5446	0.0017	0.0043	0.0071	0.0065
		0.5	0.4001	-0.0068	0.0024	0.0066	0.0061
		1.0	0.2384	-0.0150	0.0009	0.0050	0.0045
		1.5	0.1657	-0.0152	0.0033	0.0037	0.0034
		2.0	0.1237	-0.0129	0.0068	0.0028	0.0027
(1.0, 1.0)	10	0.3	0.8343	0.0114	-0.0082	0.0082	0.0074
		0.6	0.5536	0.0041	0.0044	0.0176	0.0151
		0.9	0.3750	-0.0122	0.0025	0.0163	0.0141
		1.2	0.2628	-0.0179	0.0049	0.0129	0.0114
		1.5	0.1892	-0.0167	0.0103	0.0095	0.0088
	20	0.3	0.8343	0.0067	-0.0041	0.0044	0.0042
		0.6	0.5536	0.0029	0.0026	0.0080	0.0073
		0.9	0.3750	-0.0054	0.0013	0.0075	0.0069
		1.2	0.2628	-0.0083	0.0027	0.0060	0.0056
		1.5	0.1892	-0.0097	0.0038	0.0047	0.0044
(1.0, 4.00)	10	0.5	0.8884	0.0091	-0.0127	0.0055	0.0055
		1.0	0.4056	-0.0059	0.0022	0.0176	0.0152
		1.5	0.1407	-0.0098	0.0135	0.0072	0.0071
	20	0.5	0.8884	0.0037	-0.0082	0.0030	0.0030
		0.8	0.5877	0.0031	0.0002	0.0084	0.0078
		1.1	0.3318	-0.0061	0.0007	0.0074	0.0068
		1.4	0.1753	-0.0067	0.0051	0.0046	0.0044

simulation results clearly favor the MLPD approach as smaller bias and MSE are observed. It is interesting to note that the ML approach often underestimates $R(t)$, whereas the MLPD approach almost always slightly overestimates $R(t)$.

4 APPLICATION TO LIFETIME PREDICTION

Prediction based on a known density is very simple. For example, a 95% PI for a future observation T is simply defined as (a, b) such that $\int_{-\infty}^{a} f(t)dt = 0.025$ and $\int_{b}^{\infty} f(t)dt = 0.025$, where f is the known pdf. This idea can easily be carried over to the case when f is unknown and is estimated by $\hat{f}$, say. The resulted 95% PI is thus $(\hat{a}, \hat{b})$ such that $\int_{-\infty}^{\hat{a}} \hat{f}(t)dt = 0.025$ and

$\int_{\hat{b}}^{\infty} \hat{f}(t)dt = 0.025$. The $\hat{f}$ is usually referred to as the predictive density. When the predictive density is unimodal, the 95% SPI based on $\hat{f}(t)$ is defined as $(\tilde{a}, \tilde{b})$ such that $\hat{f}(\tilde{a}) = \hat{f}(\tilde{b})$ and $\int_{\tilde{a}}^{\tilde{b}} \hat{f}(t)dt = 0.95$ [8].

When prediction of a future observation is of concern, exact frequentist PIs are available in [7]. However, as indicated in the introduction, these intervals can be too long and they are of two-sided only. To overcome these difficulties, we derive the MLPD PIs of T for various situations, i.e., one parameter unknown and both parameters unknown, and compare them with the exact frequentist ones. Note that the notation T and Y_{n+1} are now inter-changeable.

Using the MLPD, a $100(1-\alpha)\%$ equitailed MLPD PI for T is defined as $(L_e(\mathbf{Y}), U_e(\mathbf{Y}))$ such that

$$\int_0^{L_e(\mathbf{Y})} \hat{f}(t \mid \mathbf{y})dt = \alpha/2, \text{ and } \int_{U_e(\mathbf{Y})}^{\infty} \hat{f}(t \mid \mathbf{y})dt = \alpha/2. \tag{4.1}$$

When $\hat{f}(t \mid \mathbf{y})$ is unimodal as it is the case for the inverse Gaussian MLPDs, a $100(1-\alpha)\%$ SPI [8] based on the MLPD $\hat{f}(t \mid \mathbf{y})$ is defined as $\{L_s(\mathbf{Y}), U_s(\mathbf{Y})\}$ such that

$$\hat{f}(L_s(\mathbf{Y}) \mid \mathbf{y}) = \hat{f}(U_s(\mathbf{Y}) \mid \mathbf{y}), \text{ and } \int_{L_s(\mathbf{Y})}^{U_s(\mathbf{Y})} \hat{f}(t \mid \mathbf{y})dt = 1-\alpha. \tag{4.2}$$

The data sets considered in Examples 3.1 and 3.2 are used again to illustrate the new intervals and to compare them with the frequentist ones. Again the knowledge of the sample size and the MLEs are sufficient for the calculations of the PIs. The assumed 'known' parameter value is taken to be its estimated value. The 90%, 95% and 99% PIs are calculated. The results are summarized in Table 4.1. The results from the two examples show that MLPD PIs are generally superior to the exact frequentist PIs in the sense that they provide shorter and sometime much shorter intervals. The more skewed the data is, the larger the difference between the two types of intervals in general. The shortest MLPD PIs further reduce the interval length significantly, indicating the usefulness of the concept of the SPI.

It should be noted that the IG distribution not only applies to the fatigue life data, but also to some other situations as long as the physical situation is conformable with the notion of the

Table 4.1: Summary of Prediction Intervals for the Two Examples

			Shortest MLPD PI		MLPD PI		Frequentist PI	
	Ex.	1-α	Lower	Upper	Lower	Upper	Lower	Upper
μ unknown	3.1	0.90	118.4065	324.8901	132.5839	349.7623	136.5621	363.9824
		0.95	108.3470	359.3695	121.4952	385.1988	125.0446	401.2984
		0.99	91.5024	438.3416	102.8815	466.4547	105.6602	486.3372
	3.2	0.90	0.1160	9.1213	0.3635	14.2858	0.4659	32.0564
		0.95	0.0998	14.2918	0.2873	20.8617	0.3526	45.1257
		0.99	0.0795	32.3827	0.1919	42.0887	0.2211	85.7317
λ unknown	3.1	0.90	112.2316	324.1150	125.8338	346.0465	132.2020	367.7056
		0.95	97.9738	359.8110	111.8942	385.0154	117.9987	411.9660
		0.99	69.5364	449.7630	86.3496	484.7282	91.8448	529.2778
	3.2	0.90	0.1012	8.5671	0.3471	13.0987	0.4525	28.8017
		0.95	0.0852	13.1103	0.2712	18.5284	0.3381	38.5493
		0.99	0.0657	26.9080	0.1751	34.0275	0.2056	63.3901
μ and λ unknown	3.1	0.90	108.6992	333.8017	124.7491	362.0935	126.0114	397.8592
		0.95	94.4503	375.4503	110.5725	408.8904	111.2539	458.1544
		0.99	67.8166	487.6629	85.5138	536.9749	84.5699	635.4311
	3.2	0.90	0.0990	9.0617	0.3475	14.3194	0.4439	33.9674
		0.95	0.0658	14.3288	0.2718	21.0506	0.3314	48.8461
		0.99	0.0738	32.4225	0.1755	43.6147	0.2014	99.1124

first passage time of a Brownian motion. Thus, the MLPD PIs are also applied to some other data sets such as the Maximum Flood Level data in [3]; Strike Duration data for transport industry and metal manufacture industry in [6]. The basic considerations for having more examples are to see the effects of sample size and skewness. The results (not reported) generally agree with the results from the two examples given above.

The advantage can finally be given to the MLPD approach only when coverage properties of the PIs be assessed, which is what is going to be done next. Fitted models from several data sets, including the two data sets considered in Examples 3.1 and 3.2, are used as the true models for generating IG random variates. It is seen that these models cover various sample sizes and population skewness. The simulation process for prediction can be simply described as follows. In each run, a random sample of size $n+1$ is generated. The first n observations are used as the past sample and the last one is treated as the future observation. The MLPD, shortest MLPD and frequentist PIs are calculated based on the past sample, and then checked whether containing the future observation. The lengths of the intervals are also recorded. For 10,000 runs, the proportion of the intervals covering the future observation and the average length of the 10,000 intervals are recorded, and are used as, respectively, the Monte Carlo estimates of the coverage probability and the expected interval length. The above overall process is repeated for several different parameter settings and nominal levels. The simulation results given in Table 4.2 show that the MLPD PIs have coverage probabilities very close to the nominal levels, and can have lengths much shorter than the frequentist PIs.

5 DISCUSSIONS

Maximum likelihood predictive densities for the inverse Gaussian distribution are derived and applied to reliability and lifetime predictions. Numerical examples and Monte Carlo simulations show that this approach is quite promising as it gives reliability estimators with smaller bias and MSE than the MLE approach, and it gives shorter prediction intervals for a future observation than the frequentist approach. Thus, this approach deserves more attention than it actually does. In particular, more general reliability and life testing problems may be studied using this approach, such as prediction concerning a future sample and prediction in the framework of the inverse Gaussian regression. The MLPD approach is somehow similar to the Bayesian approach. Lejeune and Faulkenberry [13] have provided conditions (though stringent) for its Bayesian equivalence. However, to many practitioners the MLPD approach may be more attractive as there is no need of discussions for prior distributions. Calculations of the MLPD estimator of reliability function and MLPD prediction intervals require numerical integrations which seem to hinder the application of this approach. However, simple numerical subroutines, such as Fortran IMSL, exist for such tasks. For those who are relatively unfamiliar with programming languages, powerful and yet user-friendly statistical softwares, such as Mathematica, exist, which make the applications of the MLPD approach very easy. Note that all the plots in this paper were produced by Mathematica 3.0. Numerical integration is involved in the process of evaluating every point in each plot

Table 4.2: Summary of Simulation Results for Prediction Intervals

				MLPD PI		Frequentist PI	
n	μ	λ	$1-\alpha$	C. Prob.	A. Length	C. Prob.	A. Length
				μ unknown			
10	220.48	2708.86	0.90	0.8985	217.1935	0.8980	227.5845
			0.95	0.9499	264.2519	0.9492	276.9224
			0.99	0.9891	366.4443	0.9884	383.6481
20	0.423	5.66	0.90	0.8977	0.3865	0.8957	0.4041
			0.95	0.9519	0.4667	0.9508	0.4878
			0.99	0.9900	0.6344	0.9908	0.6626
198	7.229	2.670	0.90	0.9039	27.6632	0.9036	69.1206
			0.95	0.9470	40.7535	0.9450	93.9296
			0.99	0.9900	79.9240	0.9920	159.8345
				λ unknown			
10	220.48	2708.86	0.90	0.9055	215.8108	0.9075	232.2444
			0.95	0.9515	265.5278	0.9525	287.5875
			0.99	0.9903	386.8654	0.9902	427.6287
20	0.423	5.66	0.90	0.9010	0.3841	0.9021	0.4059
			0.95	0.9479	0.4674	0.9483	0.4948
			0.99	0.9919	0.6568	0.9912	0.7005
46	3.61	1.6667	0.90	0.9021	12.6470	0.9007	28.2361
			0.95	0.9500	18.1195	0.9484	38.1523
			0.99	0.9894	33.6108	0.9886	63.1917
102	1.012	0.119	0.90	0.9060	4.4365	0.8990	25.7640
			0.95	0.9620	7.7926	0.9590	35.8619
			0.99	0.9910	19.6666	0.9860	61.3512
198	7.229	2.670	0.90	0.8973	27.0092	0.8987	66.3844
			0.95	0.9534	39.5709	0.9532	89.4437
			0.99	0.9909	75.5424	0.9906	146.0763
				both μ and λ unknown			
10	220.48	2708.86	0.90	0.8985	217.1935	0.8980	227.5845
			0.95	0.9499	264.2519	0.9492	276.9224
			0.99	0.9891	366.4443	0.9884	383.6481
20	0.423	5.66	0.90	0.8977	0.3865	0.8957	0.4041
			0.95	0.9519	0.4667	0.9508	0.4878
			0.99	0.9900	0.6344	0.9908	0.6626
198	7.229	2.670	0.90	0.9039	27.6632	0.9036	69.1206
			0.95	0.9470	40.7535	0.9450	93.9296
			0.99	0.9900	79.9240	0.9920	159.8345

of the reliability functions. The complicated multi-plots in Figures 3.1-3.3 can be finished in less than a minute if a 300 MHz PC is used.

ACKNOWLEDGEMENTS

The author is grateful to the referees for the constructive comments that lead to improvements on the article. The prompt responses from the Editor-in-Chief, Ninoslav Stojadinović, are also greatly appreciated.

REFERENCES

1. Akman, O. and Huwang, L., Bayes computation for reliability estimation. *IEEE Trans. Reliability*, 1997, 46, 52-55.

2. Bhattacharyya, G. K. and Fries, A., Fatigue failure models - Birnbaum-Saunders vs. inverse Gaussian. *IEEE Trans. Reliability*, 1980, R-31, 439-440.

3. Cheng, R. C. H. and Amin, N. A. K., Maximum likelihood estimation of parameters in the inverse Gaussian distribution with unknown origin. *Technometrics*, 1981, 23, 257-263.

4. Chhikara, R. S. and Folks, J. L., Estimation of the inverse Gaussian distribution. *J. Amer. Statist. Assoc.*, 1974, 69, 250-254.

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

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

7. Chhikara, R.S. and Guttman, I., Prediction limits for the inverse Gaussian distribution. *Technometrics*, 1982, 24, 319-324.

8. Dahiya, R.C. and Guttman, I., Shortest confidence and prediction intervals for the log-normal. *Canadian J. Statist.*, 1982, 10, 277-291.

9. Desmond, A. F., *Local Maxima of Stationary Stochastic Processes and Stochastic Modelling of Fatigue*. Unpublished PhD thesis, University of Waterloo, 1983.

10. Geisser S., *Predictive Inference: An Introduction.* Chapman & Hall, New York London, 1991.

11. Goh, G. J., Tang, L. C. and Lim, S. C., Reliability modelling of stochastic wear-out failure. *Reliability Engineering and Systems Safety*, 1989, 25, 303-314.

12. Jain, R. K. and Jain, S., Inverse Gaussian distribution and its application to reliability. *Microelectron. Reliab.*, 1996, 36, 1323-1335.

13. Lejeune, M. and Faulkenberry, G. D., A simple predictive density function. *J. Amer. Statist. Assoc.*, 1982, 77, 654-659.

14. Padgett, W. J., Confidence bounds on reliability for the inverse Gaussian model. *IEEE Trans. Reliability*, 1979, R-28, 165-168.

15. Padgett, W. J., Bayes estimation of reliability for the inverse Gaussian model. *IEEE Trans. Reliability*, 1981, R-30. 384-385.

16. Padgett, W. J., An approximate prediction interval for the mean of future observations from the inverse Gaussian distribution. *J. Statist. Comput. Simul.*, 1982, 14, 191-199.

17. Seshadri, V., *The inverse Gaussian distribution: A case study in exponential family.* Clarendon Press · Oxford, 1993.

18. Tang, C. L. and Chang, D. S. Reliability bounds and tolerance limits of two inverse Gaussian models. *Microelectron. Reliab.*, 1994, 34, 247-259.