Predictive Inferences for Future Order Statistics Coming from an Inverse Gaussian Distribution

Nicholas A. Nechval, *Member, IAENG,* Konstantin N. Nechval, Vadim Danovich, Gundars Berzins

Abstract — **This paper provides procedures for constructing prediction limits on order statistics of future samples using the results of a previous sample from the same underlying inverse Gaussian distribution. Prediction limits are obtained from a point of view of frequentist approach. Bayesian methods are not considered here. The results have direct application in reliability theory, where the time until the first failure in a group of several items in service provides a measure of assurance regarding the operation of the items. The prediction limits are required as specifications on future life for components, as warranty limits for the future performance of a specified number of systems with standby units, and in various other applications. Prediction limit is an important statistical tool in the area of quality control. The lower prediction limits are often used as warranty criteria by manufacturers. The technique used here does not require the construction of any tables. It requires a quantile of the G_θ and F distributions and is conceptually simple and easy to use. The prediction limits obtained in the paper are generalizations of the usual prediction limits on observations or functions of observations of only one future sample. For illustration, a numerical example is given.**

Index Terms — **Future samples, order statistics, prediction limits**

I. INTRODUCTION

THE inverse Gaussian distribution is a well-known distribution whose properties and applications have a remarkable similarity to those of the normal distribution (Chhikara and Folks [1], Folks and Chhikara [2]). It arises as the distribution of the first passage time of a Browning motion with positive drift and so it is logical to use it as a lifetime model. For example, as outlined in Chhikara and Folks [1], it has been used to describe the time for a reservoir to empty, to describe the interpurchase time for a consumable commodity, to model the time to failure of a device, or to model the distribution of strikes. Dennis, Munholland, and Scott [3] used it to describe the time to extinction of endangered species. For skewed data, the inverse Gaussian distribution has advantages over the

traditional lognormal, gamma, or Weibull distributions in that a comprehensive methodology now exists similar to that of a normal distribution.

In this paper we discuss predictive inferences for the inverse Gaussian distribution, which was proposed by Chhikara and Folks [4], among others, as a lifetime model. The sampling theory and statistical methods for inverse Gaussian parametric estimation and hypothesis testing are well developed and are shown to have close analogy with their counterparts for the normal (Folks and Chhikara [2]). However, for certain applications in reliability and other fields (Banerjee and Bhattacharyya [5], Lancaster [6], Whitmore [7], and Folks and Chhikara [2]), it may be desirable to obtain confidence limits not for a parameter of the distribution, but for a future random observation drawn from the inverse Gaussian distribution itself. Such limits are called prediction limits. Statistical prediction limits have many applications in quality control and in reliability problems and the determination of these limits has been extensively investigated, particularly for the normal distribution. When the characteristic of interest is measured in time, say failure or repair time, distributions other than the normal are generally more appropriate. The gamma and Weibull distributions are frequently applied in studying the properties of life-time phenomena and there are several publications on the subject of their prediction intervals (see, for instance, Aitchison [8]).

Prediction limits can be of several forms. Hahn [9] dealt with simultaneous prediction limits on the standard deviations of all of the k future samples from a normal population. Hahn [10] considered the problem of obtaining simultaneous prediction limits on the means of all of k future samples from an exponential distribution. In addition, Hahn and Nelson [11] discussed such limits and their applications. Mann, Schafer, and Singpurwalla [12] gave an interval that contains, with probability γ, all m observations of a single future sample from the same population.

Fertig and Mann [13] constructed prediction intervals to contain at least $m - k + 1$ out of m future observations from a normal distribution with probability $1-\beta$. They considered life-test data, and the performance variate of interest is the failure time of an item. Their lower prediction limit constitutes a "warranty period".

This paper develops prediction limits (or one-sided prediction intervals) for future order statistics coming from the inverse Gaussian distribution, using classical frequentist approach. In the former case, the statistical prediction limits can be easily obtained when one parameter is assumed known and the other unknown, as well as when both are unknown.

Nicholas. A. Nechval is with the Statistics Department, EVF Research Institute, University of Latvia, Riga LV-1050, Latvia (e-mail: nechval@junik.lv).

Konstantin N. Nechval is with the Applied Mathematics Department, Transport and Telecommunication Instutute, Riga LV-1019, Latvia (e-mail: konstan@tsi.lv).

Vadim Danovich is with the Cybernetics Department, University of Latvia, Riga LV-1050, Latvia (e-mail: vadims@lu.lv).

Gundars Berzins is with the Cybernetics Department, University of Latvia, Riga LV-1050, Latvia (e-mail: gundars.berzins@lu.lv).

II. MATHEMATICAL PRELIMINARIES

A. Constructing Prediction Limits on Future Order Statistics under Certainty of Underlying Models

In this section, we consider the determination of prediction limits on future order statistics under certainty of underlying models. The following results hold.

Theorem 1. Let $Y_1 \leq \ldots \leq Y_l$ be the l ordered observations in a sample of size l from a probability distribution (continuous or discrete) with density function $g_\theta(y)$, distribution function $G_\theta(y)$, where θ is the parameter (in general, vector). Then a lower $(1-\alpha)$ prediction limit h on the kth order statistic Y_k, $k \in \{1, \ldots, l\}$, in a set of l future ordered observations $Y_1 \leq \ldots \leq Y_l$ is given by

$$h = \arg[P_\theta\{Y_k > h\} = 1-\alpha]$$

$$= \arg\left[G_\theta(h) = \frac{k}{k+(l-k+1)f_{2(l-k+1),2k;1-\alpha}} \right], \quad (1)$$

where $f_{2(l-k+1),2k;1-\alpha}$ is the quantile of order $1-\alpha$ for the F distribution.

Proof. If there is a random sample of l ordered observations $Y_1 \leq \ldots \leq Y_l$ from a known distribution (continuous or discrete) with density function $g_\theta(y)$, distribution function $G_\theta(y)$, then for the kth order statistic Y_k it is well-known that

$$P_\theta\{Y_k \leq h\} = \sum_{j=k}^{l} \binom{l}{j} [G_\theta(h)]^j [1-G_\theta(h)]^{l-j}$$

$$= I_{G_\theta(h)}(k, l-k+1)$$

$$= \frac{1}{B(k,l-k+1)} \int_0^{G_\theta(h)} u^{k-1}(1-u)^{(l-k+1)-1} du. \quad (2)$$

It follows from (2) that

$$P_\theta\{Y_k \leq h\} = \frac{\left(\frac{2(l-k+1)}{2l}\right)^{2(l-k+1)/2}}{B\left(\frac{2k}{2}, \frac{2(l-k+1)}{2}\right)} \int_0^{G_\theta(h)} u^{\frac{2(l-k+1)+2k}{2}}$$

$$\times \left(\frac{1-u}{u} \frac{2k}{2(l-k+1)}\right)^{2(l-k+1)/2-1} \frac{-2k}{2(l-k+1)} \left(-\frac{du}{u^2}\right)$$

$$= \frac{\left(\frac{2(l-k+1)}{2l}\right)^{2(l-k+1)/2}}{B\left(\frac{2(l-k+1)}{2}, \frac{2k}{2}\right)} \int_{\frac{1-G_\theta(h)}{G_\theta(h)} \frac{2k}{2(l-k+1)}}^{\infty} f^{\frac{2(l-k+1)}{2}-1}$$

$$\times (1 + \frac{2(l-k+1)}{2k} f)^{-\frac{2(l-k+1)+2k}{2}} df, \quad (3)$$

where

$$f = \frac{1-u}{u} \frac{2k}{2(l-k+1)}. \quad (4)$$

Thus,

$$P_\theta\{Y_k \leq h\} = P_\theta\left\{ F_{2(l-k+1),2k} > \frac{1-G_\vartheta(h)}{G_\theta(h)} \frac{2k}{2(l-k+1)} \right\}$$

$$= 1 - P_\theta\left\{ F_{2(l-k+1),2k} \leq \frac{1-G_\theta(h)}{G_\theta(h)} \frac{2k}{2(l-k+1)} \right\}, \quad (5)$$

where $F_{2(l-k+1),2k}$ has the F distribution with $2(l-k+1)$ and $2k$ degrees of freedom. It follows from (5) that

$$h = \arg[P_\theta\{Y_l > h\} = 1-\alpha] = \arg[P_\theta\{Y_l \leq h\} = \alpha]$$

$$= \arg\left(P_\theta\left\{ F_{2(l-k+1),2k} \leq \frac{1-G_\theta(h)}{G_\theta(h)} \frac{2k}{2(l-k+1)} \right\} = 1-\alpha \right)$$

$$= \arg\left(P_\theta\{ F_{2(l-k+1),2k} \leq f_{2(l-k+1),2k;1-\alpha} \} = 1-\alpha \right). \quad (6)$$

Since (from (6))

$$\frac{1-G_\theta(h)}{G_\theta(h)} \frac{2k}{2(l-k+1)} = f_{2(l-k+1),2k;1-\alpha}, \quad (7)$$

we have that

$$G_\theta(h) = \frac{k}{k+(l-k+1)f_{2(l-k+1),2k;1-\alpha}}. \quad (8)$$

This ends the proof.

Corollary 1.1. Let $Y_1 \leq \ldots \leq Y_l$ be the l ordered observations in a sample of size l from a known probability distribution (continuous or discrete) with density function $g_\theta(y)$, distribution function $G_\theta(y)$. Then an upper α prediction limit h on the kth order statistic Y_k, $k \in \{1, \ldots, l\}$, in a set of l ordered observations $Y_1 \leq \ldots \leq Y_l$ is given by

$$h = \arg[P_\theta\{Y_k > h\} = \alpha] = \arg[P_\theta\{Y_k \leq h\} = 1-\alpha]$$

$$= \arg\left[G_\theta(h) = \frac{k}{k+(l-k+1)f_{2(l-k+1),2k;\alpha}} \right]$$

$$= \arg\left[G_\theta(h) = \frac{k}{k+(l-k+1)/f_{2k,2(l-k+1);1-\alpha}} \right]. \quad (9)$$

The main theorem, which shows how to construct lower (upper) simultaneous prediction limit for the order statistics in all of k future samples when prediction limit for a single future sample is available, is given below.

Theorem 2. Let m_j "future" ordered observations $(Y_{1_j}, \ldots, Y_{m_j})$ represent the jth random sample from the cdf

G_θ, where θ is the parameter (in general, vector), $j \in \{1, \ldots, l\}$, and let $Y_{(k,m_j)}$ denote the kth order statistic in the jth sample of size m_j, where $m_j = m$ for all $j \in \{1, \ldots, l\}$. Assume that all of l samples from the same cdf G_θ are independent. Then a lower simultaneous $(1-\alpha)$ prediction limit h on the kth order statistics $Y_{(k,m_j)}$, $j=1, \ldots, l$, of all of l future samples is given by

$$h = \arg\left[P_\theta\left\{ Y_{(k,m_1)} > h, \ldots, Y_{(k,m_j)} > h, \ldots, Y_{(k,m_l)} > h \right\} = 1-\alpha \right]$$

$$= \arg\left[P_\theta\{ Y_{(k,m)} > h \} = \frac{(l-r+1)f_{2(l-r+1),2r;1-\alpha}}{r+(l-r+1)f_{2(l-r+1),2r;1-\alpha}} = \beta \right]$$

$$= \arg\left[G_\theta(h) = \frac{k}{k+(m-k+1)f_{2(m-k+1),2k;\beta}} \right]. \quad (10)$$

Proof. If there are l independent samples from the same known distribution (continuous or discrete) with density function $g_\theta(y)$, distribution function $G_\theta(y)$, then a lower simultaneous $(1-\alpha)$ prediction limit h on the kth order statistics $Y_{(k,m_j)}$, $j=1, \ldots, l$, of all of l future samples may be obtained as follows:

$$P_\theta\left\{ Y_{(k,m_1)} > h, \ldots, Y_{(k,m_j)} > h, \ldots, Y_{(k,m_l)} > h \right\}$$

$$= \prod_{j=1}^{l} P_\theta\left\{ Y_{(k,m_j)} > h \right\} = [P_\theta\{ Y_{(k,m)} > h \}]^l$$

$$= 1 - \sum_{j=r}^{l} \binom{l}{j} [1 - P_\theta\{ Y_{(k,m)} > h \}]^j [P_\theta\{ Y_{(k,m)} > h \}]^{l-j}$$

$$= 1 - I_{1-P_\theta\{Y_{(r,m)}>h\}}(r, l-r+1)$$

$$= 1 - \frac{1}{\mathrm{B}(r, l-r+1)} \int_0^{1-P_\theta\{Y_{(r,m)}>h\}} u^{r-1}(1-u)^{(l-r+1)-1} du$$

$$= 1 - \frac{\left(\frac{2(l-r+1)}{2l}\right)^{2(l-r+1)/2}}{\mathrm{B}\left(\frac{2r}{2}, \frac{2(l-r+1)}{2}\right)} \int_0^{1-P_\theta\{Y_{(r,m)}>h\}} u^{\frac{2(l-r+1)+2r}{2}}$$

$$\times \left(\frac{1-u}{u}\frac{2r}{2(l-r+1)}\right)^{2(l-r+1)/2-1} \frac{-2r}{2(l-r+1)}\left(-\frac{du}{u^2}\right)$$

$$= 1 - \frac{\left(\frac{2(l-r+1)}{2l}\right)^{2(l-r+1)/2}}{\mathrm{B}\left(\frac{2(l-r+1)}{2}, \frac{2r}{2}\right)} \int_{\frac{P_\theta\{Y_{(r,m)}>h\}}{1-P_\theta\{Y_{(r,m)}>h\}}\frac{2r}{2(l-r+1)}}^{\infty} f^{\frac{2(l-r+1)}{2}-1}$$

$$\times \left(1 + \frac{2(l-r+1)}{2r}f\right)^{-\frac{2(l-r+1)+2r}{2}} df, \quad (11)$$

where

$$f = \frac{1-u}{u}\frac{2r}{2(l-r+1)}, \quad (12)$$

$$r = 1. \quad (13)$$

Thus,

$$P_\theta\left\{ Y_{(k,m_1)} > h, \ldots, Y_{(k,m_j)} > h, \ldots, Y_{(k,m_l)} > h \right\}$$

$$= 1 - P_\theta\left\{ F_{2(l-r+1),2r} > \frac{P_\theta\{Y_{(k,m)} > h\}}{1 - P_\theta\{Y_{(k,m)} > h\}}\frac{2r}{2(l-r+1)} \right\}$$

$$= P_\theta\left\{ F_{2(l-r+1),2r} \le \frac{P_\theta\{Y_{(k,m)} > h\}}{1 - P_\theta\{Y_{(k,m)} > h\}}\frac{2r}{2(l-r+1)} \right\}, \quad (14)$$

where $F_{2(l-r+1),2r}$ has the F distribution with $2(l-r+1)$ and $2r$ degrees of freedom. It follows from (14) that

$$h = \arg\left(P_\theta\left\{ F_{2(l-r+1),2r} \le \frac{P_\theta\{Y_{(k,m)} > h\}}{1 - P_\theta\{Y_{(k,m)} > h\}}\frac{2r}{2(l-r+1)} \right\} = 1-\alpha \right)$$

$$= \arg\left(P_\theta\left\{ F_{2(l-r+1),2r} \le f_{2(l-r+1),2r;1-\alpha} \right\} = 1-\alpha \right). \quad (15)$$

Since (from (15))

$$\frac{P_\theta\{Y_{(k,m)} > h\}}{1 - P_\theta\{Y_{(k,m)} > h\}}\frac{2r}{2(l-r+1)} = f_{2(l-r+1),2r;1-\alpha}, \quad (16)$$

we have that

$$P_\theta\{ Y_{(k,m)} > h \} = \frac{(l-r+1)f_{2(l-r+1),2r;1-\alpha}}{r+(l-r+1)f_{2(l-r+1),2r;1-\alpha}} = \beta, \quad (17)$$

where $r=1$. Taking into account (1) and (17), we obtain via $G_\theta(h)$ a lower simultaneous $(1-\alpha)$ prediction limit h on the kth order statistics $Y_{(k,m_j)}$, $j=1, \ldots, l$, of all of l future samples as

$$h = \arg\left[P_\theta\{ Y_{(k,m)} > h \} = \beta \right]$$

$$= \arg\left[G_\theta(h) = \frac{k}{k+(m-k+1)f_{2(m-k+1),2k;\beta}} \right]. \quad (18)$$

This ends the proof.

Corollary 2.1. If $k=1$ and $m_1 = m_2 = \ldots = m_l = m$, then

$$h = \arg\left[P_\theta\left\{ Y_{(1,m_1)} > h, \ldots, Y_{(1,m_j)} > h, \ldots, Y_{(1,m_l)} > h \right\} = 1-\alpha \right]$$

$$= \arg\left[G_\theta(h) = \frac{1}{1 + mlf_{2ml,2;1-\alpha}} \right]. \qquad (19)$$

Proof.

$$P_\theta\left\{ Y_{(1,m_1)} > h, \ldots, Y_{(1,m_j)} > h, \ldots, Y_{(1,m_l)} > h \right\}$$

$$= \prod_{j=1}^{l} P_\theta\left\{ Y_{(1,m_j)} > h \right\} = \left[P_\theta\{ Y_{(1,m)} > h \} \right]^l$$

$$= \left[1 - P_\theta\{ Y_{(1,m)} \le h \} \right]^l$$

$$= \left[1 - \sum_{j=1}^{m} \binom{m}{j} [G_\theta(h)]^j [1 - G_\theta(h)]^{m-j} \right]^l$$

$$= \left[[1 - G_\theta(h)]^m \right]^l = [1 - G_\theta(h)]^{ml}$$

$$= 1 - \sum_{j=1}^{ml} \binom{ml}{j} [G_\theta(h)]^j [1 - G_\theta(h)]^{ml-j}$$

$$= 1 - \frac{1}{\mathrm{B}(1,ml)} \int_0^{G_\theta(h)} u^0 (1-u)^{ml-1} du$$

$$= 1 - P_\theta\left\{ F_{2ml,2} > \frac{1 - G_\theta(h)}{G_\theta(h)} \frac{2}{2ml} \right\}$$

$$= P_\theta\left\{ F_{2ml,2} \le \frac{1 - G_\theta(h)}{G_\theta(h)} \frac{2}{2ml} \right\} = P_\theta\{ F_{2ml,2} \le f_{2ml,2} \}, \quad (20)$$

where

$$G_\theta(h) = \frac{1}{1 + mlf_{2ml,2}}. \qquad (21)$$

This ends the proof.

Corollary 2.2. If there are l independent samples from the same known distribution (continuous or discrete) with density function $g_\theta(y)$, distribution function $G_\theta(y)$, then an upper simultaneous α prediction limit h on the kth order statistics $Y_{(k,m_j)}$, $j=1, \ldots, l$, of all of l future samples with m_j $=m$ is given by

$$h = \arg\left[P_\theta\left\{ Y_{(k,m_1)} > h, \ldots, Y_{(k,m_j)} > h, \ldots, Y_{(k,m_l)} > h \right\} = \alpha \right]$$

$$= \arg\left[P_\theta\{ Y_{(k,m)} > h \} = \frac{(l-r+1)f_{2(l-r+1),2r;\alpha}}{r + (l-r+1)f_{2(l-r+1),2r;\alpha}} = \beta_1 \right]$$

$$= \arg\left[G_\theta(h) = \frac{k}{k + (m-k+1)f_{2(m-k+1),2k;\beta_1}} \right]. \qquad (22)$$

III. THE INVERSE GAUSSIAN DISTRIBUTION

A random variable X is distributed as inverse Gaussian, denoted as $X \sim \mathrm{IG}(\mu, \lambda)$, if its probability density function is given by

$$g_{\mu,\lambda}(x) = \left(\frac{\lambda}{2\pi x^3} \right)^{1/2} \exp\left(-\frac{\lambda(x-\mu)^2}{2\mu^2 x} \right), \quad x > 0, \qquad (23)$$

where parameters μ and λ are assumed to be positive; μ is the mean of the distribution and λ is a shape parameter. The probability distribution function is given by

$$G_{\mu,\lambda}(x) = \Phi\left[\sqrt{\frac{\lambda}{x}} \left(\frac{x}{\mu} - 1 \right) \right]$$

$$+ \exp\left(\frac{2\lambda}{\mu} \right) \Phi\left[-\sqrt{\frac{\lambda}{x}} \left(1 + \frac{x}{\mu} \right) \right], \qquad (24)$$

where Φ stands for the cumulative standard normal distribution function, and

$$\overline{G}_{\mu,\lambda}(x) = 1 - G_{\mu,\lambda}(x) = \Phi\left[\sqrt{\frac{\lambda}{x}} \left(1 - \frac{x}{\mu} \right) \right]$$

$$- \exp\left(\frac{2\lambda}{\mu} \right) \Phi\left[-\sqrt{\frac{\lambda}{x}} \left(1 + \frac{x}{\mu} \right) \right]. \qquad (25)$$

When both parameters μ and λ are known, the inverse Gaussian distribution function can be evaluated using the normal distribution table.

A. Estimation of Unknown Parameters

Let $X_1, X_2, \ldots, X_n$ be a random sample from (23). Tweedie [14] showed that the maximum likelihood estimates (MLE's) of μ and λ are

$$\hat{\mu} = \overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad (26)$$

and

$$\hat{\lambda}^{-1} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{X_i} - \frac{1}{\overline{X}} \right). \qquad (27)$$

which are independently distributed with $\overline{X} \sim \mathrm{IG}(\mu, n\lambda)$ and $n\lambda/\hat{\lambda} \sim \chi^2_{n-1}$, where χ^2_v is a chi-squared random variable with v degrees of freedom. Furthermore $(\hat{\mu}, n\hat{\lambda}^{-1})$ is a complete sufficient statistic for (μ, λ).

The MLE of the probability distribution function $G_{\mu,\lambda}(h)$ is now obtained by replacing μ and λ in (24) by their estimates $\hat{\mu}$ and $\hat{\lambda}$ given in (26) and (27), respectively. Thus, the MLE of $G_{\mu,\lambda}(h)$ is

$$G_{\mu,\lambda}^{\mathrm{MLE}}(h) = \Phi\left[\sqrt{\frac{\hat{\lambda}}{h}} \left(\frac{h}{\hat{\mu}} - 1 \right) \right]$$

$$+ \exp\left(\frac{2\hat{\lambda}}{\hat{\mu}} \right) \Phi\left[-\sqrt{\frac{\hat{\lambda}}{h}} \left(1 + \frac{h}{\hat{\mu}} \right) \right]. \qquad (28)$$

Accordingly, the MLE of $\overline{G}_{\mu,\lambda}(h)$ is

$$\overline{G}_{\mu,\lambda}^{\mathrm{MLE}}(h) = 1 - G_{\mu,\lambda}^{\mathrm{MLE}}(h) = \Phi\left[\sqrt{\frac{\hat{\lambda}}{h}}\left(1-\frac{h}{\hat{\mu}}\right)\right]$$

$$- \exp\left(\frac{2\hat{\lambda}}{\hat{\mu}}\right)\Phi\left[-\sqrt{\frac{\hat{\lambda}}{h}}\left(1+\frac{h}{\hat{\mu}}\right)\right]. \qquad (29)$$

Chhikara and Folks [15] have given the minimum variance unbiased estimate (MVUE) of the inverse Gaussian distribution for different cases that might arise regarding μ and λ. When both of these parameters are unknown, it follows from their result that the MVUE of $G_{\mu,\lambda}(h)$ is given by

$$G_{\mu,\lambda}^{\mathrm{MVUE}}(h)$$

$$= \begin{cases} 0, & h < L \\ 1, & h > U \\ T_{n-2}(w_1) + \frac{n-2}{n}\left[1+\frac{4(n-1)}{nv\overline{x}}\right]^{(n-3)/2} T_{n-2}(-w_2), & \text{otherwise}, \end{cases}$$

$$\qquad (30)$$

and the MVUE of $\overline{G}_{\mu,\lambda}(h)$ is given by

$$\overline{G}_{\mu,\lambda}^{\mathrm{MVUE}}(h)$$

$$= \begin{cases} 1, & h < L \\ 0, & h > U \\ T_{n-2}(-w_1) - \frac{n-2}{n}\left[1+\frac{4(n-1)}{nv\overline{x}}\right]^{(n-3)/2} T_{n-2}(-w_2), & \text{otherwise} \end{cases}$$

$$\qquad (31)$$

where T_{n-2} denotes the cdf of the Student's t distribution with $(n-2)$ degrees of freedom and

$$v = \sum_{i=1}^{n}\left(\frac{1}{x_i}-\frac{1}{\overline{x}}\right), \qquad (32)$$

$$w_1 = \frac{\sqrt{n(n-2)}(h-\overline{x})}{\sqrt{v\overline{x}(n\overline{x}-h)h-n(h-\overline{x})^2}}, \qquad (33)$$

$$w_2 = \frac{\sqrt{n(n-2)}[\overline{x}+(n-2)h/n]}{\sqrt{v\overline{x}(n\overline{x}-h)h-n(h-\overline{x})^2}}, \qquad (34)$$

$$L = \frac{\overline{x}\left[n(2+v\overline{x})-\sqrt{4n(n-1)v\overline{x}+(nv\overline{x})^2}\right]}{2(n+v\overline{x})}, \qquad (35)$$

$$U = \frac{\overline{x}\left[n(2+v\overline{x})+\sqrt{4n(n-1)v\overline{x}+(nv\overline{x})^2}\right]}{2(n+v\overline{x})}. \qquad (36)$$

IV. NUMERICAL EXAMPLE

An industrial firm has the policy to replace a certain device, used at several locations in its plant, at the end of 500-hour intervals. It doesn't want too many of these items to fail before being replaced. Shipments of a lot of devices are made to each of three firms ($l=3$). Each firm selects a random sample of $m=2$ items and accepts his shipment if no failures occur before a specified lifetime has accumulated. The manufacturer wishes to take a random sample and to calculate the lower prediction limit so that all shipments will be accepted with a probability of $1-\alpha = 0.95$. The resulting lifetimes (in terms of 100 hours) of an initial sample of size $n=46$ from a population of such devices are given in Table 1.

TABLE I
THE RESULTING LIFETIMES (IN TERMS OF 100 HOURS)

Observations							
X_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7
Lifetime	0.2	0.3	0.5	0.5	0.5	0.5	0.6
	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
	0.6	0.7	0.7	0.7	0.8	0.8	1.0	1.0
	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}	x_{21}	x_{22}	x_{23}
	1.0	1.0	1.1	1.3	1.5	1.5	1.5	1.5
	x_{24}	x_{25}	x_{26}	x_{27}	x_{28}	x_{29}	x_{30}	x_{31}
	2.0	2.0	2.2	2.5	2.7	3.0	3.0	3.3
	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}	x_{37}	x_{38}	x_{39}
	3.3	4.0	4.0	4.5	4.7	5.0	5.4	5.4
	x_{40}	x_{41}	x_{42}	x_{43}	x_{44}	x_{45}	x_{46}
	7.0	7.5	8.8	9.0	10.3	22.0	24.5

Goodness-of-fit testing. It is assumed that

$$g_{\mu,\lambda}(x_i) = \left(\frac{\lambda}{2\pi x_i^3}\right)^{1/2}\exp\left(-\frac{\lambda(x_i-\mu)^2}{2\mu^2 x_i}\right)$$

$$(x_i,\mu,\lambda>0), \quad i=1(1)46, \qquad (37)$$

where the parameters μ and λ are unknown.

Considering the inverse Gaussian model for the data, unbiased estimates of μ and λ^{-1} are

$$\hat{\mu} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n}X_i = 3.61, \qquad (38)$$

$$\hat{\lambda}^{-1} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{1}{X_i}-\frac{1}{\overline{X}}\right) = 0.60. \qquad (39)$$

The maximum likelihood estimate of variance is given by

$$\hat{\sigma}^2 = \hat{\mu}^3/\hat{\lambda} = 28.33. \qquad (40)$$

The inverse Gaussian distribution with mean 3.61 and variance 28.33 provides a good fit to the above lifetime data, and so does the log normal distribution. The calculated value of the Kolmogorov-Smirnov test statistic is 0.0526 for the

inverse Gaussian and 0.0807 for the log normal, and these are smaller than their corresponding values expected at the five percent significance level. (The K-S test is considered in the sense that the unknown parameters of a distribution are estimated from the sample data.)

Now from (29) and (31) one is able to obtain the MLE and MVUE of $\overline{G}_\theta(x)$, respectively. These estimates are given in Fig. 1 along with the MLE of $\overline{G}_\theta(x)$ obtained from the log normal.

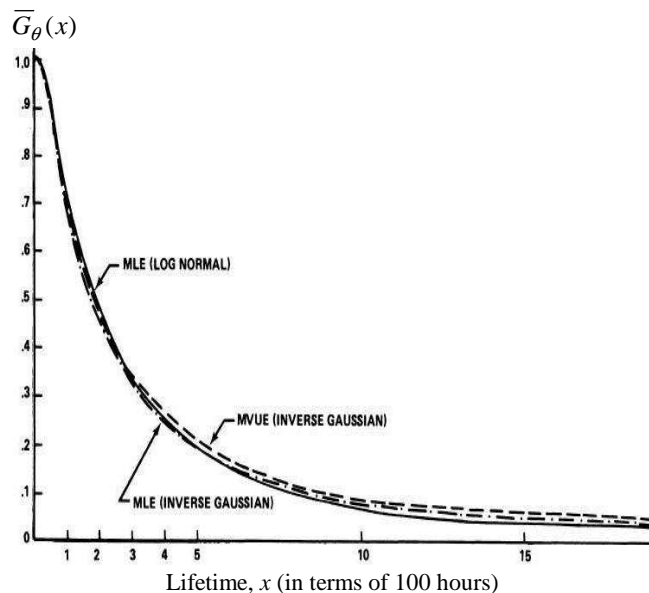**Fig. 1.** Estimates of Reliability Function $\overline{G}_\theta(x)$.

Prediction limit. A lower $(1-\alpha) = 0.95$ prediction limit h on the first order statistic Y_1 in a set of $ml = 6$ future ordered observations $Y_1 \le \dots \le Y_{ml}$ is given by

$$h = \arg[P_{\mu,\lambda}\{Y_{(1,ml)} > h\} = 1 - \alpha]$$

$$= \arg\left[G_{\mu,\lambda}^{\mathrm{MLE}}(h) = \frac{1}{1 + ml f_{2ml,2;1-\alpha}} \right]$$

$$= \arg\left[G_{\mu,\lambda}^{\mathrm{MLE}}(h) = \frac{1}{1 + 6 \times 19.4} = 0.0085 \right] = 0.22, \quad (41)$$

where

$$G_{\mu,\lambda}^{\mathrm{MLE}}(h) = \Phi\left[\sqrt{\frac{\hat{\lambda}}{h}}\left(\frac{h}{\hat{\mu}} - 1 \right) \right]$$

$$+ \exp\left(\frac{2\hat{\lambda}}{\hat{\mu}} \right)\Phi\left[-\sqrt{\frac{\hat{\lambda}}{h}}\left(1 + \frac{h}{\hat{\mu}} \right) \right] = 0.0085. \quad (42)$$

Thus, the manufacturer has 95% assurance that no failures will occur in each shipment before $h = 22$ hours.

V. Conclusion and Future Work

In this paper, we propose the technique of constructing prediction limits on future order statistics coming from the inverse Gaussian distribution under parametric uncertainty. These prediction limits are based on a previously available complete sample from the same distribution. We present an equation for this type of prediction limits which holds for any distribution and any statistic from the previous sample when a prediction limit for a single future sample is available. The prediction limits are found and illustrated with a numerical example. The methodology described here can be extended in several different directions to handle various problems that arise in practice. We have illustrated the proposed methodology for the inverse Gaussian distribution. Application to other distributions could follow directly.

References

[1] R. S. Chhikara and L. S. Folks, *The Inverse Gaussian Distribution. Theory, Methodology and Application*, New York: Marcel Dekker, 1989.
[2] J. L. Folks and R. S. Chhikara, "The inverse Gaussian distribution and its statistical applications - a review," *Journal of the Royal Statistical Society*, Ser. B, vol. 40, pp. 263 –289, 1978.
[3] B. Dennis, P. L. Munholland, and M. J. Scott, "Estimation of growth and extinction parameters for endangered species," *Ecological Monographs*, vol. 61, pp. 115 –143, 1991.
[4] R. S. Chhikara and L. S. Folks, "The inverse Gaussian distribution as a lifetime model," *Technometrics*, vol. 19, pp. 461– 468, 1977.
[5] A. K. Banerjee and G. K. Bhattacharyya, (1976), "A purchase incidence model with inverse Gaussian interpurchase times," *Journal of the American Statistical Association*, vol. 71, pp. 823 – 829, 1976.
[6] A. Lancaster, "A stochastic model for the duration of a strike," *Journal of the Royal Statistical Society*, Ser. A, vol. 135, pp. 257– 271, 1972.
[7] G. A. Whitmore, "Management applications of the inverse Gaussian distributions," *International Journal of Management Science*, vol. 4, pp. 215 –223, 1976.
[8] J. Aitchison, "Bayesian tolerance region," *Journal of the Royal Statistical Society*, Ser. B, vol. 26, pp. 161–175, 1964.
[9] G. J. Hahn, "Simultaneous prediction intervals to contain the standard deviations or ranges of future samples from a normal distribution," *Journal of the American Statistical Association*, vol. 67, pp. 938–942, 1972.
[10] G. J. Hahn, "A prediction interval on the means of future samples from an exponential distribution," *Technometrics*, vol. 17, pp. 341– 345, 1975.
[11] G. J. Hahn and W. Nelson, "A survey of prediction intervals and their applications," *Journal of Quality Technology*, vol. 5, pp. 178–188, 1973.
[12] N. R. Mann, R. E. Schafer, and J. D. Singpurwalla, *Methods for Statistical Analysis of Reliability and Life Data*. New York: John Wiley, 1974.
[13] K. W. Fertig and N. R. Mann, "One-sided prediction intervals for at least p out of m future observations from a normal population," *Technometrics*, vol. 19, pp. 167–177, 1977.
[14] M. C. K. Tweedie, "Statistical properties of inverse Gaussian distributions I," *Annals of Mathematical Statistics*, vol. 28, pp. 362– 377, 1957.
[15] R. S. Ghhikara and J. L. Folks, "Estimation of the inverse Gaussian distribution function", *J. Amer. Statist. Assoc.*, vol. 69, pp 250–254, 1974.