

Predictive Densities for the Lognormal Distribution and Their Applications

Zhenlin Yang

Department of Statistics and Applied Probability
3 Science Drive 2
National University of Singapore
Singapore 117543
email: stayzl@nus.edu.sg
phone: (65) 874-6829 fax: (65) 872-3919

Abstract: Maximum likelihood predictive densities (MLPD) for a future lognormal observation are obtained and their applications to reliability and life testing are considered. When applied to reliability and failure rate estimations, they give estimators that can be much less biased and less variable than the usual MLEs obtained by replacing the unknown parameters in the density function by their MLEs. When applied to lifetime predictions, they give prediction intervals that are shorter than the usual frequentist intervals. Using the MLPDs, it is also rather convenient to construct the shortest prediction intervals. Extensive simulations are performed for comparisons. A numerical example is given for illustration.

Key words and phrases: Failure rate; lognormal distribution; maximum likelihood predictive density; prediction interval; reliability; shortest prediction interval.

1 Introduction

The lognormal distribution may be one of the most versatile distributions. It has been seen to have applications in many fields, such as agriculture, entomology, economics, geology, industry and quality control. See Crow and Shimizu [1] or Johnson and Kolz [4, Ch14] for a complete account of the lognormal distribution. In terms of life testing and reliability, the lognormal distribution is known as a serious competitor to the Weibull distribution. A random variable (rv) X is said to have a lognormal distribution if $\log X$ is distributed as normal with mean μ and standard deviation σ . Hence, its probability density function (pdf) is of the form:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma x} \exp \left\{ -\frac{1}{2\sigma^2} (\log x - \mu)^2 \right\}, x > 0; -\infty < \mu < \infty, \sigma > 0. \quad (1)$$

Letting $\tau = \exp(\sigma^2)$, the mean, variance, skewness and kurtosis of the lognormal distribution are, respectively, e^μ , $e^{2\mu}\tau(\tau - 1)$, $(\tau - 1)^{1/2}(\tau + 2)$ and $\tau^4 + 2\tau^3 + 3\tau^2 - 6$. From the skewness expression, it is clear that the population skewness γ can go up very fast as σ increases. For examples, when $\sigma = 1$, we have $\gamma = 6.18$, and when $\sigma = 2$, we have $\gamma = 414.4$. Hence, from a practical point of view it is very unlikely to have $\sigma > 2$ for the lognormal distribution.

As the logarithm of a lognormal random variable has a normal distribution, many statistical methods, such as methods for computing statistical intervals, for the normal distribution can be directly translated to suit for the lognormal distribution [3, p207]. Problems related to reliability and failure rate of the lognormal have been considered by Rohn [7] and Sweet [9], among the others. Dahiya and Guttman [2] provided shortest confidence and prediction intervals for this distribution. The Bayesian approach has also been used for estimating the mean of the lognormal [10], estimating the reliability function [6], and constructing prediction intervals [8], etc.

In this article, the maximum likelihood predictive densities (MLPD) for a future lognormal observation based on an independent past sample are derived. Based

on these MLPDs, one can easily obtain estimators for the reliability and failure rate functions and construct prediction interval (PI) and shortest prediction interval (SPI) for a future observation. Yang [11] derived MLPDs for the inverse Gaussian distribution and applied them to reliability and lifetime predictions. He showed that the MLPD estimators of reliability possess smaller bias and mean squared error (MSE) than the usual MLEs. The PIs obtained through the MLPDs can have much shorter lengths than the usual frequentist PIs while preserving good coverage properties.

Section 2 derives the MLPDs for the lognormal distribution for the cases of one or both parameters unknown. Section 3 discusses the applications of these MLPDs to reliability and life testing studies, including the reliability and failure rate estimations, and PI and SPI constructions. Extensive simulations are performed to compare the MLPD estimators of reliability and failure rate with the usual MLEs obtained by simply replacing the parameters in the pdf by their MLEs. A numerical example is provided in Section 4 for illustration. Finally, a discussion is given in Section 5.

Compared with the usual maximum likelihood method, the MLPD method gives reliability estimators with smaller bias and MSE, and failure rate estimators with much smaller bias and MSE. This latter finding is quite promising as an accurate estimation of the failure mechanism is very important to the reliability practitioners. The MLPD method also gives prediction intervals with shorter length compared to the usual frequentist PI though with slightly lesser coverage. One advantage of the MLPD approach in constructing PIs is that it allows us to obtain shortest PIs in a rather simple manner.

2 The MLPDs for the Lognormal Distribution

Let $\mathbf{X} = (X_1, X_2, \dots, X_n)$ be a sample of past observations from a lognormal population with pdf specified in (1). Let Y or X_{n+1} , whichever convenient, be a future

observation from the same population. According to Lejeune and Faulkenberry [5], the maximum likelihood predictive density (MLPD) of Y given $\mathbf{X} = \mathbf{x}$ is defined as,

$$\tilde{f}(y | \mathbf{x}) \propto \max_{\theta} f(y; \theta) f(\mathbf{x}; \theta), \quad (2)$$

where $\theta = (\mu, \sigma)$, or μ , or σ , representing the unknown parameter(s) of interest.

The joint pdf of \mathbf{X} and Y can be written as

$$f(y; \mu, \sigma) f(\mathbf{x}; \mu, \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^{n+1}} \left(\prod_{i=1}^{n+1} \frac{1}{x_i} \right) \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n+1} (\log x_i - \mu)^2 \right] \quad (3)$$

Maximizing (3) with respect to μ and σ , or μ only, or σ only, gives the MLPDs of Y for the cases of both parameters unknown or one parameter unknown, respectively. For the lognormal distribution, it is well known that the MLEs of μ and σ based on the past sample are

$$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^n \log X_i \quad \text{and} \quad \hat{\sigma}_n = \frac{1}{n} \sum_{i=1}^n (\log X_i - \hat{\mu}_n)^2,$$

and the restricted MLEs of μ for a given σ and of σ for a given μ are, respectively,

$$\hat{\mu}_n(\sigma) = \hat{\mu}_n \quad \text{and} \quad \hat{\sigma}_n(\mu) = \frac{1}{n} \sum_{i=1}^n (\log X_i - \mu)^2.$$

We now derive the MLPDs by starting with the most practical case.

Case of both parameters unknown. Let $\hat{\mu}_{n+1}$ and $\hat{\sigma}_{n+1}$ denote the MLEs of μ and σ based on all the $n + 1$ observations. The joint pdf (3) achieves its maximum at $\hat{\mu}_{n+1}$ and $\hat{\sigma}_{n+1}$. Simple algebra shows that

$$\begin{aligned} \hat{\mu}_{n+1} &= \frac{1}{n+1} \sum_{i=1}^{n+1} \log X_i = \frac{1}{n+1} \log Y + \frac{n}{n+1} \hat{\mu}_n, \\ \hat{\sigma}_{n+1} &= \frac{1}{n+1} \sum_{i=1}^{n+1} (\log X_i - \hat{\mu}_{n+1})^2 = \frac{1}{n+1} \left[\frac{n}{n+1} (\log Y - \hat{\mu}_n)^2 + n \hat{\sigma}_n^2 \right] \end{aligned}$$

Substituting $\hat{\mu}_{n+1}$ and $\hat{\sigma}_{n+1}$ into (3), we have

$$f(y | \mathbf{x}) \propto \max_{\mu, \sigma} f(y; \mu, \sigma) f(\mathbf{x}; \mu, \sigma)$$

$$\begin{aligned}
 &\propto \frac{1}{(\sqrt{2\pi}\hat{\sigma}_{n+1})^{n+1}} \left(\prod_{i=1}^{n+1} \frac{1}{x_i} \right) \exp \left[-\frac{1}{2\hat{\sigma}_{n+1}^2} \sum_{i=1}^{n+1} (\log x_i - \hat{\mu}_{n+1})^2 \right] \\
 &\propto \frac{1}{x_{n+1}} (\hat{\sigma}_{n+1}^2)^{-(n+1)/2} \\
 &\propto \frac{1}{y} \left[\frac{n}{n+1} (\log y - \hat{\mu}_n)^2 + n\hat{\sigma}_n^2 \right]^{-(n+1)/2} \\
 &\propto \frac{1}{y} \left[1 + \frac{(\log y - \hat{\mu}_n)^2}{(n+1)\hat{\sigma}_n^2} \right]^{-(n+1)/2},
 \end{aligned}$$

Integrating the last expression (through a t -integration) gives the normalizing constant and hence the final form of the MLPD for Y :

$$\tilde{f}(y | \mathbf{x}) = \frac{\Gamma[(n+1)/2]}{\sqrt{(n+1)\pi\hat{\sigma}_n}\Gamma(n/2)} \frac{1}{y} \left[1 + \frac{(\log y - \hat{\mu}_n)^2}{(n+1)\hat{\sigma}_n^2} \right]^{-(n+1)/2} \quad (4)$$

Case of only μ unknown. When σ is known but μ is unknown, it is easy to see that (3) is maximized at $\hat{\mu}_{n+1}$, the restricted MLE of μ based on all $n+1$ observations. Using the identity $\sum_{i=1}^{n+1} (\log x_i - \hat{\mu}_{n+1})^2 = \frac{n}{n+1} (\log y - \hat{\mu}_n)^2 + n\hat{\sigma}_n^2$, some simple algebra gives the MLPD for this case:

$$\tilde{f}(y | \mathbf{x}; \sigma) = \frac{\sqrt{n}}{\sqrt{2(n+1)\pi\sigma}} \frac{1}{y} \exp \left[-\frac{n}{2(n+1)\sigma^2} (\log y - \hat{\mu}_n)^2 \right] \quad (5)$$

Case of only σ unknown. For the last case where μ is known but σ is unknown, the maximum of (3) happens at $\hat{\sigma}_{n+1}^2(\mu) = \frac{1}{n+1} \sum_{i=1}^{n+1} (\log x_i - \mu)^2$, the restricted MLE of σ for a given μ based on all $n+1$ observations. Similar algebraic work as in the case of both parameters unknown gives the MLPD:

$$\tilde{f}(y | \mathbf{x}; \mu) = \frac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\hat{\sigma}_n(\mu)\Gamma(n/2)} \frac{1}{y} \left[1 + \frac{(\log y - \mu)^2}{n\hat{\sigma}_n^2(\mu)} \right]^{-(n+1)/2} \quad (6)$$

Notice that the MLPD $\tilde{f}(y | \mathbf{x}; \sigma)$ for the σ known case is a lognormal and hence its properties are well known. In particular, it is unimodal. The MLPDs $\tilde{f}(y | \mathbf{x})$ and $\tilde{f}(y | \mathbf{x}; \mu)$ for the other two cases have an identical structure

$$g(y) = \frac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\sigma^* \Gamma(n/2)} \frac{1}{y} \left[1 + \frac{(\log y - \mu^*)^2}{n\sigma^{*2}} \right]^{-(n+1)/2} \quad (7)$$

which is termed by Dahiya and Guttman as the log- t distribution with n degrees of freedom and with parameters μ^* and σ^* . Their properties are summarized in the following theorem.

Theorem 1. Let Y be a random variable having a log- t distribution with pdf $g(y)$ given in (7). We have,

- i) $g(y)$ is a decreasing function of y if $\sigma^* \geq (n + 1)/2\sqrt{n}$,
and has two stationary points y_1 and y_2 if $\sigma^* < (n + 1)/2\sqrt{n}$;
- ii) the two stationary points are, respectively, the local minima and maxima;
- iii) the probability $P(Y < y_1)$ is negligible if n is large relative to σ^* .

Proof. Part i) of the theorem is from Dahiya and Guttman. Part ii) can be proved by showing that $d^2g(y)/dy^2 > 0$ at y_1 and < 0 at y_2 which is easy following the result of Dahiya and Guttman. To prove iii), notice that

$$c_1 = \frac{\log y_1 - \mu^*}{\sigma^*} = -\frac{1}{2\sigma^*} \left[n + 1 + \sqrt{(n + 1)^2 - 4n\sigma^*} \right]$$

Hence $P(Y < y_1) = P(t_n < c_1)$ is negligible if n is large relative to σ^* , where t_n denotes a t random variable with n degrees of freedom.

The part ii) of the theorem simply says that as y moves from 0 to ∞ , $g(y)$ first decreases down to the local minimum, then increases up to the local maximum, and then decreases again down to zero. The last part of the theorem tells that the MLPDs $\tilde{f}(y | \mathbf{x})$ and $\tilde{f}(y | \mathbf{x}; \mu)$ are essentially unimodal as in practice n is usually larger than 10 and σ is smaller than 2.

Given the MLPDs (4)-(6), it is easy to show, through t - or normal integrations, that the cumulative MLPDs are:

$$\tilde{F}(y | \mathbf{x}) = \Psi_n \left(\frac{\log t - \hat{\mu}_n}{\sqrt{1 + n^{-1}\hat{\sigma}_n^2}} \right) \quad (8)$$

$$\tilde{F}(y | \mathbf{x}; \sigma) = \Phi \left(\frac{\log t - \hat{\mu}_n}{\sqrt{1 + n^{-1}}\sigma} \right) \quad (9)$$

$$\tilde{F}(y | \mathbf{x}; \mu) = \Psi_n \left(\frac{\log t - \mu}{\hat{\sigma}_n(\mu)} \right) \quad (10)$$

where Ψ_n denotes the cumulative distribution function (CDF) of a t distribution with n degrees of freedom, and Φ denotes the CDF of the standard normal distribution.

3 Applications to Reliability and Life Testing

The MLPDs can have many applications in studying the problems related to reliability and life testing, such as reliability estimation, failure rate estimate, constructing prediction intervals, etc. No doubt, the simple forms of the MLPDs and cumulative MLPDs derived earlier make these studies handy.

3.1 Reliability estimation

The reliability function of the lognormal distribution is seen to be of the form:

$$R(t) = 1 - \Phi[(\log t - \mu)/\sigma]$$

The simplest method for estimating $R(t)$ may be the maximum likelihood estimation (MLE) method where the unknown parameter(s) are replaced by their MLEs in the expression of $R(t)$. Thus, the MLEs for $R(t)$ are,

$$\hat{R}(t) = 1 - \Phi[(\log t - \hat{\mu}_n)/\hat{\sigma}_n]$$

$$\hat{R}(t, \sigma) = 1 - \Phi[(\log t - \hat{\mu}_n)/\sigma]$$

$$\hat{R}(t, \mu) = 1 - \Phi[(\log t - \mu)/\hat{\sigma}_n(\mu)]$$

respectively for the cases of i) both parameters unknown, ii) only μ unknown, and iii) only σ unknown.

Now, considering the MLPDs derived in Section 2 as estimators of the unknown pdf, we propose the MLPD estimators of the reliability function $R(t)$ as

$$\begin{aligned}\tilde{R}(t) &= 1 - \tilde{F}(t | \mathbf{x}) \\ \tilde{R}(t, \sigma) &= 1 - \tilde{F}(t | \mathbf{x}, \sigma) \\ \tilde{R}(t, \mu) &= 1 - \tilde{F}(t | \mathbf{x}, \mu),\end{aligned}$$

respectively, for the three cases.

A comparison for both parameter unknown case. Clearly, comparing two reliability estimators is one of the major themes of this paper. Since the case of both parameters unknown is the most realistic case, we thus concentrate on this situation for comparisons. Comparisons for the other cases can be carried out in a similar way. An analytical comparison may be difficult, we thus turn to Monte Carlo simulations to perform this task. For each of the sample size chosen, various combinations of the parameter values are considered to give populations ranging from moderately skewed to very skewed. For each parameter configuration, a number of 'time' points are considered that cover a large portion of the population values. Each row of the simulation results are based on 10,000 samples generated from a specified lognormal population. Table 1 lists the results for $n = 10$ and 20. The results show that the MLPD estimator is generally less biased and less variable than the MLE, especially when the population is very skewed. The bias of the MLPD estimator can be much smaller than that of the MLE, especially when the population is very skewed and the time point is within the middle 80% of the population. For example, for the parameter configuration $(n, \mu, \sigma) = 10, 1.0, 2.0$, the bias at $t = 3.5$ is -0.0038 for the MLE, but is only -0.0001 for the MLPD estimator. Increasing the sample size reduces the discrepancies between the two methods.

All the simulations are carried out using F90 on a mainframe machine in the

National University of Singapore. The IMSL subroutines ANORDF and TDF are used for calculating the CDFs of the standard normal and t distributions. The Fortran code is available from the author upon request.

3.2 Failure rate estimation

The failure rate function $r(t)$ for the lognormal distribution has the form:

$$r(t) = \frac{(\sqrt{2\pi}\sigma x)^{-1} \exp[-(\log x - \mu)^2/(2\sigma^2)]}{1 - \Phi[(\log t - \mu)/\sigma]}.$$

Its MLEs of are

$$\begin{aligned} \hat{r}(t) &= \frac{(\sqrt{2\pi}\hat{\sigma}_n t)^{-1} \exp[-(\log t - \hat{\mu}_n)^2/(2\hat{\sigma}_n^2)]}{1 - \Phi[(\log t - \hat{\mu}_n)/\hat{\sigma}_n]}, \\ \hat{r}(t, \sigma) &= \frac{(\sqrt{2\pi}\sigma t)^{-1} \exp[-(\log t - \hat{\mu}_n)^2/(2\sigma^2)]}{1 - \Phi[(\log t - \hat{\mu}_n)/\sigma]}, \\ \hat{r}(t, \mu) &= \frac{(\sqrt{2\pi}\hat{\sigma}_n(\mu)t)^{-1} \exp[-(\log t - \mu)^2/(2\hat{\sigma}_n^2(\mu))]}{1 - \Phi[(\log t - \mu)/\hat{\sigma}_n(\mu)]}, \end{aligned}$$

respectively, for the cases of both parameters unknown, only μ unknown and only σ unknown. The corresponding MLPD estimators of $r(t)$ are given by:

$$\begin{aligned} \tilde{r}(t) &= \frac{k_n(\mathbf{x})t^{-1}[1 + (\log t - \hat{\mu}_n)^2/((n+1)\hat{\sigma}_n^2)]^{-(n+1)/2}}{1 - \Psi_n[(\log t - \hat{\mu}_n)/(\sqrt{1+n^{-1}}\hat{\sigma}_n)]}, \\ \tilde{r}(t, \sigma) &= \frac{k_n(\mathbf{x}, \sigma)t^{-1} \exp[-n(\log t - \hat{\mu}_n)^2/(2(n+1)\sigma^2)]}{1 - \Phi[(\log t - \hat{\mu}_n)/(\sqrt{1+n^{-1}}\sigma)]}, \\ \tilde{r}(t, \mu) &= \frac{k_n(\mathbf{x}, \mu)t^{-1}[1 + (\log t - \mu)^2/(n\hat{\sigma}_n^2(\mu))]^{-(n+1)/2}}{1 - \Psi_n[(\log t - \mu)/\hat{\sigma}_n(\mu)]}, \end{aligned}$$

where $k_n(\mathbf{x})$, $k_n(\mathbf{x}, \sigma)$ and $k_n(\mathbf{x}, \mu)$ are the normalizing constants given in (4)-(6).

A comparison for both parameters unknown case. Again, we concentrate on the case of both parameters unknown and use Monte Carlo simulations to compare the performance of the two estimators of the failure rate. Similar parameter configurations as in the reliability estimation case are used. The results are reported in Table 2 for $n = 10$ and 20 .

Table 1: Simulation Results for Reliability Estimators

n	(μ, σ)	t	$R(t)$	Bias		MSE	
				MLE	MLPD	MLE	MLPD
10	(1.0, 0.5)	1.5	0.8828	0.0074	-0.0144	0.0059	0.0060
		2.0	0.7303	0.0100	-0.0059	0.0140	0.0124
		2.5	0.5665	0.0046	-0.0003	0.0193	0.0168
		3.5	0.3066	-0.0118	0.0019	0.0154	0.0135
		4.0	0.2199	-0.0121	0.0065	0.0113	0.0102
		4.5	0.1567	-0.0101	0.0110	0.0081	0.0077
	(1.0, 1.0)	1.0	0.8413	0.0104	-0.0107	0.0083	0.0078
		2.0	0.6205	0.0106	0.0017	0.0179	0.0155
		3.0	0.4607	-0.0030	-0.0001	0.0196	0.0170
		4.0	0.3496	-0.0091	0.0017	0.0173	0.0150
		5.0	0.2711	-0.0138	0.0021	0.0137	0.0121
		6.5	0.1917	-0.0106	0.0093	0.0102	0.0093
	(1.0, 1.5)	0.5	0.8705	0.0090	-0.0127	0.0066	0.0065
		1.5	0.6541	0.0090	-0.0020	0.0166	0.0145
		2.5	0.5223	0.0020	0.0004	0.0193	0.0167
		4.5	0.3684	-0.0059	0.0034	0.0171	0.0149
		6.0	0.2988	-0.0100	0.0040	0.0157	0.0138
		9.5	0.2021	-0.0114	0.0080	0.0106	0.0096
	(1.0, 2.0)	0.5	0.8014	0.0117	-0.0079	0.0104	0.0094
		1.5	0.6169	0.0061	-0.0023	0.0177	0.0154
		3.5	0.4497	-0.0038	-0.0001	0.0186	0.0162
6.5		0.3315	-0.0092	0.0027	0.0167	0.0146	
11.5		0.2354	-0.0125	0.0053	0.0125	0.0111	
15.0		0.1965	-0.0115	0.0082	0.0104	0.0095	
20	(1.0, 0.5)	1.5	0.8828	0.0039	-0.0080	0.0031	0.0031
		2.0	0.7303	0.0084	0.0000	0.0068	0.0063
		2.5	0.5665	0.0003	-0.0021	0.0085	0.0079
		3.5	0.3066	-0.0060	0.0011	0.0075	0.0069
		4.0	0.2199	-0.0066	0.0033	0.0055	0.0052
		4.5	0.1567	-0.0051	0.0062	0.0042	0.0041
	(1.0, 1.0)	1.0	0.8413	0.0052	-0.0062	0.0043	0.0041
		2.0	0.6205	0.0046	0.0001	0.0082	0.0076
		3.0	0.4607	-0.0004	0.0011	0.0088	0.0081
		4.0	0.3496	-0.0051	0.0004	0.0080	0.0074
		5.0	0.2711	-0.0063	0.0020	0.0068	0.0064
		6.5	0.1917	-0.0066	0.0041	0.0051	0.0048
	(1.0, 1.5)	0.5	0.8705	0.0050	-0.0068	0.0034	0.0034
		1.5	0.6541	0.0037	-0.0020	0.0080	0.0075
		2.5	0.5223	0.0024	0.0015	0.0085	0.0079
		4.5	0.3684	-0.0047	0.0002	0.0081	0.0075
		6.0	0.2988	-0.0053	0.0020	0.0072	0.0067
		9.5	0.2021	-0.0045	0.0058	0.0055	0.0052
	(1.0, 2.0)	0.5	0.8014	0.0071	-0.0034	0.0054	0.0051
		1.5	0.6169	0.0050	0.0007	0.0082	0.0077
		3.5	0.4497	-0.0019	-0.0001	0.0086	0.0080
6.5		0.3315	-0.0059	0.0003	0.0077	0.0072	
11.5		0.2354	-0.0059	0.0034	0.0061	0.0057	
15.0		0.1965	-0.0049	0.0056	0.0052	0.0049	

The results are rather surprising: the MLPD estimator performs much better than the MLE, almost uniformly across the time points. The bias and MSE of the MLPD estimator can be much smaller than those of the MLE. For example, at the parameter configuration $(n, \mu, \sigma, t) = (10, 1.0, 2.0, 15.0)$, the bias and MSE for the MLPD estimator are 0.0012 and 0.0003, whereas for the MLE, they are 0.0138 and 0.0012, respectively. It seems that the MLPD estimator performs even better at the two ends of the distribution. This is somehow in contrast to the MLPD estimator for reliabilities. Increasing the sample size reduces the discrepancy between the two estimators.

3.3 Prediction intervals

For the lognormal distribution, exact frequentist prediction intervals (PI) for the log of a future observation can easily be constructed based on the normal theory. Exponentiate these intervals give PIs for the future observation in the original scale. Let $Z_{\alpha/2}$ be the $100(\alpha/2)\%$ point of the standard normal distribution and $t_{\nu}(\alpha/2)$ be that for a t distribution with ν degrees of freedom. When both parameters are unknown, an exact $100(1 - \alpha)\%$ PI for Y is given by:

$$\exp \left\{ \hat{\mu}_n \pm t_{n-1}(\alpha/2) \hat{\sigma}_n \sqrt{(n+1)/(n-1)} \right\}. \quad (11)$$

When σ is known but μ is unknown, an $100(1 - \alpha)\%$ PI for Y has the form:

$$\exp \left\{ \hat{\mu}_n \pm Z_{\alpha/2} \sigma \sqrt{1 + n^{-1}} \right\}. \quad (12)$$

Finally, when μ is known but σ is unknown, an $100(1 - \alpha)\%$ PI for Y has the form:

$$\exp \left\{ \mu \pm t_n(\alpha/2) \hat{\sigma}_n(\mu) \sqrt{1 + n^{-1}} \right\}. \quad (13)$$

The corresponding PIs based on the MLPDs can be easily constructed. An equitailed PI for a future observation $Y > 0$ based on a predictive density $p(y)$, say, is

Table 2: Simulation Results for Failure Rate Estimators

n	(μ, σ)	t	$r(t)$	Bias		MSE	
				MLE	MLPD	MLE	MLPD
10	(1.0, 0.5)	1.5	0.2971	-0.0153	-0.0110	0.0248	0.0201
		2.0	0.4525	0.0271	0.0062	0.0482	0.0390
		2.5	0.5556	0.0806	0.0354	0.0936	0.0661
		3.5	0.6544	0.1638	0.0479	0.1862	0.0808
		4.0	0.6731	0.1946	0.0301	0.2439	0.0765
		4.5	0.6807	0.2111	0.0002	0.2834	0.0647
	(1.0, 1.0)	1.0	0.2876	-0.0041	-0.0071	0.0204	0.0168
		2.0	0.3067	0.0360	0.0149	0.0234	0.0178
		3.0	0.2872	0.0589	0.0264	0.0309	0.0182
		4.0	0.2647	0.0655	0.0234	0.0324	0.0151
		5.0	0.2444	0.0649	0.0154	0.0288	0.0108
		6.5	0.2190	0.0673	0.0072	0.0295	0.0078
	(1.0, 1.5)	0.5	0.3232	-0.0120	-0.0102	0.0285	0.0232
		1.5	0.2506	0.0250	0.0094	0.0155	0.0121
		2.5	0.2034	0.0341	0.0154	0.0132	0.0089
		4.5	0.1516	0.0374	0.0145	0.0105	0.0052
		6.0	0.1291	0.0331	0.0093	0.0075	0.0032
		9.5	0.0978	0.0288	0.0033	0.0054	0.0016
	(1.0, 2.0)	0.5	0.3479	0.0011	-0.0078	0.0274	0.0226
		1.5	0.2063	0.0256	0.0111	0.0112	0.0084
		3.5	0.1257	0.0260	0.0114	0.0059	0.0034
6.5		0.0842	0.0200	0.0062	0.0029	0.0014	
11.5		0.0568	0.0162	0.0031	0.0017	0.0006	
15.0		0.0470	0.0138	0.0014	0.0012	0.0003	
20	(1.0, 0.5)	1.5	0.2971	-0.0071	-0.0055	0.0110	0.0097
		2.0	0.4525	0.0138	0.0032	0.0168	0.0152
		2.5	0.5556	0.0349	0.0149	0.0256	0.0220
		3.5	0.6544	0.0722	0.0250	0.0489	0.0342
		4.0	0.6731	0.0840	0.0185	0.0613	0.0374
		4.5	0.6807	0.0877	0.0038	0.0637	0.0344
	(1.0, 1.0)	1.0	0.2876	-0.0016	-0.0038	0.0084	0.0075
		2.0	0.3067	0.0160	0.0063	0.0074	0.0066
		3.0	0.2872	0.0236	0.0103	0.0078	0.0063
		4.0	0.2647	0.0276	0.0109	0.0076	0.0056
		5.0	0.2444	0.0257	0.0063	0.0068	0.0046
		6.5	0.2190	0.0263	0.0030	0.0062	0.0036
	(1.0, 1.5)	0.5	0.3232	-0.0089	-0.0084	0.0123	0.0109
		1.5	0.2506	0.0119	0.0046	0.0050	0.0045
		2.5	0.2034	0.0141	0.0060	0.0036	0.0030
		4.5	0.1516	0.0158	0.0066	0.0026	0.0019
		6.0	0.1291	0.0144	0.0049	0.0020	0.0014
		9.5	0.0978	0.0120	0.0018	0.0013	0.0008
	(1.0, 2.0)	0.5	0.3479	0.0006	-0.0046	0.0113	0.0102
		1.5	0.2063	0.0120	0.0053	0.0035	0.0031
		3.5	0.1257	0.0112	0.0052	0.0016	0.0012
6.5		0.0842	0.0084	0.0029	0.0007	0.0005	
11.5		0.0568	0.0067	0.0015	0.0004	0.0003	
15.0		0.0470	0.0057	0.0007	0.0003	0.0002	

defined as $\{L, U\}$ such that $\int_0^L p(y)dy = \alpha/2$ and $\int_U^\infty p(y)dy = \alpha/2$. Thus, when both parameters are unknown, a $100(1 - \alpha)\%$ equitailed MLPD PI for a future lognormal observation is defined as $\{L_e, U_e\}$ such that $\tilde{F}(L_e | \mathbf{x}) = \alpha/2$ and $\tilde{F}(U_e | \mathbf{x}) = 1 - \alpha/2$. The resulted MLPD PI for Y is easily seen to be of the form:

$$\exp \left\{ \hat{\mu}_n \pm t_n(\alpha/2) \hat{\sigma}_n \sqrt{1 + n^{-1}} \right\}. \quad (14)$$

Similarly, the MLPD PI for the case of only μ unknown has the form:

$$\exp \left\{ \hat{\mu}_n \pm Z_{\alpha/2} \sigma \sqrt{1 + n^{-1}} \right\}, \quad (15)$$

and the MLPD PI for the case of only σ unknown is

$$\exp \left\{ \mu \pm t_n(\alpha/2) \hat{\sigma}_n(\mu) \right\}. \quad (16)$$

The MLPD PI for the σ known case is the same as the frequentist PI, and the MLPD PIs for the other two cases are slightly different from the frequentist PIs. Closer examination shows that in these two cases the MLPD PIs are shorter with slightly lower coverage than the nominal levels. Again, we would like to concentrate on the both parameters unknown case for comparisons. The coverage probability for the frequentist PI is $1 - \alpha$ and the coverage probability for the MLPD PI can be calculated as follows:

$$\begin{aligned} CP &= P\{\hat{\mu}_n - t_n(\alpha/2) \hat{\sigma}_n \sqrt{1 + n^{-1}} \leq \log Y \leq \hat{\mu}_n + t_n(\alpha/2) \hat{\sigma}_n \sqrt{1 + n^{-1}}\} \\ &= P\{-t_n(\alpha/2) \leq \frac{\log Y - \hat{\mu}_n}{\hat{\sigma}_n \sqrt{1 + n^{-1}}} \leq t_n(\alpha/2)\} \\ &= P\{-t_n(\alpha/2) \sqrt{1 - n^{-1}} \leq t_{n-1} \leq t_n(\alpha/2) \sqrt{1 - n^{-1}}\}. \end{aligned}$$

The true coverage probability of the MLPD PI is slightly smaller than the nominal levels. It depends only on n . The larger the n is, the closer the coverage probability is to the nominal level. It can be easily calculated using a certain statistical software such as MINITAB. For $n = 10$, the coverage probabilities for the 90%, 95% and

Table 3: A Summary of the Ratio of the Interval Lengths

n	α	a	b	R				
				$\hat{\sigma}_n = 0.5$	1.0	1.5	2.0	2.5
10	0.1	6.6163	6.4769	1.0146	1.0226	1.0327	1.0436	1.0547
	.05	10.2962	9.9415	1.0216	1.0364	1.0541	1.0726	1.0916
	.01	28.4986	26.2276	1.0457	1.0868	1.1327	1.1807	1.2307
20	0.1	5.9438	5.9168	1.0032	1.0048	1.0069	1.0092	1.0115
	.05	8.6490	8.5862	1.0046	1.0075	1.0110	1.0147	1.0184
	.01	19.0865	18.7822	1.0090	1.0163	1.0244	1.0327	1.0410
50	0.1	5.6302	5.6265	1.0005	1.0007	1.0010	1.0013	1.0017
	.05	7.9361	7.9278	1.0007	1.0011	1.0016	1.0021	1.0026
	.01	15.8383	15.8031	1.0013	1.0023	1.0033	1.0045	1.0056

99% nominal levels are, respectively, 0.8804, 0.9363 and 0.9852; for $n = 20$, they are 0.8909, 0.9438 and 0.9879; and for $n = 50$, they are 0.8965, 0.9476 and 0.9892. This shows that the coverages of the MLPD PIs are very close to the nominal level even when n is very small. Now, we like to see how the two intervals compare in term of the interval lengths. The ratio of the lengths of the frequentist and the MLPD PIs is:

$$R = \frac{a^{\hat{\sigma}_n} - a^{-\hat{\sigma}_n}}{b^{\hat{\sigma}_n} - b^{-\hat{\sigma}_n}}$$

where $a = \exp[t_{n-1}(\alpha/2)\sqrt{(n+1)/(n-1)}]$ and $b = \exp[t_n(\alpha/2)\sqrt{(n+1)/(n)}]$. The ratio R depends on the sample size n , the nominal level and the value $\hat{\sigma}_n$, but does not depend on the value of $\hat{\mu}_n$. It is easy to see that R is always larger than 1, suggesting that the MLPD PI is always shorter than the frequentist PI. From the results in Table 3 we see that R increases as α or $\hat{\sigma}_n$ increases, but decreases to 1 as n goes large. The frequentist PI can be as high as 23% longer than the MLPD PI. Reducing the skewness of the population or increasing the sample size can reduce the discrepancy between the two PIs significantly.

3.4 Shortest prediction intervals

Let $p(y)$ be a predictive density function of a random variable Y . Dahiya and Guttman defines the shortest prediction interval of level $1 - \alpha$ for Y as $\{L_s, U_s\}$ such that (i) $P(L_s \leq Y \leq U_s) = 1 - \alpha$ and (ii) for $y_1 \in \{L_s, U_s\}$ and $y_2 \ni \{L_s, U_s\}, p(y_1) \geq p(y_2)$. In the case that the predictive density is unimodal, the conditions are reduced to (i) $p(L_s) = p(U_s)$ and (ii) $\int_{L_s}^{U_s} p(y)dy = 1 - \alpha$.

When σ is known but μ is unknown, the MLPD $\tilde{f}(y | \mathbf{x}; \sigma)$ is unimodal, hence the shortest MLPD PI for Y is defined as $\{L_s(\sigma), U_s(\sigma)\}$ such that

$$\begin{aligned} \Phi \left\{ \frac{\log U_s(\sigma) - \hat{\mu}_n}{\sqrt{1 + n^{-1}\sigma}} \right\} - \Phi \left\{ \frac{\log L_s(\sigma) - \hat{\mu}_n}{\sqrt{1 + n^{-1}\sigma}} \right\} &= 1 - \alpha \\ \exp \left\{ \frac{[\log U_s(\sigma) - \hat{\mu}_n]^2 - [\log L_s(\sigma) - \hat{\mu}_n]^2}{2(1 + n^{-1})\sigma^2} \right\} &= \frac{L_s(\sigma)}{U_s(\sigma)} \end{aligned} \quad (17)$$

For the other two cases, the MLPDs are not unimodal, but are either purely decreasing functions or functions with two stationary points. Hence the shortest PIs are either one-sided, or sets consisting of two sub-intervals. In these cases, the Theorem 4 of Dahiya and Guttman can be easily used to give MLPD SPIs.

Clearly, the shortest MLPD PIs should possess similar coverage properties as the corresponding equitailed MLPD PIs. Thus, it is only necessary to compare the length of the MLPD PI with that of the equitailed PI. The examples given in next section will help to see the difference.

4 An Illustrative Example

The times to failures (in hours) for 20 guidance systems were considered by Hanh and Shapiro (1967) to fit a lognormal distribution and again by Dahiya and Guttman (1982) for illustrating their shortest prediction intervals. The data: 1, 4, 5, 6, 15, 20, 40, 40, 60, 93, 95, 106, 125, 151, 200, 268, 459, 827, 840, 1089, is quite

Table 4: Results for the Guidance System Example

t	$\hat{R}(t)$	$\tilde{R}(t)$	$\hat{r}(t)$	$\tilde{r}(t)$
1	0.9817	0.9726	0.023013	0.027345
5	0.8997	0.8870	0.019685	0.019838
10	0.8241	0.8128	0.015807	0.015581
20	0.7197	0.7118	0.011785	0.011500
40	0.5920	0.5887	0.008257	0.008004
80	0.4537	0.4554	0.005498	0.005280
150	0.3325	0.3386	0.003668	0.003473
250	0.2451	0.2542	0.002585	0.002408
500	0.1493	0.1613	0.001568	0.001418
1000	0.0825	0.0953	0.000929	0.000808
2000	0.0412	0.0527	0.000540	0.000448
5000	0.0139	0.0222	0.000257	0.000198

skewed. The MLEs are $\hat{\mu}_n = 4.1511$ and $\hat{\sigma}_n = 1.9856$, respectively. The estimated values for the reliability and failure rate at 12 different time points are calculated for both MLE and MLPD methods. The results are summarized in Table 4. From the results we see that the MLE and MLPD estimators are quite conformable when the time points are in the center half of the population. Significant discrepancies are observed at the two end areas, particularly for the failure rate estimators. Figure 1 presents plots of the estimated reliability and failure rate functions based on this data set. The plots clearly reveal the discrepancies of the two methods, particularly for the failure rate functions.

Insert Figure 1 near here

The 90% frequentist, MLPD equitailed and MLPD shortest PIs are, respectively, $\{1.7188, 2346.08\}$, $\{1.9003, 2122.18\}$ and $\{0.0, 941.62\}$. Notice that the shortest MLPD PI is one-sided and is much shorter than the other two PIs. This indicates the need for constructing the shortest possible prediction interval for a very skewed distribution.

5 Discussions

Continuing on the work of Yang [11] for the inverse Gaussian distribution, the usefulness of the MLPD approach in reliability and life testing studies is demonstrated once more by the lognormal distribution. Closed and simple forms of the MLPD and cumulative MLPD exist for lognormal distribution, which makes it technically more attractive than the inverse Gaussian distribution. The good property of the MLPD estimator of the failure rate deserves more attention and more detailed investigations. Furthermore, this property signals the feasibility of obtaining better reliability and failure rate estimators for other lifetime distributions besides lognormal and inverse Gaussian.

Acknowledgements

The author wishes to thank the referees and the Editor-in-Chief, Ninoslav Stojadinović, for the comments and help.

References

1. Crow EL, Shimizu K. *Lognormal Distributions: Theory and Applications*. New York: Marcel Dekker, 1988.
2. Dahiya RC, Guttman I. Shortest confidence and prediction intervals for the log-normal. *Canadian J Statist* 1982; 10:277-91.
3. Hahn GJ, Meeker WQ. *Statistical Intervals, A Guide for Practitioners*. John Wiley & Sons, Inc, 1991.
4. Johnson NL, Kotz S, Balakrishnan N. *Continuous Univariate Distributions, 2nd ed.* John Wiley & Sons, Inc, 1994.
5. Lejeune M, Faulkenberry GD. A simple predictive density function. *J Amer Statist Assoc* 1982; 77:654-59.

6. Padgett WJ, Wei LJ. Bayes estimation of reliability for the two-parameter lognormal distribution. *Commu Statist-Theory Meth* 1977; 6:443-57.
7. Rohn WB. Reliability prediction for complex systems. *Proceedings of the fifth national symposium on reliability and quality control in electronics* 1959; 381-88.
8. Sinha SK. Bayesian inference about the prediction/credible intervals and reliability function for lognormal distribution. *J Indian Statist Assoc* 1989; 27:73-8.
9. Sweet AL. On the hazard rate of the lognormal distribution. *IEEE Trans Reliability* 1990; 39:325-28.
10. Wang HZ, Ma BH, Shi JS, Estimation of environmental factors for the lognormal distribution. *Microelectron Reliab* 1992; 32:679-85.
11. Yang, Z. Maximum likelihood predictive densities for the inverse Gaussian distribution with applications to reliability and lifetime predictions. *Microelectron Reliab* 1999; 37:1413-21.