Malaysian Journal of Mathematical Sciences 9(1): 127-144 (2015)

MALAYSIAN JOURNAL OF MATHEMATICAL SCIENCES

Journal homepage: http://einspem.upm.edu.my/journal

A Coverage Probability on the Parameters of the Log-Normal Distribution in the Presence of Left-Truncated and Right-Censored Survival Data

[1*]Thirunanthini Manoharan, [2]Jayanthi Arasan, [1]Habshah Midi and [2]Mohd Bakri Adam

[1]Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

[2]Laboratory of Computational Statistics and Operational Research, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

E-mail: mthirunanthini@gmail.com

*Corresponding author

ABSTRACT

The log-normal distribution is often used to model lifetime data due to its non-monotonic hazard rate. However, with left-truncated data the normal approximation fails due to the increased skewness in this distribution. This sometimes results in the poor performance of the confidence interval estimation based on the asymptotic normality of the maximum likelihood estimates, especially when the sample sizes are small. The purpose of this research is to compare and analyze the performance of the Wald, likelihood ratio and jackknife confidence intervals based on the widths of the intervals for the parameters of the log-normal model with fixed covariates through a coverage probability study. A lifetime data is therefore simulated under six different settings; model 1 (no truncation with exact observations), model 2 (low truncation with exact observations), model 3 (high truncation with exact observations), model 4 (no truncation with low censoring), model 5 (low truncation with low censoring) and model 6 (high truncation with low censoring). The comparative study indicates that the Wald, likelihood ratio and jackknife intervals performed reasonably well when no truncation or truncation is present and exact observations are available (model 1, model 2 and model 3) compared to when no truncation or truncation is observed with the presence of censoring (model 4, model 5 and model 6). Additionally, it is also evident from the results that the jackknife method outperformed the Wald and likelihood ratio methods specifically for the covariate parameter of the log-normal model even with small sample sizes when data is left-truncated with the presence of low censoring.

Keywords: Log-normal distribution, left-truncated and right censored, Wald, likelihood ratio, Jackknife

1. INTRODUCTION

The statistical distribution that has gained a popular use among medical practitioners to model lifetime data is the log-normal distribution which is known to have a non-monotone hazard rate; the hazard rate that increases to a maximum and later decreases. Studies on survival times after cancer diagnosis i.e lung and breast cancer diagnosis, age of onset Alzheimer disease and latency periods of diseases have often been shown to follow a log-normal distribution. This has been highlighted by authors Tai *et.al* (2003), Royston (2001) and Limpert *et.al* (2001). Conversely, lifetime of an observation is left-truncated when it is not feasible to follow an individual from the beginning time point of the study, $t = 0$ but at some time point u due to cost or time constraint. Subsequently, only those who experience some transitional event i.e. diagnosed with lung cancer are recruited into the study and followed prospectively until the event of interest i.e. death occurs in addition to the usual right censoring. This type of data is also known as left-truncated and right censored (LTRC) which is usually encountered in prevalence cohort study and has an extensive use in the field of survival studies (Grover and Sabharwal, 2012). Since time to onset might be random for each individual, observations may enter the study at random time points or delayed entry occurs (Shen (2009)).

Information on lifetime of an observation is only considered upon time of entry, u or left-truncated at u. In other words, if t is the lifetime of an individual, under left-truncation $t \geq u$, and individuals with $t < u$ remains unknown or unobserved by researchers. Thus, when a log-normal model is fitted to LTRC data, some of the observations on the left-tail of the distribution will be disregarded consequently increasing the existing skewness of the log-normal distribution (Cain *et.al*, 2011). In other words, the selection mechanism applied to the study design would result in reduced data and subsequently result in the poor performance of the normal approximation method in constructing the confidence intervals for the parameters.

Many authors have shown interest in determining suitable inferential methods for the parameters of the log-normal distribution with censored data. Lawless (1982) has indicated that the Wald method would perform poorly with small samples particularly with heavy censoring. As an alternative, he proposed the likelihood ratio method which is based on asymptotic chi-square distribution as the method which often outperforms the Wald for small to moderate samples. However, the likelihood ratio method is

computationally intensive and provide interval limits that is usually close to the one obtained using the Wald method when the sample sizes are large. Further Schmee *et.al* (1985) also presented that asymptotic based intervals result in anticonservative intervals for small samples. Doganaksoy and Schmee (1993) showed that the likelihood ratio method performed better with parameters of the log-normal distribution compared to the Wald method. Additionally, Mitra (2013) proposed the parametric bootstrap technique for parameters of the log-normal model fitted with LTRC data without covariates. Arasan and Lunn (2008) concluded that the jackknife method worked well than any of the bootstrap techniques for censored samples. Although many of the research works on coverage probability study are focused on censored samples, there is very limited work in investigating the LTRC survival data.

On this basis, a coverage probability study is conducted to assess the performance of the Wald, likelihood ratio (LR) and jackknife (JK) based confidence interval estimation methods for the parameters in the log-normal model for a simulated lifetime data under six different settings; model 1 (no truncation with exact observations), model 2 (low truncation with exact observations), model 3 (high truncation with exact observations), model 4 (no truncation with moderate censoring), model 5 (low truncation with moderate censoring) and model 6 (high truncation with moderate censoring).

2. LOG NORMAL MODEL WITH LEFT-TRUNCATION AND COVARIATES

In this study, we considered a single fixed covariate. Following that, the density and survival function are given in (1) and (2) correspondingly as follows:

$$f(t_{\mathrm{i}}) = \frac{1}{t_i \sigma \sqrt{2\pi}}\, e^{-\frac{1}{2}\left[\left(\frac{\log t_i - (\beta_0+\beta_1 x_i)}{\sigma}\right)^2\right]} \tag{1}$$

and

$$S(t_{\mathrm{i}}) = 1-\Phi\left(\frac{\log t_i - (\beta_0+\beta_1 x_i)}{\sigma}\right) \tag{2}$$

for $i = 1,2,...,n$. Following that, the likelihood function consisting both exact and right censored observations with and without left-truncation and $\boldsymbol{\theta} = (\sigma, \beta_0, \beta_1)$ is given in (3) and (4) respectively as follows:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} \left\{ \frac{f(t_i)}{S(u_i)} \right\}^{c_i} \left\{ \frac{S(r_i)}{S(u_i)} \right\}^{1-c_i} \tag{3}$$

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} \left\{ f(t_i) \right\}^{c_i} \left\{ S(r_i) \right\}^{1-c_i} \tag{4}$$

with failure times (t_i), right censored times (r_i) and left truncated times (u_i). Also the censoring indicators is defined in (5) as follows:

$$c_i = \begin{cases} 0 \text{ if subject is right-censored} \\ 1 \text{ otherwise} \end{cases} \tag{5}$$

Therefore the log-likelihood function for observations with or without left-truncation attribute can be derived by combining the likelihood function as given in (3), (4) and (5) with a truncation indicator variable (v_i). This is defined in (6) and (7).

$$\begin{aligned} l(\boldsymbol{\theta}) = & \sum_{\mathrm{i}=1}^{\mathrm{n}} \left[-c_i \log(t_i \sigma) - \frac{c_i}{2\sigma^2} \left(\log t_i - \mu \right)^2 \right] \\ & + \sum_{\mathrm{i}=1}^{\mathrm{n}} \left[\left(1 - c_i \right) \log \left(1 - \Phi \left(\frac{\log t_i - \mu}{\sigma} \right) \right) \right] \\ & - \sum_{\mathrm{i}=1}^{\mathrm{n}} \left[\left(1 - v_i \right) \log \left(1 - \Phi \left(\frac{\log u_i - \mu}{\sigma} \right) \right) \right] \end{aligned} \tag{6}$$

with $\mu = \beta_0 + \beta_1 x_i$ and the truncation indicator,

$$\mathrm{v}_i = \begin{cases} 0 \text{ if subject is right-censored} \\ 1 \text{ otherwise} \end{cases} \tag{7}$$

By utilizing the result in (6) and (7), the log-likelihood for each model can be further simplified by choosing the appropriate values for c_i and v_i, e.g. when no truncation and exact observations are available, $c_i = 1$ and $\mathrm{v}_i = 1$.

3. CONFIDENCE INTERVAL ESTIMATES

The exact confidence intervals (CI) are in practice difficult to construct and unavailable under Type-I and random censoring (Doganaksoy and Schmee, 1993). Thus, as an alternative, researchers opt for Wald based C.I for parameters followed by the likelihood ratio (LR) method. Arasan (2008) suggested that parameterization based confidence intervals such as $\log(\sigma)$ for the scale parameter σ produce intervals that are more symmetrical as one should expect that Wald based interval estimates for parameter σ to be highly asymmetrical due to a sharp boundary in the parameter space. The parameterization of $\log(\sigma)$ method addressed as PLS in this article is equally explored. We should anticipate that the jackknife (JK) method to perform the best compared to all the other proposed methods. The reason being, the jackknife adaptation of the consistent root of the maximum likelihood equation has equivalent asymptotic distribution as the consistent root; in addition to that the jackknife estimate of the variance of the asymptotic distribution of the consistent root is itself consistent, refer to Reeds (1978).

The suitability of the proposed CI methods is assessed based on the least number of asymmetrical, conservative and anticonservative intervals produced. The following section discusses on methods of constructing confidence intervals (C.I) estimates for parameter σ which would equally apply to the rest of the parameters β_0 and β_1 of the log-normal model.

3.1 Wald method

Let $\hat{\sigma}$ be the maximum likelihood estimate (mle) of σ. The $100(1-\alpha)\%$ C.I for the parameter σ is given by (8) as follows:

$$\hat{\sigma} - z_{1-\frac{\alpha}{2}}\sqrt{\mathrm{var}(\hat{\sigma})} < \sigma < \hat{\sigma} + z_{1-\frac{\alpha}{2}}\sqrt{\mathrm{var}(\hat{\sigma})} \qquad (8)$$

with $\text{var}(\hat{\sigma})$ the first diagonal element of the inverse of the observed Fisher information matrix $\mathbf{I}^{-1}(\hat{\boldsymbol{\theta}})$ and $\hat{\boldsymbol{\theta}}$ be the mle of vector of parameter $\boldsymbol{\theta}$. By utilizing the same principle, the PLS C.I for $\log(\sigma)$ is given by

$$\log(\hat{\sigma}) \pm z_{1-\frac{\alpha}{2}} \sqrt{\text{var}\left[\log(\hat{\sigma})\right]} \tag{9}$$

where the variance of $\log(\sigma)$ can be estimated using the delta method and is given as $\text{var}\left[\log(\hat{\sigma})\right] \approx \dfrac{\text{var}(\hat{\sigma})}{\left[\exp\left[\log(\hat{\sigma})\right]\right]^2} \approx \dfrac{\text{var}(\hat{\sigma})}{\hat{\sigma}^2}$.

Therefore the $100(1-\alpha)\%$ C.I for the parameter σ using the result from (9) can be obtained using the back transformation method given by,

$$\hat{\sigma}\exp\left(-z_{1-\frac{\alpha}{2}} \frac{\sqrt{\text{var}(\hat{\sigma})}}{\hat{\sigma}}\right) < \sigma < \hat{\sigma}\exp\left(z_{1-\frac{\alpha}{2}} \frac{\sqrt{\text{var}(\hat{\sigma})}}{\hat{\sigma}}\right) \tag{10}$$

Note that the PLS method is only applied to the shape parameter σ.

3.2 Likelihood Ratio (LR) method

The LR statistics in inspecting the null hypothesis $H_0: \sigma = \tilde{\sigma}$ versus $H_1: \sigma \neq \tilde{\sigma}$ for parameter σ can be written as,

$$\psi(\tilde{\sigma}) = -2\left[l(\tilde{\sigma},\tilde{\boldsymbol{\eta}}) - l(\hat{\sigma},\hat{\boldsymbol{\eta}})\right] \sim \chi^2_{(1,1-\alpha)} \tag{11}$$

with l the likelihood function, $\boldsymbol{\eta} = (\beta_0, \beta_1)$ the vector of nuisance parameters, $(\tilde{\sigma},\tilde{\boldsymbol{\eta}})$ maximizes $l(\sigma,\boldsymbol{\eta})$ under the null hypothesis and $(\hat{\sigma},\hat{\boldsymbol{\eta}})$ is the mle of $(\sigma,\boldsymbol{\eta})$. Thus, the $100(1-\alpha)\%$ CI for σ can be estimated as a set of values of $\tilde{\sigma}$; a lower bound by $\tilde{\sigma}_L < \hat{\sigma}$ and upper bound $\tilde{\sigma}_U > \hat{\sigma}$ for which the null hypothesis will not be rejected or equivalently by finding set of values of $\tilde{\sigma}$ so that $l(\hat{\sigma},\hat{\boldsymbol{\eta}}) - \frac{1}{2}\chi^2_{(1,1-\alpha)} \leq l(\tilde{\sigma},\tilde{\boldsymbol{\eta}})$, refer to (11).

3.3 Jackknife (JK) method

Let $w=(w_1,w_2,...,w_n)$ be the original dataset with n observations and $\hat{\sigma}$ be the mle of σ obtained from this dataset. The i^{th} jackknife sample is constructed by excluding the i^{th} observation from the original dataset. Thus each jackknife sample would consist of $n-1$ observations. The i^{th} jackknife sample with the i^{th} observation removed can be expressed as $w_{(i)}=(w_1,w_2,...,w_{i-1},w_{i+1,...,}w_n)$ and $\hat{\sigma}_{(i)}$ is the mle of σ obtained from this sample. Thus, the $100(1-\alpha)\%$ CI for σ using the JK method is given by,

$$\hat{\sigma}_{jk} - t_{(1-\alpha/2,n-1)}\,\widehat{se}_{jk}(\hat{\sigma}) < \sigma < \hat{\sigma}_{jk} + t_{(1-\alpha/2,n-1)}\,\widehat{se}_{jk}(\hat{\sigma}) \qquad (12)$$

with $\hat{\sigma}_{(jk)}=\hat{\sigma}-(n-1)(\hat{\sigma}_{(.)}-\hat{\sigma})$ and $\hat{\sigma}_{(.)}=\sum_{i=1}^{n}\hat{\sigma}_{(i)}/n$. Also, the jackknife estimate of the bias, $\hat{\sigma}_{(i)}-\hat{\sigma}_{(.)}$ and the standard error, $\widehat{se}_{jk}(\hat{\sigma})$ is obtained from the jackknife sample with $\widehat{se}_{jk}(\hat{\sigma})=\sqrt{\frac{n-1}{n}\sum_{i=1}^{n}\left(\hat{\sigma}_{(i)}-\hat{\sigma}_{(.)}\right)^2}$.

4. SIMULATION AND COVERAGE PROBABILITY STUDY

The simulation study on LTRC survival data proposed by Mitra (2013) is adopted and modified to mimic the small cell lung cancer survival data studied by Tai *et.al* (2003) which provides a satisfactory fit with the log-normal distribution.

The estimates from the proposed model are used as the true parameter values for the simulation study namely $\boldsymbol{\theta}=(\sigma,\beta_0,\beta_1)=(0.50,2.87,0.05)$ to obtain more realistic survival times. The month of truncation or the beginning time point of the study, *y* is fixed.

A set of random number of months which basically represents the month of diagnosis of the lung cancer is simulated with unequal probabilities with replacement; before $\left(y_{b_k}\right)$ and after $\left(y_{a_j}\right)$ the month of truncation where k = 1,2,...,n_1 and j=1,2,...,n_2.

In other words, y_{b_k} represents all prevalence cohort or left-truncated at y with $y > y_{b_k}$ and is fixed at 20% and 60% to determine the effect of high and low truncation on the length of the confidence interval estimates. The remaining observations are incidence cohort, y_{a_j} observed from the beginning time point of the study with $y = 0$ and $y < y_{a_j}$. Note that in this simulation study the total observation is determined as $n = n_1 + n_2$.

The lifetimes for the prevalence cohort, t_k are simulated from the log-normal distribution as $t_k = \exp(\sigma \times \Phi^{-1}(1 - z_k) + \mu)$ for $k = 1, 2, ..., n_1$ with $z_k \sim unif(0,1)$, Φ^{-1} the inverse of the cumulative distribution function of the normal distribution, σ and μ are the shape and the location parameter respectively. Further, the lifetimes, t_k are added to y_{b_k}; if the resulting failure times are less than *y*, these months of diagnosis are removed and a new set of random values of y_{b_k}, t_k, and z_k are simulated. Following that, the left truncation times u_k are obtained as $u_k = y - y_{b_k}$ and for all the left-truncated observations $t_k > u_k$. Also, for all left-truncated observations in the study, additional parameters, β_0 and β_1 are modeled through μ as $\mu = \beta_0 + \beta_1 x_k$ with covariate $x_k \sim N(0,1)$.

The lifetimes for incidence cohort, $t_j = \exp(\sigma \times \Phi^{-1}(1 - z_j) + \mu)$ for $j = 1, 2, ..., n_2$ with $\mu = \beta_0 + \beta_1 x_j$ are simulated in the same manner as above. Note that for the incidence cohort however $u_j = 0$ as all the individuals are observed from $y = 0$.

Subsequently, as the method of simulation adopted for t_k and t_j are the same and $n = n_1 + n_2$, the lifetimes for n independent random samples can be simulated by $t_i = \exp(\sigma \times \Phi^{-1}(1 - z_i) + \mu)$ with $z_i \sim unif(0,1)$, $x_i \sim N(0,1)$ and $\mu = \beta_0 + \beta_1 x_i$ for $i = 1, 2, ..., n$. The censoring times, c_i are simulated as $c_i \sim \exp(\lambda)$, where the value of λ is adjusted to yield approximately 10% of censored data.

A coverage probability study is conducted to analyze and compare the performance of the Wald, LR and JK C.I estimates for the parameters σ, β_0 and β_1 with the nominal probability error (npe), $\alpha = 0.05$.

A coverage probability is the probability of a confidence interval containing the true parameter value, and we desire this value to be close to α, the nominal error probability. A coverage probability study is a simulation study conducted to evaluate the performance of a confidence interval estimation procedure. In any coverage probability study, we do not want a conservative (anticonservative) interval, which generates coverage probability that is greater (smaller) than $(1-\alpha)$. Further, we do not want an asymmetrical interval where when the larger error probability is less than 1.5 times the smaller one.

Following that, we generated 2000 samples of size $n = 20, 30, 80, 100, 200$ and 250 with the nominal probability error (npe), $\alpha = 0.05$ for model 1 (no truncation with exact observations), model 2 (20% truncation with exact observations), model 3 (60% truncation with exact observations), model 4 (no truncation with 10% censoring), model 5 (20% truncation with 10% censoring) and model 6 (60% truncation with 10% censoring). The estimated error probabilities on the left (lep) and right (rep) for parameter σ is calculated by adding the number of times the left (right) endpoint was more (less) than the true parameter value divided by the number of simulations; 2000 times.

Thus, for the Wald, LR , JK and PLS CI method this can be written as in (13), (14),(15) and (16) respectively as follows:

$$\begin{aligned} \text{lep} &= \#\left\{\hat{\sigma} - z_{1-\alpha/2}\sqrt{\text{var}(\hat{\sigma})} > \sigma\right\}/2000 \\ \text{rep} &= \#\left\{\hat{\sigma} + z_{1-\alpha/2}\sqrt{\text{var}(\hat{\sigma})} < \sigma\right\}/2000 \end{aligned} \tag{13}$$

$$\begin{aligned} \text{lep} &= \#\left\{\psi(\sigma) > \chi^2_{(1-\alpha)} \text{ and } \hat{\sigma} > \sigma\right\}/2000 \\ \text{rep} &= \#\left\{\psi(\sigma) > \chi^2_{(1-\alpha)} \text{ and } \hat{\sigma} < \sigma\right\}/2000 \end{aligned} \tag{14}$$

$$\begin{aligned} \text{lep} &= \#\left\{\hat{\sigma}_{jk} - t_{(1-\alpha/2,n-1)}\, \widehat{se}_{jk}(\hat{\sigma}) > \sigma\right\}/2000 \\ \text{rep} &= \#\left\{\hat{\sigma}_{jk} + t_{(1-\alpha/2,n-1)}\, \widehat{se}_{jk}(\hat{\sigma}) < \sigma\right\}/2000 \end{aligned} \tag{15}$$

$$\begin{aligned} \text{lep} &= \#\left\{\hat{\sigma}\exp(-z_{1-\alpha/2}\sqrt{\text{var}(\hat{\sigma})}/\hat{\sigma}) > \sigma\right\}/2000 \\ \text{rep} &= \#\left\{\hat{\sigma}\exp(z_{1-\alpha/2}\sqrt{\text{var}(\hat{\sigma})}/\hat{\sigma}) > \sigma\right\}/2000 \end{aligned} \tag{16}$$

Therefore, the estimated total error probability (tep) for σ is simply the sum of lep and rep. Following that outcome, a CI method is termed anticonservative (AC) if $\text{tep} > \alpha + 2.58 \times \text{se}(\hat{\alpha})$, conservative (C) if $\text{tep} < \alpha\text{-}2.58 \times \text{se}(\hat{\alpha})$ with $se(\hat{\alpha}) = \sqrt{\alpha(1-\alpha)/N}$. Also, the estimated error probabilities are asymmetric (AS) when the larger error probabilities on one side of the interval is greater than 1.5 times the smaller one.

A preferred confidence interval method produces least number of AS, CV and AC intervals, the value of the lep and rep closer to 0.025 and the value of the tep closer to npe of 0.05, (Doganaksoy and Schmee (1993)) .

In this study, it is assumed that t_i , u_i and c_i are non-informative and independent of each other. Also, the exact month of diagnosis is assumed to be known for all observations in this study. The analysis is done with *R* statistical software and the parameter estimates are obtained using the Newton-Raphson iteration procedure.

5. RESULTS AND DISCUSSIONS

The results in Table 1 indicates that the Wald, LR, JK and PLS methods produced AS intervals with parameter σ although none of the proposed CI methods produced C intervals. Subsequently, it is also evident that the number of AC intervals decreased with the increase in sample size, see Table 2. However, higher number of AC intervals are produced when the percentage of truncation is high (model 3) compared to when no or low truncation is observed (model 1 and 2) specifically with the Wald and PLS method. Also, the parameterization of $\log(\sigma)$ for parameter σ did not improve the performance of the Wald method as the number of AC intervals remained the same for model 1 and 2 and increased under model 3, refer to Tables 1 and 2. It is interesting to note that the LR based intervals appear to be more symmetrical for parameter σ provided that the number of exact observations are extremely large, preferably greater than 250. Conversely, with real lifetime or clinical data this may not be plausible.

TABLE 1: Number of AC, C and AS confidence intervals for parameters in model 1, 2 and 3

Method	Parameter	Model 1			Model 2			Model 3		
		AC	C	AS	AC	C	AS	AC	C	AS
Wald	σ	3	0	6	2	0	6	4	0	6
	β_0	2	0	0	2	0	0	3	0	3
	β_1	1	0	0	2	0	0	2	0	0
LR	σ	2	0	5	3	0	6	3	0	6
	β_0	2	0	0	0	0	0	2	0	1
	β_1	2	0	0	1	0	0	1	0	0
JK	σ	2	0	6	2	0	6	2	0	6
	β_0	0	0	0	0	0	0	1	0	1
	β_1	0	0	0	1	0	0	0	0	0
PLS	σ	3	0	6	2	0	6	6	0	6

TABLE 2: Estimated error probabilities for parameter σ under models 1, 2 and 3

	n	Model 1			Model 2			Model 3		
		lep	rep	tep	lep	rep	tep	lep	rep	tep
Wald	20	0.001	0.139	0.140	0.001	0.131	0.131	0.000	0.148	0.148
	30	0.003	0.104	0.107	0.002	0.106	0.108	0.003	0.113	0.116
	80	0.005	0.063	0.068	0.008	0.053	0.061	0.005	0.075	0.079
	100	0.008	0.051	0.059	0.005	0.055	0.060	0.008	0.057	0.065
	200	0.012	0.050	0.061	0.009	0.053	0.062	0.012	0.046	0.058
	250	0.012	0.042	0.054	0.011	0.048	0.059	0.013	0.042	0.055
LR	20	0.006	0.061	0.067	0.010	0.064	0.074	0.006	0.070	0.076
	30	0.016	0.058	0.074	0.011	0.060	0.071	0.012	0.058	0.069
	80	0.014	0.040	0.054	0.016	0.048	0.064	0.014	0.050	0.063
	100	0.014	0.039	0.052	0.014	0.037	0.051	0.016	0.037	0.053
	200	0.017	0.033	0.050	0.015	0.039	0.053	0.018	0.031	0.049
	250	0.020	0.029	0.049	0.016	0.036	0.051	0.018	0.034	0.051
JK	20	0.009	0.071	0.080	0.009	0.069	0.078	0.010	0.078	0.088
	30	0.011	0.060	0.071	0.011	0.048	0.059	0.008	0.071	0.079
	80	0.013	0.041	0.054	0.016	0.048	0.063	0.012	0.048	0.060
	100	0.014	0.045	0.059	0.017	0.043	0.059	0.012	0.043	0.055
	200	0.013	0.037	0.050	0.015	0.039	0.054	0.016	0.031	0.046
	250	0.020	0.036	0.056	0.013	0.041	0.054	0.016	0.035	0.051
PLS	20	0.000	0.147	0.147	0.000	0.139	0.139	0.000	0.126	0.126
	30	0.000	0.084	0.084	0.000	0.147	0.147	0.001	0.179	0.180
	80	0.003	0.077	0.080	0.001	0.055	0.056	0.005	0.109	0.114
	100	0.002	0.051	0.053	0.001	0.055	0.056	0.005	0.071	0.076
	200	0.004	0.051	0.055	0.005	0.054	0.059	0.010	0.055	0.065
	250	0.006	0.039	0.045	0.006	0.045	0.051	0.012	0.054	0.066

In contrast, the Wald, LR and JK methods do not produce any AS and C intervals for parameter β_0 under model 1 and 2 with AC intervals detected only for smaller sample sizes e.g. $n = 20$ and 30 and at higher proportion of truncation (model 3), refer to Tables 1 and 3. Alternatively, the

JK method outperform all the remaining methods by producing the least AC C.I for parameter β_0 under models 1,2 and 3 as the estimated tep closer to 0.05 and the error probabilities appears to be more symmetrical compared to the Wald or the LR method, refer to Tables 1 and 3.

Subsequently all the proposed CI methods work well for the covariate parameter β_1 , but as the estimated error probabilities are more symmetrical in addition that the estimated tep are closer to 0.05, we can say that the JK performs the best, refer to Tables 1 and 4.

On the other hand, the presence of both censoring and truncation affect the performance of the Wald, LR and JK CI specifically for parameter σ and β_0 as more AS and AC CI are produced regardless of large sample sizes under models 4, 5 and 6, refer to the output in Tables 5, 6 and 7. Also, the tpe is far from npe of 0.05 and the distance increases with the increase in the sample size, refer to Tables 6 and 7.

TABLE 3: Estimated error probabilities for parameter β_0 under model 1, 2 and 3

	n	Model 1			Model 2			Model 3		
		lep	rep	tep	lep	rep	tep	lep	rep	tep
Wald	20	0.044	0.038	0.082	0.035	0.043	0.078	0.051	0.028	0.078
	30	0.035	0.032	0.067	0.044	0.041	0.085	0.043	0.025	0.068
	80	0.025	0.027	0.052	0.028	0.029	0.057	0.042	0.014	0.056
	100	0.025	0.026	0.050	0.022	0.030	0.052	0.027	0.026	0.052
	200	0.026	0.021	0.047	0.026	0.030	0.056	0.029	0.022	0.051
	250	0.029	0.026	0.055	0.023	0.027	0.050	0.036	0.031	0.067
LR	20	0.037	0.033	0.070	0.031	0.027	0.058	0.038	0.030	0.067
	30	0.030	0.027	0.057	0.028	0.034	0.062	0.033	0.027	0.059
	80	0.034	0.024	0.058	0.029	0.025	0.053	0.031	0.016	0.047
	100	0.027	0.032	0.059	0.021	0.030	0.051	0.024	0.028	0.052
	200	0.031	0.033	0.064	0.025	0.030	0.055	0.027	0.025	0.052
	250	0.030	0.022	0.051	0.023	0.027	0.050	0.034	0.033	0.067
JK	20	0.024	0.029	0.052	0.027	0.022	0.049	0.032	0.022	0.054
	30	0.022	0.024	0.046	0.030	0.024	0.054	0.031	0.020	0.050
	80	0.018	0.023	0.041	0.031	0.024	0.055	0.029	0.021	0.050
	100	0.026	0.030	0.056	0.019	0.025	0.043	0.024	0.024	0.048
	200	0.025	0.019	0.044	0.025	0.030	0.054	0.027	0.024	0.050
	250	0.028	0.022	0.050	0.023	0.024	0.047	0.035	0.030	0.065

TABLE 4: Estimated error probabilities for parameter β_1 under model 1, 2 and 3

	n	Model 1			Model 2			Model 3		
		lep	rep	tep	lep	rep	tep	lep	rep	tep
Wald	20	0.047	0.034	0.081	0.039	0.036	0.075	0.039	0.039	0.077
	30	0.031	0.030	0.060	0.040	0.033	0.073	0.030	0.037	0.067
	80	0.030	0.025	0.054	0.026	0.027	0.052	0.032	0.030	0.062
	100	0.034	0.024	0.057	0.027	0.025	0.051	0.027	0.020	0.047
	200	0.021	0.026	0.047	0.026	0.026	0.052	0.030	0.032	0.061
	250	0.026	0.025	0.051	0.023	0.026	0.048	0.022	0.030	0.052
LR	20	0.038	0.027	0.065	0.038	0.027	0.065	0.034	0.034	0.068
	30	0.031	0.033	0.064	0.029	0.032	0.061	0.027	0.033	0.059
	80	0.027	0.027	0.054	0.030	0.022	0.052	0.032	0.028	0.060
	100	0.022	0.025	0.046	0.024	0.024	0.048	0.026	0.020	0.046
	200	0.019	0.023	0.042	0.026	0.025	0.050	0.029	0.031	0.060
	250	0.025	0.028	0.053	0.023	0.026	0.048	0.020	0.030	0.050
JK	20	0.036	0.019	0.055	0.028	0.027	0.054	0.023	0.027	0.049
	30	0.030	0.027	0.057	0.034	0.036	0.070	0.026	0.034	0.059
	80	0.028	0.023	0.051	0.021	0.030	0.051	0.028	0.025	0.053
	100	0.022	0.025	0.046	0.026	0.023	0.049	0.026	0.019	0.045
	200	0.021	0.026	0.047	0.026	0.024	0.050	0.031	0.030	0.060
	250	0.024	0.025	0.048	0.023	0.027	0.049	0.022	0.030	0.052

Nevertheless, it is observed that the PLS method performs fairly well for parameter σ at moderate sample sizes e.g. $n = 80$ and 100 particularly for model 4 as there is reduction in the number of AC and the intervals are more symmetric; even so, when low or high proportion of truncation is observed with censoring (model 5 and model 6) the PLS method generated many AC intervals even at larger sample sizes and C intervals are produced at moderate sample sizes, see Tables 5 and 6.

Conversely, the presence of censoring do not affect the confidence intervals of parameter β_1 as the Wald, LR and JK method perform well specifically for model 4 as all the proposed CI method do not produce any AS or C intervals; AC intervals are produced only for small samples $n = 20$ and 30, refer to Tables 5 and 8. However the JK method offers a better option for parameter β_1 particularly when data is equally left-truncated with censored observations (models 5 and 6) as least number of AS, C and AC intervals are generated and the estimated error probabilities are more symmetric and tpe closer to npe of 0.05, refer to Tables 5 and 8.

TABLE 5: Number of AC, C and AS confidence intervals for parameters in model 4, 5 and 6

Method	Parameter	Model 4			Model 5			Model 6		
		AC	C	AS	AC	C	AS	AC	C	AS
Wald	σ	5	0	6	5	0	6	5	0	6
	β_0	6	0	6	6	0	6	6	0	6
	β_1	2	0	0	2	0	1	2	0	1
LR	σ	6	0	4	6	0	5	6	0	5
	β_0	6	0	6	6	0	6	6	0	6
	β_1	2	0	0	2	0	1	2	0	1
JK	σ	4	0	6	4	0	6	4	0	6
	β_0	4	0	6	4	0	6	5	0	6
	β_1	0	0	1	0	0	1	0	0	1
PLS	σ	2	0	5	5	1	5	4	2	5

TABLE 6: Estimated error probabilities for parameter σ under model 4, 5 and 6

	n	Model 4			Model 5			Model 6		
		lep	rep	tep	lep	rep	tep	lep	rep	tep
Wald	20	0.002	0.091	0.093	0.004	0.091	0.094	0.001	0.103	0.104
	30	0.009	0.061	0.070	0.007	0.065	0.072	0.004	0.070	0.073
	80	0.045	0.016	0.061	0.038	0.013	0.051	0.041	0.014	0.055
	100	0.051	0.015	0.066	0.052	0.018	0.069	0.051	0.019	0.069
	200	0.114	0.003	0.117	0.114	0.006	0.120	0.111	0.005	0.115
	250	0.142	0.005	0.147	0.154	0.004	0.158	0.155	0.002	0.156
LR	20	0.028	0.039	0.066	0.037	0.038	0.075	0.028	0.040	0.068
	30	0.040	0.025	0.065	0.042	0.023	0.065	0.045	0.024	0.069
	80	0.081	0.011	0.092	0.080	0.013	0.093	0.071	0.012	0.083
	100	0.084	0.009	0.093	0.095	0.010	0.104	0.095	0.010	0.105
	200	0.158	0.002	0.160	0.156	0.005	0.161	0.161	0.003	0.163
	250	0.181	0.003	0.184	0.194	0.003	0.197	0.184	0.002	0.186
JK	20	0.016	0.045	0.061	0.011	0.046	0.057	0.018	0.051	0.068
	30	0.023	0.036	0.059	0.022	0.036	0.058	0.022	0.039	0.061
	80	0.062	0.012	0.073	0.064	0.014	0.078	0.049	0.012	0.061
	100	0.067	0.008	0.075	0.075	0.011	0.086	0.069	0.013	0.082
	200	0.134	0.005	0.139	0.133	0.005	0.137	0.130	0.004	0.133
	250	0.165	0.005	0.170	0.174	0.003	0.177	0.160	0.001	0.160
PLS	20	0.002	0.091	0.093	0.004	0.091	0.094	0.001	0.103	0.104
	30	0.009	0.061	0.070	0.007	0.065	0.072	0.004	0.070	0.073
	80	0.045	0.016	0.061	0.038	0.013	0.051	0.041	0.014	0.055
	100	0.051	0.015	0.066	0.052	0.018	0.069	0.051	0.019	0.069
	200	0.114	0.003	0.117	0.114	0.006	0.120	0.111	0.005	0.115
	250	0.142	0.005	0.147	0.154	0.004	0.158	0.155	0.002	0.156

TABLE 7: Estimated error probabilities for parameter β_0 under model 4, 5 and 6

	n	Model 4			Model 5			Model 6		
		lep	rep	tep	lep	rep	tep	lep	rep	tep
Wald	20	0.061	0.024	0.084	0.062	0.026	0.087	0.082	0.020	0.101
	30	0.061	0.015	0.076	0.065	0.017	0.082	0.076	0.008	0.084
	80	0.102	0.003	0.105	0.099	0.011	0.110	0.085	0.007	0.092
	100	0.098	0.004	0.102	0.124	0.004	0.128	0.097	0.004	0.101
	200	0.182	0.001	0.183	0.179	0.004	0.183	0.157	0.002	0.159
	250	0.210	0.001	0.211	0.194	0.001	0.195	0.195	0.001	0.196
LR	20	0.068	0.015	0.083	0.059	0.012	0.071	0.060	0.017	0.076
	30	0.075	0.012	0.087	0.066	0.015	0.081	0.061	0.014	0.074
	80	0.092	0.003	0.095	0.092	0.003	0.095	0.078	0.007	0.085
	100	0.126	0.006	0.132	0.105	0.006	0.111	0.090	0.004	0.094
	200	0.193	0.002	0.195	0.177	0.004	0.181	0.148	0.000	0.148
	250	0.208	0.003	0.210	0.200	0.001	0.201	0.174	0.001	0.175
JK	20	0.039	0.012	0.051	0.043	0.010	0.053	0.049	0.012	0.061
	30	0.047	0.008	0.055	0.053	0.010	0.063	0.059	0.011	0.069
	80	0.081	0.003	0.084	0.084	0.005	0.089	0.094	0.006	0.099
	100	0.112	0.006	0.118	0.095	0.006	0.101	0.082	0.002	0.084
	200	0.170	0.001	0.171	0.169	0.004	0.173	0.157	0.002	0.158
	250	0.204	0.000	0.204	0.197	0.001	0.198	0.190	0.002	0.191

TABLE 8: Estimated error probabilities for parameter β_1 under model 4, 5 and 6

	n	Model 4			Model 5			Model 6		
		lep	rep	tep	lep	rep	tep	lep	rep	tep
Wald	20	0.044	0.036	0.080	0.039	0.034	0.072	0.041	0.038	0.079
	30	0.034	0.031	0.064	0.041	0.038	0.078	0.031	0.040	0.071
	80	0.028	0.024	0.052	0.025	0.024	0.049	0.021	0.033	0.054
	100	0.031	0.024	0.055	0.028	0.031	0.058	0.036	0.031	0.067
	200	0.029	0.022	0.051	0.029	0.018	0.047	0.032	0.027	0.059
	250	0.029	0.022	0.050	0.024	0.027	0.051	0.033	0.020	0.052
LR	20	0.040	0.028	0.068	0.032	0.035	0.067	0.034	0.032	0.066
	30	0.038	0.029	0.066	0.032	0.034	0.066	0.027	0.031	0.058
	80	0.022	0.018	0.040	0.036	0.026	0.062	0.033	0.034	0.067
	100	0.025	0.020	0.045	0.035	0.024	0.059	0.036	0.029	0.065
	200	0.026	0.019	0.045	0.028	0.018	0.046	0.035	0.022	0.056
	250	0.028	0.021	0.048	0.025	0.025	0.050	0.032	0.020	0.052
JK	20	0.032	0.020	0.052	0.029	0.027	0.056	0.025	0.027	0.052
	30	0.025	0.023	0.047	0.035	0.024	0.058	0.033	0.026	0.059
	80	0.026	0.024	0.050	0.021	0.022	0.043	0.026	0.026	0.052
	100	0.025	0.022	0.047	0.031	0.025	0.055	0.028	0.017	0.045
	200	0.023	0.020	0.043	0.029	0.018	0.047	0.030	0.025	0.055
	250	0.029	0.022	0.051	0.026	0.025	0.051	0.028	0.029	0.057

6. CONCLUSIONS

In general, the Wald, LR and JK CI methods generate least number of AS,C and AC observed when exact observations are available (models 1,2 and 3) compared to when data is censored (models 4,5 and 6).

We recommend the JK method for parameter β_0 and β_1 when observations are truncated and exact failure times are available as the estimated error probabilities are all symmetrical with the tpe closer to 0.05 although the proportion of truncation is high.

Further, as all the suggested CI methods produce error probabilities that are not symmetrical for parameter σ under models 1, 2 and 3, a search for an alternative C.I method is therefore necessary.

In conclusion, since left-truncated data is often skewed, fitting them to a symmetrical distribution would disregard some observations on the left-tail of a specified distribution as demonstrated by Cain et.al, 2011. Also as data is equally skewed, the assumption of normality often fails as it can't fully capture the sampling distribution of the sample statistics being studied, subsequently resulting in the poor performance of the Wald and likelihood ratio methods in constructing the confidence intervals for the parameter estimates specifically when higher proportion of truncation and censoring is present in the data. Under these circumstances, one may opt for the JK CI method instead; as the estimated tpe are closer to the npe of 0.05 compared to the Wald method.

As a rule of thumb, we propose that the JK based confidence intervals are only used with parameter β_1 when the proportion of censoring is small or equally the number of observed failures are large in the presence of left-truncation.

The parameterization of $\log(\sigma)$ may improve the performance of the Wald method and is therefore recommended for moderate sample sizes when no truncation is observed in the presence of censoring. Nevertheless, one may have to look into alternative methods such as bootstrap in estimating confidence limits for parameters σ and β_0 or equally for parameter β_1 characteristically when higher proportion of censoring is observed in the presence of left truncation, as bootstrap intervals are based on distribution of data in hand and not asymptotic normality.

We also recommend to include more number of new cases rather than using existing cases alone in a way to reduce truncation and to decrease sampling bias. This may equally help to improve the performance of the proposed CI methods.

ACKNOWLEDGMENTS

We would like to extend our gratitude to Fundamental Research Grant Scheme (FRGS), VOT 5524226, University Putra Malaysia and Dr.Patricia Tai, University of Saskatchewan, Saskatoon, Canada.

REFERENCES

Arasan, J. (2008). Parallel Weibull Regression Model. *Pertanika Journal of Science & Technology*. **16**: 83-95.

Arasan, J., and Lunn, M. (2008). Alternative interval estimation for parameters of bivariate exponential model with time varying covariate. *Computational Statistics*. **23**: 605-622.

Cain, K .C., Harlow, S. D., Little, R . J., Nan, B., Yosef, M., Taffe, J . R., and Elliott, M . R. (2011). Bias due to left truncation and left censoring in longitudinal studies of developmental and disease processes. *American journal of epidemiology*. **173**: 1078-84.

Doganaksoy, N., and Schmee J. (1993). Comparison of approximate confidence interval distributions used in life-data analysis. *Technometrics* **2**: 175-84 .

Grover, G., and Sabharwal, A. (2012). A parametric approach to estimate survival time of diabetic nephropathy with left truncated and right censored data. *International Journal of Statistics and Probability*. **1**:128-137.

Lawless, J.F. (1982). *Statistical Model and Methods for Lifetime Data.* Wiley: New York.

Limpert, E., Stahel W.A., and Abbt M. (2001). Log-normal distributions across the sciences: Keys and clues. *BioScience.* **51**: 341-52 .

Mitra, D. (2013). *Likelihood Inference for Left-Truncated and Right Censored Lifetime Data*. PhD thesis: McMaster University.

Reeds, J.A. (1978). Jackknifing maximum likelihood estimates. *The Annals of Statistics*. **6**: 727-739.

Royston, P. (2001). The log-normal distribution as a model for survival time in cancer, with an emphasis on prognostic factors. *Stat Neerlandica*. **55**: 89-104.

Schmee, J., Gladstein, D., and Nelson, W. (1985). Confidence Limits for Parameters of a Normal Distribution from Singly Censored Sampels, Using Maximum Likelihood. *Technometrics*. **27**: 119-128 .

Shen, P. (2009). Semiparametric analysis of survival data with left truncation and right censoring. *Computational Statistics and Data Analysis*. **53**: 4417-32.

Tai, P., Tonita, J., Yu, E., and Skarsgard, D. (2003). Twenty-year folow-up study of long-term survival of limited-stage small-cell lung cancer and overview of prognostic and treatment factors. *Int J.Radiatioan Oncology Bio/Phys*. **56**: 626-633.