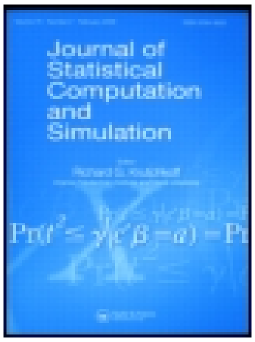

Journal of Statistical Computation and Simulation

ISSN: 0094-9655 (Print) 1563-5163 (Online) Journal homepage: http://www.tandfonline.com/loi/gscs20

Confidence intervals for the mean and a percentile based on zero-inflated lognormal data

Md Sazib Hasan & K. Krishnamoorthy

JOURNAL OF STATISTICAL COMPUTATION AND SIMULATION, 2018
https://doi.org/10.1080/00949655.2018.1439033

Confidence intervals for the mean and a percentile based on zero-inflated lognormal data

Md Sazib Hasan and K. Krishnamoorthy

Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA, USA

ABSTRACT

The problems of estimating the mean and an upper percentile of a lognormal population with nonnegative values are considered. For estimating the mean of a such population based on data that include zeros, a simple confidence interval (CI) that is obtained by modifying Tian's [Inferences on the mean of zero-inflated lognormal data: the generalized variable approach. Stat Med. 2005;24:3223—3232] generalized CI, is proposed. A fiducial upper confidence limit (UCL) and a closed-form approximate UCL for an upper percentile are developed. Our simulation studies indicate that the proposed methods are very satisfactory in terms of coverage probability and precision, and better than existing methods for maintaining balanced tail error rates. The proposed CI and the UCL are simple and easy to calculate. All the methods considered are illustrated using samples of data involving airborne chlorine concentrations and data on diagnostic test costs.

ARTICLE HISTORY
Received 27 January 2018
Accepted 6 February 2018

KEYWORDS
Coverage probability; delta-lognormal; fiducial approach; generalized variable approach; one-side tolerance limit; quantile; upper confidence limit

1. Introduction

There are practical situations where one encounters data that are skewed and contain a relatively high proportion of zeros. For example, the diagnostic test charges data in a study by Callahan et al. [1] include zeroes because some patients had no diagnostic tests during the study and nonzero diagnostic testing charges were highly skewed to the right. In workplace exposure/pollution assessment studies, measurements collected from individual workers or locations within a working facility often include zeroes, and positive measurements are highly skewed. In both situations, the problem of interest is to estimate the population mean. In diagnostic test cost analysis, the mean of diagnostic charges can be used to estimate the total charge, which reflects the entire diagnostic expenditure in a given patient population [2]. In exposure/pollution assessment, an upper confidence limit (UCL) of the mean or an UCL of an upper percentile of the exposure distribution is used to check if the exposure levels are within the federal standards; see Owen and DeRouen [3] and Krishnamoorthy et al. [4]. To make inference based on such zero-inflated data, delta-lognormal distribution is commonly postulated. The delta-lognormal distribution is a generalized form of the two-parameter lognormal distribution in which a proportion δ of

CONTACT K. Krishnamoorthy krishna@louisiana.edu Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA 70504, USA

the observations may be zeros and the nonzero values follow a lognormal distribution with parameters μ and σ, denoted by LN (μ, σ^2). The distribution function associated with a delta-lognormal population is given by

$$G(x; \delta, \mu, \sigma) = \begin{cases} \delta & \text{if } x = 0, \\ \delta + (1-\delta)F(x; \mu, \sigma) & \text{if } x > 0, \end{cases} \tag{1}$$

where $F(x; \mu, \sigma)$ is the LN(μ, σ^2) distribution function [5].

The problems of interest here are estimation of the mean and an upper percentile. The population mean is given by

$$M = (1-\delta)\exp(\mu + \sigma^2/2). \tag{2}$$

The pth quantile, denoted by q_p, of the delta-lognormal distribution is determined by the equation $G(q_p; \mu, \sigma, \delta) = p$. It follows from Equation (1) that $q_p = 0$ if $p < \delta$. For $p > \delta$, using the relation between the normal and lognormal distributions, we see that q_p is root of the equation

$$\Phi\left(\frac{\ln(q_p) - \mu}{\sigma}\right) = \frac{p-\delta}{1-\delta},$$

where Φ is the standard normal distribution function. The above equation yields

$$q_p = \exp\left(\mu + \Phi^{-1}\left(\frac{p-\delta}{1-\delta}\right)\sigma\right), \quad p > \delta. \tag{3}$$

Owen and DeRouen [3] have proposed a large sample approach to find a confidence interval (CI) for the delta-lognormal mean, which appears to be conservative even for large samples. Zhou and Tu [2] have proposed CIs based on the likelihood approach, bootstrap method and profile-likelihood. Among these CIs, profile-likelihood CI appears to be better than other two CIs. However, computation of the profile-likelihood CI is technically involved and is very liberal for small σ^2 (variance on the log scale). Tian [6] has proposed a simple *generalized variable* approach, which is a special case of the fiducial approach, to find a CI for the mean. This CI seems to be satisfactory except that it could be too conservative for $\sigma^2 \leq 5$. Recently, Li et al. [7] have proposed a modified version of Tian's generalized CI. Zou et al. [8] have proposed an approximate CI based on the method of variance estimate recovery (MOVER). These CIs are in general over cover on the left tail and under cover on the right. This means that one-sided lower confidence limits based on these methods should be conservative and the one-sided UCLs should be liberal. As noted earlier, one-sided confidence limits are used to assess the exposure/pollution levels in workplaces.

Tian [6] has obtained the so-called generalized pivotal quantity (which is the same as fiducial quantity) for $M = (1-\delta)\exp(\mu + \sigma^2/2)$ by replacing the parameters with their fiducial quantities (FQs). In particular, Tian [6] and Li et al. [7] have proposed different CIs for M based on two different FQs for δ. These fiducial CIs are too conservative in some cases and liberal in other cases. Such performances of the proposed fiducial CIs are mainly due to the fiducial quantity that is used for δ, because the fiducial quantity that is used for $\mu + .5\sigma^2$ in (2) produced CIs for a lognormal mean that are practically exact [9]. These findings suggest that a better fiducial CI may be obtained by using a better fiducial quantity for δ. Stevens [10] has proposed an alternative fiducial quantity for δ, which produced better inferences for problems involving binomial distributions [11]. In this article, we shall investigate the properties of such fiducial CIs for the delta-lognormal mean based on a different fiducial quantity for δ. Furthermore, we develop an UCL for an upper percentile using the fiducial approach and also provide a simple closed-form UCL.

The rest of the article is organized as follows. In the following section, we provide some preliminary results by describing fiducial distributions for normal parameters and binomial parameter. In Section 3, we present fiducial and MOVER CIs for the mean. In Section 4, we address the problem of finding an UCL for a percentile. The proposed CIs along with other CIs are evaluated for their accuracy in terms of coverage probability and precision in Section 5. The methods are illustrated using airborne chlorine concentration data and diagnostic test charges data in Section 6. Some concluding remarks are given in Section 7.

2. Fiducial distributions for lognormal and binomial parameters

Tian [6] has proposed the generalized variable approach for finding CI for the mean. As noted in [12], the generalized variable approach is a special case of the fiducial approach. Furthermore, the generalized variable approach, as described in [13,14], is not clear as to developing a so-called generalized pivotal quantity for a parameter of a discrete distribution. So we shall describe fiducial distributions for various parameters involved in our present estimation problem.

2.1. Normal parameters

The one-to-one relation between the lognormal distribution and the normal distribution allows us to find fiducial distributions for lognormal parameters based on a log-transformed sample. Let $(\bar{X}, S^2)$ denote the (mean, variance) based on a log-transformed sample of size m from an $LN(\mu, \sigma^2)$ distribution. Let $(\bar{x}, s^2)$ be an observed value of $(\bar{X}, S^2)$. To describe the fiducial distributions for the μ and σ^2 using Dawid and Stone's [15] method, we first note that

$$\bar{X} \stackrel{d}{=} \mu + Z\frac{\sigma}{\sqrt{m}} \quad \text{and} \quad S^2 \stackrel{d}{=} \sigma^2 \frac{\chi^2_{m-1}}{m-1},$$

where $Z \sim N(0,1)$ independently of χ^2_{m-1} random variable and the notation '$\stackrel{d}{=}$' means 'distributed as '. Furthermore, Z and χ^2_{m-1} are independent random variables. Solving the above equations for μ and σ^2, and replacing $(\bar{X}, S)$ with $(\bar{x}, s)$, we obtain the fiducial quantity (FQ) Q_μ for μ and FQ Q_{σ^2} for σ^2 as follows:

$$Q_\mu = \bar{x} - Z\frac{\sqrt{m-1}}{\sqrt{\chi^2_{m-1}}}\frac{s}{\sqrt{m}} \quad \text{and} \quad Q_{\sigma^2} = \frac{(m-1)s^2}{\chi^2_{m-1}}. \tag{4}$$

The quantities Q_μ and Q_{σ^2} are referred to as the FQs, and the conditional distributions of the FQs given $(\bar{x}, s^2)$ are called fiducial distributions. Some authors interpret the fiducial distribution as the posterior distribution without assuming prior distributions [16].

Notice that the fiducial distributions for the parameters do not depend on any unknown parameters, and they depend only on the observed value $(\bar{x}, s)$ and the sample size m. A fiducial quantity for a real-valued function, say, $h(\mu, \sigma^2)$ can be obtained by simple substitution as $h(G_\mu, G_{\sigma^2})$. For example, if $(\bar{x}, s)$ is based on a log-transformed sample from an LN(μ, σ^2) distribution, then the fiducial distribution for the lognormal mean is the conditional distribution of

$$\exp(Q_\mu + .5Q_\sigma^2) \tag{5}$$

given $(\bar{x}, s)$. The lower and upper αth quantiles of the FQ form a $1-2\alpha$ CI for the mean of the lognormal distribution [9].

2.2. Binomial parameter

Let $K \sim$ binomial (m, π) distribution with number of trials m and 'success' probability π. Let k be an observed value of K. Let $B_{a,b}$ denote the beta random variable with shape parameters a and b. Using the following relations that

$$P(K \geq k \,|\, m, \pi) = P(B_{k,m-k+1} < \pi) \quad \text{and} \quad P(K \leq k) = P(B_{k+1,m-k} > \pi),$$

two fiducial distributions, namely

$$Q_\pi^l = B_{k,m-k+1} \quad \text{and} \quad Q_\pi^u = B_{k+1,m-k} \tag{6}$$

were proposed in the literature. The fiducial distribution Q_π^l was used to find a lower confidence limit for π, and Q_π^u was used to find an UCL for π [17]. Stevens [10] suggested that one could use a distribution that is stochastically between $B_{k,m-k+1}$ and $B_{k+1,m-k}$ as an approximate fiducial distribution for π. A natural choice is

$$Q_\pi = B_{k+.5,m-k+.5}, \tag{7}$$

which is also the posterior distribution with Jeffreys prior [18]. Applications of the above approximate fiducial distribution for various problems related to binomial and Poisson distributions can be found in [11,19].

3. Confidence intervals for the mean

Let $Y_1, \ldots, Y_n$ be a sample with n_0 zeros from a delta-lognormal distribution. Let us assume that $n_0 \sim$ binomial(n, δ) so that $n_1 = n - n_0$ sample measurements are greater than zero. Let $(\bar{X}_1, S_1^2)$ denote the (mean, variance) based on the log-transformed positive sample measurements, and let $(\bar{x}_1, s_1^2)$ be an observed value of $(\bar{X}_1, S_1^2)$.

3.1. Fiducial CL for the mean

Recall that the mean of the delta-lognormal distribution is $M = (1-\delta)\exp(\mu + \sigma^2/2)$, and so the estimation problem simplifies to estimating

$$\theta = \ln(M) = \ln(1-\delta) + \mu + \sigma^2/2. \tag{8}$$

3.2. Fiducial CLs

A FQ for θ can be obtained by replacing the parameters by their FQs. Let Q_δ be an FQ for δ. Then, a FQ for θ is expressed as follows:

$$Q_\theta = \ln(1 - Q_\delta) + \bar{x}_1 - \frac{Z}{U_1}\frac{s_1}{\sqrt{n_1}} + \frac{s_1^2}{2U_1^2}, \tag{9}$$

where $Z \sim N(0,1)$ independently of $U_1^2 \sim \chi^2_{n_1-1}/(n_1 - 1)$.

3.3. Fiducial-1 CI

Tian [6] has obtained an FQ for θ by substituting two different fiducial distributions for δ. Specifically, using Equations (4) and (6), Tian [6] has proposed

$$G_\theta^l = \ln(1 - B_{n_0,n_1+1}) + \bar{x}_1 - \frac{Z}{U_1}\frac{s_1}{\sqrt{n_1}} + \frac{s_1^2}{2U_1^2}$$

to find a lower confidence limit for θ, and

$$G_\theta^u = \ln(1 - B_{n_0+1,n_1}) + \bar{x}_1 - \frac{Z}{U_1}\frac{s_1}{\sqrt{n_1}} + \frac{s_1^2}{2U_1^2}$$

to find an UCL for θ. This CI is too conservative when σ^2 (variance on the log scale) is small. Recently, Li et al. [7] have proposed another fiducial quantity for δ, namely $.5(B_{n_0,n_1+1} + B_{n_0+1,n_1})$. This leads to the FQ for θ given in Equation (9) with

$$\ln(1 - Q_\delta) = \ln[1 - .5(B_{n_0,n_1+1} + B_{n_0+1,n_1})]. \tag{10}$$

We shall refer to the CI on the basis of the FQ (9) and (10) as Fiducial-1 CI.

3.4. Fiducial-2 CI

Li et al. [7] have also proposed another fiducial quantity for $1 - \delta$ on the basis of the Wilson score CI for a binomial parameter, and is given by

$$1 - \frac{n_0 + Z^2/2}{n + Z^2} + \frac{Z}{n + Z^2}\left\{n_0\left(1 - \frac{n_0}{n}\right) + \frac{Z^2}{4}\right\}^{1/2}. \tag{11}$$

By replacing $1 - Q_\delta$ in Equation (9) with the above FQ, we can obtain a FQ for θ. Let us refer to the CI on the basis of this FQ for θ as Fiducial-2 CI.

3.5. Fiducial-3 CI

Our new CI for θ is obtained by using $B_{n_0+.5,n_1+.5}$ as a fiducial distribution for δ, instead $.5(B_{n_0,n_1+1} + B_{n_0+1,n_1})$. Since

$$1 - B_{n_0+.5,n_1+.5} \stackrel{d}{=} B_{n_1+.5,n_0+.5} \tag{12}$$

for $1-\delta$, we propose our new FQ for θ as follows:

$$\ln(B_{n_1+.5,n_0+.5}) + \bar{x}_1 - \frac{Z}{U_1}\frac{s_1}{\sqrt{n_1}} + \frac{s_1^2}{2U_1^2}. \tag{13}$$

We shall refer to the CI based on the above FQ as Fiducial-3 CI.

All the fiducial CIs can be estimated using Monte Carlo simulation in a straightforward manner. The following R code can be used to estimate a 90% CL for θ based on Equation (13). Other fiducial CIs can be found similarly.

```
Usq = rchisq(N, n1-1)/(n1-1) # N = number of
      simulation runs
U = sqrt(Usq)
Qtheta = log(rbeta(N,n1+.5,n0+.5))+ xbar-rnorm(N)/U*s1/
         sqrt(n1)+.5*s1^2/Usq
ci = exp(quantile(Qtheta,c(.05,.95)))
```

3.6. MOVER CLs for the mean

As the MOVER is relatively new, we shall first describe this approach as given in [20,21]. MOVER is a method to find a CI for a linear combination of parameters based on individual CIs of the parameters. Graybill and Wang [22] first proposed such method to find a CL for a linear combination variance components, and referred to the approach as the modified large sample approach. Zou and co-authors gave a different argument so as to justify the validity of the MOVER CL for any set of parameters including location parameters.

To describe the MOVER CI in a general set-up, consider a set of parameters $\xi_1, \ldots, \xi_g$. Let $\hat{\xi}_i$ be an estimator of ξ_i, $i = 1, \ldots, g$. Assume that $\hat{\xi}_1, \ldots, \hat{\xi}_g$ are independent. Furthermore, let (l_i, u_i) denote the $100(1-\alpha)\%$ CL for ξ_i, $i = 1, \ldots, k$. The $100(1-\alpha)\%$ MOVER CL (L, U) for $\sum_{i=1}^{k} c_i\xi_i$ can be expressed as follows:

$$L = \sum_{i=1}^{g} c_i\hat{\xi}_i - \sqrt{\sum_{i=1}^{g} c_i^2(\hat{\xi}_i - l_i^*)^2}, \quad \text{with } l_i^* = \begin{cases} l_i & \text{if } c_i > 0, \\ u_i & \text{if } c_i < 0, \end{cases} \tag{14}$$

and

$$U = \sum_{i=1}^{g} c_i\hat{\xi}_i + \sqrt{\sum_{i=1}^{g} c_i^2(\hat{\xi}_i - u_i^*)^2}, \quad \text{with } u_i^* = \begin{cases} u_i & \text{if } c_i > 0, \\ l_i & \text{if } c_i < 0. \end{cases} \tag{15}$$

Recall that the parameter to be estimated in our present set-up is $\theta = \ln(1-\delta) + \mu + .5\sigma^2$. Let $\gamma_1 = \ln(1-\delta)$ and $\gamma_2 = \mu + .5\sigma^2$ so that

$$\theta = \gamma_1 + \gamma_2.$$

Let $\hat{\delta} = n_0/n$, $\hat{\gamma}_1 = \ln(1-\hat{\delta})$ and $\hat{\gamma}_2 = \bar{X}_1 + .5S_1^2$. Since $X_1, \ldots, X_n$ are independent log-transformed delta lognormal random variables, $\hat{\delta}$, $\bar{X}_1$ and S_1^2 are independent, and so $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are independent. To find a CI for γ_1, Zou et al. [8] have used the following CI to estimate $\delta' = 1 - \delta$:

$$(\hat{\delta}'_l, \hat{\delta}'_u) = \frac{\hat{\delta}' + .5z_{1-\alpha/2}^2/n \pm z_{1-\alpha/2}\sqrt{\left[\hat{\delta}'(1-\hat{\delta}') + .25z_{1-\alpha/2}^2/n\right]/n}}{1 + z_{1-\alpha/2}^2/n}, \tag{16}$$

where $\hat{\delta}' = 1 - \hat{\delta} = n_1/n$. The CI for $\gamma_1 = \ln(1-\delta)$ is given by

$$(\hat{\gamma}_{1l}, \hat{\gamma}_{1u}) = \left(\ln(\hat{\delta}'_l), \ln(\hat{\delta}'_u)\right) \quad \text{with} \quad \hat{\gamma}_1 = \ln(\hat{\delta}'). \tag{17}$$

An approximate $1-\alpha$ CI $(\hat{\gamma}_{2l}, \hat{\gamma}_{2u})$ for γ_2 proposed in [8] is given by

$$\hat{\gamma}_{2l} = \hat{\gamma}_2 - S_1\sqrt{\frac{1}{n_1}z_{1-\alpha/2}^2 + \frac{S_1^2}{4}\left(1 - \frac{n_1-1}{\chi^2_{n_1-1;1-\alpha/2}}\right)^2} \tag{18}$$

and

$$\hat{\gamma}_{2u} = \hat{\gamma}_2 + S_1\sqrt{\frac{1}{n_1}z_{1-\alpha/2}^2 + \frac{S_1^2}{4}\left(1 - \frac{n_1-1}{\chi^2_{n_1-1;\alpha/2}}\right)^2}, \tag{19}$$

where $\hat{\gamma}_2 = \bar{X}_1 + \frac{1}{2}S_1^2$. Using the above CIs for γ_1 and γ_2 and $c_1 = c_2 = 1$ in Equations (14) and (15), we can find a MOVER CI for $\theta = \ln(1-\delta) + \mu + .5\sigma^2$ as follows:

$$\hat{\theta}_{ML} = \hat{\gamma}_1 + \hat{\gamma}_2 - \sqrt{(\hat{\gamma}_1 - \hat{\gamma}_{1l})^2 + (\hat{\gamma}_2 - \hat{\gamma}_{2l})^2} \tag{20}$$

and

$$\hat{\theta}_{MU} = \hat{\gamma}_1 + \hat{\gamma}_2 + \sqrt{(\hat{\gamma}_1 - \hat{\gamma}_{1u})^2 + (\hat{\gamma}_2 - \hat{\gamma}_{2u})^2}. \tag{21}$$

We shall refer the CI $(\hat{\theta}_{ML}, \hat{\theta}_{Mu})$ for θ as MOVER CI.

4. Upper confidence limits for an upper percentile

It follows from Equation (3) that estimation of the pth quantile simplifies to estimation of

$$\lambda_p = \mu + \Phi^{-1}\left(\frac{p-\delta}{1-\delta}\right)\sigma.$$

An FQ for $\eta = (p-\delta)/(1-\delta)$ can be found by replacing δ by its FQ. Let Q_η be an FQ for η. A FQ for λ_p can be expressed as follows:

$$\begin{aligned} Q_{\lambda_p} &= Q_\mu + \Phi^{-1}(Q_\eta)Q_\sigma \\ &= \bar{x} + \frac{Z}{U_1}\frac{s_1}{\sqrt{n_1}} + \Phi^{-1}(Q_\eta)\frac{s_1}{U_1} \\ &= \bar{x} + \frac{s_1}{\sqrt{n_1}}\left(\frac{Z + \Phi^{-1}(Q_\eta)\sqrt{n_1}}{U_1}\right), \end{aligned} \tag{22}$$

where Z and U_1 are as defined in Equation (9). Notice that the fiducial quantity Q_η must be in the interval (0, 1), or equivalently,

$$0 < \frac{p - Q_\delta^*}{1 - Q_\delta^*} < 1,$$

where Q_δ^* is the fiducial quantity for δ. Given n_0, it follows from Equation (12) that Q_δ^* has the probability distribution of the $B_{n_0+.5,n_1+.5}$ random variable whose values are bounded above by p. For a given n_0 and $0 < x < p$, the distribution function of Q_δ^* is given by

$$\begin{aligned} P(Q_\delta^* \leq x) &= P(B_{n_0+.5,n_1+.5} \leq x \mid B_{n_0+.5,n_1+.5} < p) \\ &= \frac{1}{\text{Beta}(n_0+.5, n_1+.5)} \int_0^x \frac{u^{n_0+.5-1}(1-u)^{n_1+.5-1}}{H(p; n_0+.5, n_1+.5)} du \\ &= C(x; n_0, n_1), \quad \text{say}, \end{aligned}$$

where $H(x; a, b)$ is the beta distribution function with shape parameters a and b. Thus, Q_δ^* can be generated by generating V from uniform(0, 1) and then calculating

$$\begin{aligned} Q_\delta^* &= C^{-1}(V; n_0, n_1) \\ &= H^{-1}(VH(p; n_0+.5, n_1+.5); n_0+.5, n_1+.5). \end{aligned} \tag{23}$$

To estimate the percentiles of Q_{λ_p}, it is enough to estimate the percentiles of the term within parentheses in Equation (22). For a given $(n_0, n_1, \bar{x}_1, s_1, p, \alpha)$, the following R code can be used to find the $1 - \alpha$ UCL for $\exp(\lambda_p)$, which is the $100p$ percentile of the delta-lognormal distribution:

```
V = runif(N) # N = number of simulation runs
Qd = qbeta(V*pbeta(p,n0+.5,n1+.5),n0+.5,n1+.5)
eta = (p-Qd)/(1-Qd)
U1 = sqrt(rchisq(N, n1-1)/(n1-1))
perc = quantile((rnorm(N)+qnorm(eta)*sqrt(n1))/U1, 1-alpha)
UCL = exp(xbar1+s1*perc/sqrt(n1))
```

4.1. *An approximate UCL*

To avoid simulation, an approximate UCL can be found by replacing δ in $\eta = (p - \delta)/(1 - \delta)$ by an estimate based on n_0 and n. Since we used the interval $(B_{n_0+.5,n_1+.5;\alpha}, B_{n_0+.5,n_1+.5;1-\alpha})$ to estimate δ, we shall use $\hat{\delta}_{.5} = B_{n_0+.5,n_1+.5;.5}$, the median of the beta distribution with shape parameters $n_0 + .5$ and $n_1 + .5$, as a point estimate of δ. Using $\hat{\eta} = (p - \hat{\delta}_{.5})/(1 - \hat{\delta}_{.5})$ in (22) an approximate fiducial distribution for Q_{λ_p} can be obtained as follows:

$$\begin{aligned} Q_{\lambda_p} &\sim \bar{x}_1 + \left(\frac{Z + \Phi^{-1}(\hat{\eta})\sqrt{n_1}}{U_1} \right) \frac{s_1}{\sqrt{n_1}} \quad \text{approximately}, \\ &= \bar{x}_1 + t_{n_1-1}(z_{\hat{\eta}}\sqrt{n_1}) \frac{s_1}{\sqrt{n_1}}, \end{aligned} \tag{24}$$

where $z_{\hat{\eta}}$ is the $100\hat{\eta}$ percentile of the standard normal distribution, and $t_m(d)$ denotes the noncentral t distribution with df $= m$ and the noncentrality parameter d. This approximate fiducial distribution leads to the $1-\alpha$ UCL for $\lambda_p = \mu + z_\eta \sigma$ as follows:

$$\bar{x}_1 + t_{n_1-1;1-\alpha}(z_{\hat{\eta}}\sqrt{n_1})\frac{s_1}{\sqrt{n_1}}, \tag{25}$$

where $t_{m;q}(d)$ denotes the qth quantile of the noncentral t distribution with df $= m$ and the noncentrality parameter d.

5. Coverage and precision studies

To judge the accuracy of the proposed CIs and to compare with other CIs, we carried out simulation studies. To evaluate the various CIs for the mean, we adapted the sample size and parameter configurations as given in [2,6]. In particular, we chose $\mu = -.5\sigma^2$ and $\sigma^2 = .1, .5, 1, 2, 3$. The coverage probabilities and precisions of fiducial CIs were estimated as follows. For a given (n, σ^2, δ), we first generated 10,000 sets of $(\bar{x}_1, s_1, n_0)$. For each of the generated sets, we used the R code in Section 3 with $N = 10{,}000$ runs to compute a CI. The percentage of these 10,000 CIs that include the assumed mean is an estimate of the coverage probability. The average length and error rates were estimated similarly.

For comparison studies, we will include all three CIs based on the fiducial approach and the recent MOVER CI by Zou et al. [8]. The large sample CI by Owen and DeRouen [3] and the profile-likelihood CI by Zhou and Tu [2] are not included for comparison studies because theses CIs are known to be either liberal, conservative or technically involved to compute them. The estimated coverage probabilities and average lengths of 95% Fiducial-1 and Fiducial-2 CIs by Li et al. [7], our Fiducial-3 CI, and MOVER CI by Zou et al. [8] are reported in Table 1 for sample sizes ranging from 10 to 40 and $\delta = .1, .2, .3$ and.4. Furthermore, we reported results only for $\sigma^2 \le 3$, because σ^2, being the variance of log-transformed random variable, is usually small, and our simulation results (not reported here) indicated that all CIs preform very similar for $\sigma^2 \ge 3$. Examination of values in Table 1 clearly indicate that our fiducial CI (Fiducial-3) controls coverage probabilities very close to the nominal 95% level with well balanced error rates on both tails. For small values of σ^2 and $\delta \ge .2$, Fiducial-1 and Fiducial-2 CIs by Li et al. [7] are liberal. The coverage probabilities of these two CIs could go as low as.875 when the nominal level is.95; see the values for the cases $\delta = .4$ and $n = 40$ in Table 1. For $\sigma^2 = 3$, the performances of all three fiducial CIs are very similar in terms coverage probabilities and precision. The coverage probabilities of the MOVER CI are very close to the nominal level.95 for all the cases considered; however, the MOVER CI over cover on the left tail and under cover on the right for smaller values of δ. This means that one-sided lower confidence limits based on the MOVER could be conservative and one-sided UCLs could be liberal. So MOVER one-sided confidence limits may be inappropriate to set one-sided bounds on the mean. Regarding the precisions of the CIs, we see that all CIs have approximately similar expected widths provided their coverage probabilities are close to the nominal level.95. Otherwise, the one with smaller coverage probabilities appears to have smaller expected widths than those with coverage probabilities close to the nominal level.

As the Fiducial-3 CI performs better than other CIs for $\delta \le .4$, it is desired to check its performance when the proportion of zeros is somewhat larger. As suggested by a reviewer,

Table 1. Coverage probabilities, error rates and average lengths of 95% CIs for the mean.

$\delta = .1$		Fiducial-1				Fiducial-2				Fiducial-3				MOVER			
n	σ^2	ER_L	CP	ER_R	AL	ER_L	CP	ER_R	AL	ER_L	CP	ER_R	AL	ER_L	CP	ER_R	AL
10	.1	.018	.934	.048	0.614	.020	.937	.043	0.616	.010	.957	.033	0.725	.001	.951	.047	0.728
	.5	.027	.943	.030	1.420	.024	.946	.030	1.414	.020	.953	.027	1.485	.010	.954	.035	1.465
	1	.027	.944	.029	2.294	.024	.950	.026	2.315	.023	.949	.028	2.339	.013	.955	.032	2.339
	2	.028	.949	.023	4.025	.027	.945	.028	4.002	.026	.952	.022	4.055	.016	.955	.028	4.045
	3	.024	.950	.025	5.659	.029	.946	.026	5.708	.024	.951	.025	5.682	.017	.955	.028	5.737
15	.1	.025	.923	.051	0.469	.025	.924	.051	0.469	.018	.950	.032	0.550	.003	.952	.044	0.560
	.5	.025	.946	.029	1.037	.026	.944	.030	1.030	.020	.953	.026	1.082	.012	.951	.037	1.071
	1	.030	.943	.027	1.631	.028	.943	.029	1.625	.027	.948	.025	1.664	.015	.954	.031	1.654
	2	.027	.948	.025	2.753	.028	.949	.023	2.770	.025	.950	.025	2.774	.017	.954	.029	2.782
	3	.028	.944	.028	3.845	.028	.949	.023	3.885	.028	.945	.027	3.860	.019	.954	.028	3.894
30	.1	.049	.894	.057	0.378	.036	.914	.050	0.313	.023	.948	.029	0.463	.010	.950	.040	0.371
	.5	.032	.934	.034	0.743	.027	.940	.033	0.670	.024	.949	.027	0.793	.015	.951	.034	0.693
	1	.028	.945	.027	1.121	.031	.944	.025	1.035	.023	.953	.024	1.157	.018	.951	.031	1.046
	2	.026	.948	.026	1.835	.025	.949	.026	1.698	.024	.950	.026	1.858	.019	.953	.028	1.712
	3	.028	.948	.024	2.555	.026	.950	.024	2.361	.026	.950	.023	2.572	.021	.951	.028	2.362
40	.1	.034	.916	.050	0.269	.038	.914	.048	0.266	.019	.950	.031	0.311	.012	.949	.039	0.315
	.5	.027	.941	.032	0.566	.029	.940	.031	0.562	.022	.949	.028	0.589	.017	.950	.033	0.589
	1	.029	.943	.028	0.871	.023	.949	.028	0.870	.026	.947	.027	0.886	.019	.951	.030	0.885
	2	.029	.946	.025	1.425	.029	.945	.026	1.425	.028	.947	.025	1.435	.020	.952	.027	1.440
	3	.024	.953	.023	1.971	.024	.950	.026	1.973	.023	.954	.022	1.979	.021	.952	.027	1.976
$\delta = .2$																	
10	.1	.045	.904	.051	0.741	.041	.910	.050	0.739	.023	.948	.029	0.907	.007	.956	.037	0.880
	.5	.033	.938	.030	1.654	.031	.936	.033	1.642	.024	.950	.026	1.760	.012	.954	.034	1.754
	1	.028	.945	.026	2.660	.028	.946	.027	2.687	.022	.954	.024	2.738	.015	.955	.030	2.811
	2	.029	.944	.026	4.718	.031	.945	.024	4.788	.027	.947	.026	4.772	.017	.955	.028	4.902
	3	.027	.949	.024	6.649	.028	.945	.027	6.669	.025	.951	.023	6.690	.019	.955	.026	6.994
15	.1	.047	.904	.048	0.566	.045	.903	.052	0.565	.027	.947	.026	0.694	.011	.950	.040	0.678
	.5	.029	.940	.031	1.173	.032	.935	.033	1.164	.020	.954	.026	1.253	.013	.953	.034	1.215
	1	.028	.946	.026	1.837	.027	.943	.030	1.830	.024	.951	.025	1.893	.015	.953	.031	1.864
	2	.028	.946	.026	3.097	.025	.949	.026	3.111	.025	.950	.025	3.134	.018	.953	.028	3.124
	3	.027	.948	.025	4.340	.026	.947	.027	4.315	.025	.951	.024	4.368	.019	.954	.027	4.368
30	.1	.049	.894	.057	0.378	.047	.898	.055	0.378	.023	.948	.029	0.463	.016	.948	.036	0.460
	.5	.032	.934	.034	0.743	.035	.931	.034	0.734	.024	.949	.027	0.793	.017	.950	.033	0.780
	1	.028	.945	.027	1.121	.030	.942	.028	1.122	.023	.953	.024	1.157	.018	.951	.031	1.155
	2	.026	.948	.026	1.835	.030	.943	.027	1.853	.024	.950	.026	1.858	.019	.952	.029	1.866
	3	.028	.948	.024	2.555	.025	.950	.026	2.547	.026	.950	.023	2.572	.021	.952	.028	2.571
40	.1	.046	.901	.053	0.325	.049	.899	.052	0.324	.025	.949	.026	0.398	.017	.949	.034	0.394
	.5	.030	.937	.033	0.626	.032	.937	.031	0.620	.022	.951	.027	0.668	.018	.950	.033	0.664
	1	.029	.944	.027	0.947	.032	.939	.029	0.942	.024	.951	.025	0.976	.019	.950	.030	0.970
	2	.028	.947	.025	1.541	.029	.948	.024	1.547	.027	.950	.023	1.561	.021	.950	.029	1.561
	3	.023	.954	.023	2.140	.025	.946	.029	2.109	.022	.956	.022	2.154	.021	.950	.028	2.142

$\delta = .3$		Fiducial-1				Fiducial-2				Fiducial-3				MOVER			
n	σ^2	ER_L	CP	ER_R	AL	ER_L	CP	ER_R	AL	ER_L	CP	ER_R	AL	ER_L	CP	ER_R	AL
10	.1	.057	.894	.049	0.891	.057	.894	.048	0.896	.026	.950	.025	1.114	.014	.960	.026	1.145
	.5	.040	.928	.031	2.013	.040	.929	.031	2.029	.026	.949	.025	2.162	.015	.957	.028	2.580
	1	.040	.933	.027	3.319	.040	.932	.029	3.343	.029	.945	.025	3.430	.015	.957	.028	4.344
	2	.031	.942	.027	5.835	.033	.942	.025	5.920	.027	.947	.026	5.914	.017	.956	.027	7.566
	3	.030	.944	.026	8.530	.033	.942	.026	8.452	.027	.947	.026	8.595	.018	.956	.027	10.80
15	.1	.059	.887	.054	0.673	.061	.884	.055	0.669	.028	.949	.023	0.847	.016	.951	.033	0.812
	.5	.035	.931	.034	1.330	.038	.926	.036	1.331	.022	.952	.025	1.443	.016	.953	.031	1.403
	1	.035	.936	.028	2.096	.034	.936	.030	2.076	.027	.948	.026	2.181	.017	.954	.029	2.163
	2	.032	.943	.025	3.589	.027	.945	.028	3.566	.029	.947	.024	3.645	.018	.954	.028	3.612
	3	.028	.947	.025	5.058	.029	.942	.029	5.011	.026	.949	.025	5.103	.019	.954	.027	5.060
30	.1	.058	.885	.057	0.446	.056	.889	.056	0.448	.025	.952	.023	0.565	.019	.949	.031	0.556
	.5	.042	.924	.035	0.825	.043	.920	.037	0.823	.028	.946	.026	0.901	.019	.950	.031	0.885

(*continued*).

Table 1. Continued.

$\delta = .3$		Fiducial-1				Fiducial-2				Fiducial-3				MOVER			
n	σ^2	ER_L	CP	ER_R	AL	ER_L		ER_R	AL	ER_L	CP	ER_R	AL	ER_L	CP	ER_R	AL
	1	.035	.935	.031	1.245	.033	.939	.028	1.244	.028	.946	.026	1.299	.019	.951	.030	1.282
	2	.025	.948	.028	2.035	.030	.944	.025	2.055	.022	.952	.026	2.071	.021	.952	.028	2.062
	3	.030	.943	.027	2.815	.028	.944	.028	2.812	.028	.946	.026	2.842	.021	.951	.028	2.840
40	.1	.060	.886	.054	0.386	.056	.887	.057	0.384	.024	.951	.024	0.486	.020	.950	.031	0.479
	.5	.035	.932	.034	0.699	.040	.926	.033	0.693	.023	.952	.025	0.762	.019	.949	.032	0.750
	1	.034	.937	.029	1.040	.036	.936	.028	1.036	.027	.948	.025	1.086	.019	.952	.029	1.077
	2	.026	.948	.026	1.689	.028	.948	.025	1.687	.024	.952	.024	1.718	.020	.951	.029	1.712
	3	.028	.945	.027	2.312	.028	.943	.029	2.317	.026	.947	.026	2.335	.021	.952	.027	2.345
$\delta = .4$																	
10	.1	.066	.891	.042	1.126	.062	.893	.045	1.139	.028	.951	.021	1.408	.019	.965	.016	1.936
	.5	.043	.927	.030	2.725	.049	.923	.029	2.751	.025	.951	.024	2.915	.018	.960	.023	5.836
	1	.039	.933	.028	4.589	.040	.932	.028	4.722	.028	.947	.025	4.737	.018	.956	.025	11.23
	2	.032	.942	.025	8.513	.033	.940	.027	8.572	.026	.950	.024	8.620	.018	.957	.025	21.34
	3	.030	.945	.025	12.47	.027	.947	.026	12.30	.026	.949	.025	12.55	.018	.957	.025	30.29
15	.1	.059	.890	.051	0.801	.067	.881	.052	0.802	.023	.953	.024	1.027	.021	.954	.025	0.983
	.5	.041	.924	.034	1.569	.046	.924	.031	1.584	.025	.949	.026	1.720	.019	.954	.027	1.754
	1	.040	.932	.028	2.498	.037	.932	.031	2.517	.028	.947	.024	2.612	.019	.953	.028	2.693
	2	.034	.938	.028	4.309	.034	.940	.026	4.336	.028	.945	.027	4.389	.018	.955	.027	4.694
	3	.032	.941	.026	6.137	.030	.942	.028	6.109	.028	.946	.025	6.201	.019	.955	.026	6.637
30	.1	.066	.874	.060	0.531	.068	.876	.057	0.530	.025	.948	.027	0.685	.021	.950	.029	0.666
	.5	.048	.920	.032	0.929	.046	.919	.035	0.939	.028	.949	.023	1.032	.020	.951	.029	1.010
	1	.032	.938	.030	1.411	.036	.932	.032	1.409	.023	.952	.025	1.485	.019	.952	.029	1.452
	2	.033	.942	.025	2.282	.031	.941	.028	2.308	.029	.948	.024	2.333	.020	.952	.027	2.327
	3	.029	.946	.025	3.159	.029	.945	.026	3.162	.026	.950	.024	3.198	.021	.951	.028	3.195
40	.1	.064	.875	.062	0.455	.067	.874	.058	0.455	.026	.947	.027	0.586	.021	.951	.028	0.574
	.5	.043	.919	.038	0.780	.043	.921	.036	0.788	.025	.949	.026	0.867	.020	.951	.029	0.855
	1	.035	.934	.031	1.157	.037	.933	.029	1.156	.025	.950	.025	1.220	.021	.950	.029	1.209
	2	.028	.943	.029	1.875	.029	.944	.028	1.885	.025	.949	.027	1.918	.021	.951	.028	1.913
	3	.027	.947	.026	2.605	.032	.943	.025	2.582	.025	.950	.025	2.637	.022	.950	.027	2.611

Note: ER_L = error rates on the left tail; CP = coverage probability; ER_R = error rates on the right tail; AL = average length
Fiducial-1 & Fiducial-2: CIs by Li et al. [7]; Fiducial-3: our new fiducial CI; MOVER: CI by Zou et al. [8]

Table 2. Coverage probabilities, error rates and average lengths of 95% Fiducial-3 CIs for the mean.

$\delta = .5$	$n = 10$				$n = 15$				$n = 30$				$n = 40$			
σ^2	ER_L	CP	ER_R	AL	ER_L	CP	ER_R	AL	ER_L	CP	ER_R	AL	ER_L	CP	ER_R	AL
.1	.030	.950	.020	1.934	.023	.956	.021	1.260	.028	.947	.025	0.819	.025	.951	.023	0.705
.5	.026	.948	.024	4.815	.027	.948	.025	2.205	.028	.946	.026	1.218	.027	.949	.024	1.013
1	.025	.948	.027	8.902	.026	.949	.026	3.451	.026	.950	.024	1.726	.028	.946	.027	1.407
2	.025	.949	.026	15.85	.023	.949	.028	5.821	.025	.946	.028	2.728	.028	.948	.024	2.214
3	.025	.949	.025	23.32	.027	.947	.026	8.360	.026	.950	.024	3.718	.026	.949	.025	3.016
$\delta = .6$																
.1	.027	.966	.007	3.48	.029	.952	.019	1.64	.028	.948	.024	1.00	.024	.950	.026	0.85
.5	.030	.951	.019	12.7	.026	.949	.025	3.37	.025	.951	.024	1.47	.027	.949	.024	1.21
1	.024	.949	.025	28.3	.022	.952	.026	5.49	.026	.949	.025	2.15	.024	.950	.025	1.70
2	.031	.949	.020	53.0	.027	.946	.027	10.0	.027	.948	.025	3.36	.023	.950	.027	2.62
3	.029	.950	.021	75.6	.027	.948	.025	15.7	.024	.947	.029	4.58	.024	.947	.029	3.60

Note: ER_L = error rates on the left tail; CP = coverage probability; ER_R = error rates on the right tail; AL = average length

we evaluated the coverage probabilities, error rates and precision of the Fiducial-3 CIs for $\delta = .5$ and.6 and reported them in Table 2. These estimated values in Table 2 indicate that the Fiducial-3 CI is very satisfactory in controlling error rates and coverage probabilities. In general, the new Fiducial-3 CI is expected to work satisfactorily as long as the number

Table 3. Coverage probabilities and [averages] of 95% UCLs for $\mu + \Phi^{-1}\left(\frac{p-\delta}{1-\delta}\right)\sigma$.

		$n = 10$		$n = 15$		$n = 30$		$n = 50$	
	σ^2	Approx	Fiducial	Approx	Fiducial	Approx	Fiducial	Approx	Fiducial
		$p = .90$							
$\delta = .1$	.1	.948 [0.719]	.949 [0.712]	.947 [0.613]	.948 [0.618]	.947 [0.542]	.947 [0.550]	.947 [0.502]	.950 [0.514]
	.5	.948 [1.608]	.949 [1.612]	.949 [1.370]	.949 [1.374]	.947 [1.213]	.949 [1.234]	.947 [1.122]	.948 [1.132]
	1	.947 [2.275]	.947 [2.283]	.947 [1.937]	.948 [1.942]	.947 [1.715]	.949 [1.726]	.947 [1.586]	.948 [1.467]
	2	.948 [3.217]	.948 [3.215]	.946 [2.738]	.948 [2.744]	.946 [2.424]	.947 [2.434]	.948 [2.243]	.949 [2.253]
	5	.947 [5.081]	.947 [5.104]	.946 [4.330]	.947 [4.337]	.947 [3.834]	.949 [3.842]	.947 [3.549]	.948 [3.601]
$\delta = .2$	.1	.946 [0.719]	.946 [0.716]	.946 [0.623]	.947 [0.629]	.946 [0.526]	.944 [0.513]	.945 [0.484]	.946 [0.500]
	.5	.946 [1.612]	.947 [1.617]	.946 [1.392]	.946 [1.393]	.946 [1.180]	.946 [1.179]	.946 [1.082]	.946 [1.091]
	1	.946 [2.274]	.947 [2.284]	.946 [1.971]	.945 [1.962]	.946 [1.666]	.945 [1.667]	.945 [1.531]	.946 [1.535]
	2	.947 [3.225]	.947 [3.235]	.945 [2.784]	.945 [2.754]	.946 [2.356]	.947 [2.346]	.945 [2.164]	.946 [2.169]
	5	.946 [5.091]	.948 [5.181]	.946 [4.406]	.947 [4.416]	.946 [3.727]	.945 [3.711]	.945 [3.424]	.946 [3.434]
$\delta = .4$	.1	.943 [0.743]	.944 [0.743]	.942 [0.600]	.944 [0.608]	.940 [0.486]	.942 [0.494]	.940 [0.437]	.944 [0.445]
	.5	.944 [1.660]	.946 [1.673]	.941 [1.340]	.942 [1.345]	.940 [1.086]	.943 [1.096]	.941 [0.978]	.941 [0.981]
	1	.943 [2.344]	.945 [2.354]	.942 [1.898]	.942 [1.903]	.941 [1.536]	.941 [1.541]	.940 [1.383]	.943 [1.392]
	2	.943 [3.315]	.944 [3.321]	.941 [2.682]	.943 [2.685]	.940 [2.170]	.942 [2.174]	.941 [1.958]	.941 [1.955]
	5	.944 [5.259]	.946 [5.302]	.941 [4.237]	.942 [4.247]	.940 [3.434]	.941 [3.433]	.940 [3.094]	.943 [3.106]
		$p = .95$							
$\delta = .1$	.1	.948 [0.901]	.950 [0.910]	.948 [0.795]	.951 [0.800]	.949 [0.687]	.950 [0.690]	.948 [0.637]	.949 [0.640]
	.5	.947 [2.012]	.950 [2.020]	.948 [1.775]	.950 [1.777]	.949 [1.536]	.951 [1.540]	.949 [1.427]	.950 [1.430]
	1	.949 [2.855]	.951 [2.900]	.947 [2.508]	.950 [2.519]	.948 [2.172]	.953 [2.180]	.948 [2.016]	.948 [2.014]
	2	.949 [4.029]	.951 [4.040]	.948 [3.552]	.952 [3.496]	.948 [3.070]	.951 [3.090]	.948 [2.853]	.951 [2.863]
	5	.947 [6.373]	.949 [6.473]	.948 [5.613]	.949 [5.620]	.948 [4.856]	.949 [4.901]	.949 [4.514]	.952 [4.520]
$\delta = .2$	.1	.949 [0.921]	.948 [0.922]	.948 [0.795]	.948 [0.788]	.948 [0.677]	.953 [0.687]	.948 [0.626]	.954 [0.634]
	.5	.947 [2.055]	.949 [2.029]	.946 [1.774]	.947 [1.767]	.948 [1.514]	.952 [1.535]	.948 [1.400]	.951 [1.417]
	1	.949 [2.904]	.951 [2.903]	.949 [2.515]	.947 [2.456]	.948 [2.144]	.953 [2.151]	.947 [1.980]	.953 [2.003]
	2	.949 [4.113]	.951 [4.207]	.949 [3.553]	.948 [3.505]	.947 [3.029]	.953 [3.310]	.948 [2.800]	.953 [4.476]
	5	.949 [6.509]	.951 [6.506]	.949 [5.629]	.951 [5.609]	.947 [4.791]	.947 [4.806]	.949 [4.426]	.953 [6.333]
$\delta = .4$	.1	.949 [1.015]	.949 [1.010]	.946 [0.805]	.951 [0.792]	.946 [0.656]	.947 [0.673]	.947 [0.594]	.951 [0.605]
	.5	.948 [2.273]	.948 [2.277]	.948 [1.805]	.953 [1.818]	.946 [1.465]	.952 [1.485]	.946 [1.330]	.951 [1.332]
	1	.949 [3.210]	.950 [3.136]	.947 [2.550]	.948 [2.554]	.945 [2.070]	.952 [2.110]	.945 [1.881]	.951 [1.910]
	2	.948 [4.561]	.950 [4.570]	.947 [3.602]	.948 [3.617]	.946 [2.928]	.947 [2.917]	.945 [2.660]	.951 [2.675]
	5	.948 [7.195]	.953 [7.205]	.947 [5.705]	.949 [5.712]	.946 [4.630]	.952 [6.701]	.946 [4.204]	.949 [4.213]
		$p = .99$							
$\delta = .1$	.1	.949 [1.249]	.950 [1.255]	.948 [1.103]	.951 [1.148]	.950 [0.960]	.950 [0.969]	.949 [0.896]	.949 [0.903]
	.5	.949 [2.798]	.949 [2.801]	.951 [2.471]	.953 [2.564]	.949 [2.147]	.948 [2.144]	.950 [2.002]	.951 [1.999]
	1	.948 [3.955]	.950 [3.966]	.950 [3.486]	.950 [3.621]	.950 [3.035]	.950 [3.041]	.948 [2.832]	.950 [2.842]
	2	.949 [5.598]	.950 [5.612]	.949 [4.936]	.951 [5.118]	.949 [4.291]	.949 [4.299]	.949 [4.005]	.951 [4.010]
	5	.950 [8.843]	.949 [8.799]	.949 [7.804]	.949 [8.106]	.950 [6.787]	.950 [6.823]	.951 [6.343]	.950 [6.353]
$\delta = .2$	.1	.950 [1.296]	.949 [1.290]	.950 [1.122]	.949 [1.122]	.948 [0.960]	.949 [0.963]	.949 [0.892]	.951 [0.890]
	.5	.949 [2.897]	.949 [2.901]	.950 [2.506]	.950 [2.506]	.949 [2.148]	.950 [2.152]	.950 [1.995]	.950 [1.989]
	1	.949 [4.103]	.951 [4.110]	.950 [3.546]	.949 [3.546]	.949 [3.038]	.947 [3.031]	.950 [2.821]	.950 [2.821]
	2	.948 [5.794]	.951 [5.784]	.949 [5.013]	.949 [5.013]	.950 [4.296]	.951 [4.290]	.949 [3.986]	.950 [3.981]
	5	.949 [9.155]	.951 [9.159]	.950 [7.929]	.951 [7.929]	.949 [6.796]	.952 [6.806]	.950 [6.304]	.951 [6.311]
$\delta = .4$	.1	.950 [1.536]	.949 [1.526]	.948 [1.191]	.950 [1.191]	.948 [0.967]	.949 [0.962]	.948 [0.883]	.950 [0.887]
	.5	.949 [3.433]	.948 [3.421]	.949 [2.658]	.951 [2.658]	.947 [2.162]	.949 [2.169]	.948 [1.975]	.951 [1.981]
	1	.950 [4.856]	.950 [4.861]	.949 [3.764]	.949 [3.764]	.949 [3.059]	.947 [3.039]	.948 [2.791]	.951 [2.796]
	2	.950 [6.882]	.950 [6.871]	.949 [5.322]	.949 [5.322]	.949 [4.326]	.950 [4.329]	.949 [3.951]	.950 [3.949]
	5	.951 [10.87]	.951 [10.91]	.949 [8.403]	.949 [8.403]	.948 [6.832]	.949 [6.829]	.948 [6.244]	.950 [6.252]

of positive observations in the sample is three or more, which is a reasonable requirement to estimate the variance σ^2 effectively. Overall, we see that our Fiducial-3 CI has coverage probabilities very close to the nominal level.95 and maintain well-balanced tail error rates.

The coverage probabilities and expectations of 95% UCLs for the 100pth percentile of a delta-lognormal distribution are estimated along the lines for the mean described above, and they are reported in Table 3. The reported values are for $\delta = .1, .2, .4$,

$\sigma^2 = .1, .5, .1, 2, 5,\ n = 10,15,30,50$ and $p = .90,.95$ and.99. Both fiducial UCL based on Equation (22) and the approximate one (25) perform like an exact UCL with respect to coverage probability for all sample sizes and $p = .95$ and.99. For estimating the 90th percentile of a delta-lognormal distribution, the coverage probabilities of both UCLs are slightly smaller than the nominal level.95, but not below.94. In general, we see that the fiducial UCL and the approximate UCL perform very similar in terms of coverage probability and precision. Thus, for practical applications, approximate UCL can be recommended as they are straightforward to compute and perform very similar to the fiducial UCL.

We have also carried out simulation studies for 90% and 99% UCLs for 100pth percentile when $p = .90,.95,.99$. The results for these cases are not reported here as they are very similar to those for 95% UCLs. In general, the coverage probabilities are slightly smaller than the nominal level when $p = .90$, and they are practically equal to the nominal levels when $p = .95$ and .99.

6. Examples

Example 6.1: This example and the data are taken from Owen and DeRouen [3]. Measurements of the concentration of airborne chlorine were collected by occupational health personnel from an industrial site in the United States. A sample of 15 chlorine determinations (in parts per million) were collected over the period of a working day, and the data are as follows:

Measurement	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Chlorine reading (ppm)	6	0	6	9	6.5	0	0	0	1	.5	2	2	0	0	1

On the basis of this sample measurements, it is desired to determine whether the company was in compliance with the federal standard threshold limit value (TLV) of 1.0 ppm for chlorine. The 9 positive measurements fit a lognormal model very well, and so a delta-lognormal model can be used to estimate the mean or an upper percentile of airborne concentrations. The sample sizes and the statistics based on the log-transformed positive measurements are as follows:

$$n = 15, \quad n_0 = 6, \quad n_1 = 9, \quad \bar{x}_1 = 0.92730 \quad \text{and} \quad s_1 = 1.02802.$$

Using the R code in Section 3, we computed the 95% Fiducial-3 UCL for the mean $\exp(\theta)$ as 9.24. Since the UCL is not less than 1, we see that the data do not provide enough evidence to support that the company was in compliance. We have reported one-sided confidence limits for the mean chlorine concentration based on various methods in the following table. The 95% lower confidence limits (LCLs) based on all methods are greater than 1, indicating that the mean concentration is exceeding the federal standard at the level.05. Since the Fiducial-1 CI could be slightly liberal, it is the shortest among all three fiducial CIs.

A 95% UCL for the 95th percentile of exposure distribution is also used to assess the contaminant level in a workplace. To compute the UCL (22), we used the R code in Section 4 with $N = 100{,}000$ runs, and found the 95% UCL for $\lambda_{.95} = \mu + \Phi^{-1}((p - \delta)/(1 - \delta))\sigma$ as 3.6056. This yields $\exp(3.6056) = 36.80$ as the 95% UCL for the 95th percentile of

95% one-sided confidence limits for the mean chlorine concentration

Method	95% LCL	95% UCL
Fiducial-1	1.37	9.19
Fiducial-2	1.29	9.22
Fiducial-3	1.28	9.24
MOVER	1.32	8.88

chlorine concentration distribution. This means that 95% of chlorine concentrations are no more that 37 ppm with confidence 95%. To compute the approximate UCL for the 95th percentile, note that $\hat{\delta}_{.5} = B_{6.5,9.5;.5} = .400251$, $\hat{\eta} = (p - \hat{\delta}_{.5})/(1 - \hat{\delta}_{.5}) = .91635$ and $z_{\hat{\eta}}\sqrt{n_1} = 4.142795$. The noncentral t factor is $t_{8,.95}(4.142795)/3 = 2.6103$. Furthermore,

$$\bar{x} + \frac{1}{\sqrt{n_1}} t_{n_1-1,.95}(z_{\hat{\eta}}\sqrt{n_1})s = 0.92730 + 2.6103 \times 1.02802 = 3.6107.$$

Thus, the 95% UCL for the 95th percentile is $\exp(3.6107) = 36.99$. Notice that this approximate UCL is practically the same as the one based on simulation given above.

Example 6.2: In this example, we shall use the data set of diagnostic test charges in a study by Callahan et al. [1]. The data set includes 40 patients and 10 of them had no diagnostic tests during the study period. Zhou and Tu [2] have shown that the data fit a lognormal distribution and constructed CIs for the mean cost. Tian [6] has also used the data for illustration purpose. Relevant statistics to calculate CIs are

$$n = 40, \quad n_0 = 10, \quad n_1 = 30, \quad \bar{x}_1 = 6.8535 \quad \text{and} \quad s_1^2 = 1.8696.$$

The 95% CIs for the mean based on various methods are as follows.

Fiducial-1	(1005.9, 4634.0)	Fiducial-3	987.8, 4654.2)
Fiducial-2	(981.7, 4650.4)	MOVER	(981.1, 4573.3)

All fiducial CIs were computed using simulation consisting of 1,000,000 runs. Notice that the MOVER CI lies on the left side of Fiducial-3 CI. This is in agreement with our simulation results which indicated that the MOVER CIs over cover on the left tail and under cover on the right tail. Furthermore, we observe that Fiducial-1 and Fiducial-2 CIs are shorter than the Fiducial-1 CI, because the former CIs are in general slightly liberal.

7. Concluding remarks

Fiducial distribution for the parameter of a discrete distribution is not unique, and so by considering different fiducial distributions for the proportion of zeroes in the lognormal population one could find different fiducial CIs for the mean. In this article, we proposed a CI for the mean of a delta-lognormal distribution by using the fiducial distribution suggested by Stevens [10], and modifying Tian's [6] CI. The proposed CI for the mean is practically exact in terms of coverage probability and maintains a well-balanced tail error rates. The proposed fiducial CI, like other CIs, is simple to compute and works well even for small sample sizes. We have also addressed the problem of constructing an UCL for an upper percentile by providing a simple closed-form UCL that is very satisfactory for estimating the usual upper percentiles of practical interest. As noted earlier, the fiducial approach, in the name of generalized variable approach, has been used in numerous articles to find CIs for a function of parameters such as quantile in the random effects model and mixed models, to analyse interlaboratory trials [23,24], to develop tests for establishing bioequivalence [25], and to find CIs for various correlation coefficients [26]. Thus, the fiducial approach is versatile and provides simple yet accurate solutions to many complex problems.

Acknowledgments

The authors are grateful to two reviewers and an associate editor for providing useful comments and suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] Callahan CM, Kesterson JG, Tierney WM. Association of symptoms of depression with diagnostic test charges. Ann Intern Med. 1997;126:426–432.
[2] Zhou XH, Tu W. Confidence intervals for the mean of diagnostic test charge data containing zeros. Biometrics. 2000;56:1118–1125.
[3] Owen WJ, DeRouen TA. Estimation of the mean for lognormal data containing zeroes and left-censored values, with applications to the measurement of worker exposure to air contaminants. Biometrics. 1980;36:707–719.
[4] Krishnamoorthy K, Mallick A, Mathew T. Model-based imputation approach for data analysis in the presence of non-detects. Ann Occup Hyg. 2009;59:249–268.
[5] Aitchison J. On the distribution of a positive random variable having discrete probability mass at the origin. J Am Stat Assoc. 1955;50:901–908.
[6] Tian L. Inferences on the mean of zero-inflated lognormal data: the generalized variable approach. Stat Med. 2005;24:3223–3232.
[7] Li X, Zhou X, Tian L. Interval estimation for the mean of lognormal data with excess zeros. Stat Probabil Lett. 2013;83:2447–2453.
[8] Zou GY, Huo CY, Taleban J. Confidence interval estimation for lognormal data with application to health economics. Comput Stat Data Anal. 2009;53:3755–3764.
[9] Krishnamoorthy K, Mathew T. Inferences on the means of lognormal distributions using generalized p-values and generalized confidence intervals. J Stat Plan Inference. 2003;115:103–121.
[10] Stevens WL. Fiducial limits of the parameter of a discontinuous distribution. Biometrika. 1950;37:117–129.
[11] Krishnamoorthy K, Lee M. Inference for functions of parameters in discrete distributions based on fiducial approach: binomial and Poisson cases. J Stat Plan Inference. 2010;140:1182–1192.
[12] Hannig J. On generalized fiducial inference. Stat Sin. 2009;19:491–544.
[13] Tsui K-W, Weerahandi S. Generalized p-values in siginificance testing of hypotheses in the presence of nuisance parameters. J Am Stat Assoc. 1989;84:602–607.
[14] Weerahandi S. Generalized confidence intervals. J Am Stat Assoc. 1993;88:899–905.
[15] Dawid AP, Stone M. The functional-model basis of fiducial inference. Ann Stat. 1982;10:154–1074.
[16] Efron B. R.A. Fisher in the 21st century. Stat Sci. 1998;13:95–122.
[17] Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of binomial. Biometrika. 1934;26:404–413.

[18] Brown LD, Tony Cai T, Das Gupta A. Interval estimation for a binomial proportion (with discussion). Stat Sci. 2001;16:101–133.
[19] Krishnamoorthy K, Lee M, Zhang D. Closed-form fiducial confidence intervals for some functions of independent binomial parameters with comparisons. Stat Methods Med Res. 2017;26:43–63.
[20] Zou GY, Donner A. Construction of confidence limits about effect measures: a general approach. Stat Med. 2008;27:1693–1702.
[21] Zou GY, Huo CY, Taleban J. Simple confidence intervals for lognormal means and their differences with environmental applications. Environmetrics. 2009;20:172–180.
[22] Graybill FA, Wang C-M. Confidence intervals on nonnegative linear combinations of variances. J Am Stat Assoc. 1980;75:869–873.
[23] Krishnamoorthy K, Mathew T. One-Sided tolerance limits in balanced and unbalanced one-way random models based on generalized confidence limits. Technometrics. 2004;46:44–52.
[24] Iyer HK, Jack Wang CM, Mathew T. Models and confidence intervals for true values in interlaboratory trials. J Am Stat Assoc. 2004;99:1060–1071.
[25] McNally RJ, Iyer H, Mathew T. Tests for individual and population bioequivalence based on generalized p-values. Stat Med. 2003;22:31–53.
[26] Krishnamoorthy K, Xia Y. Inferences on correlation coefficients: one-sample, independent and correlated cases. J Stat Plan Inference. 2007;137:2362–2379.