Journal of Modern Applied Statistical Methods

Volume 16 | Issue 2 Article 32

December 2017

Inferential Procedures for Log Logistic Distribution with Doubly Interval Censored Data

Yue Fang Loh
Universiti Putra Malaysia, Seri Kembangan, Malaysia, yuefangloh@yahoo.com

Jayanthi Arasan
Universiti Putra Malaysia, Seri Kembangan, Malaysia, jayanthi@upm.edu.my

Habshah Midi
Universiti Putra Malaysia, Seri Kembangan, Malaysia, habshah@upm.edu.my

M. R. Abu Bakar
Universiti Putra Malaysia, Seri Kembangan, Malaysia, mrizam@upm.edu.my

Inferential Procedures for Log Logistic Distribution with Doubly Interval Censored Data

Cover Page Footnote
We gratefully acknowledge financial support from the Ministry of Education Malaysia. The research leading to these results has received funding from the Fundamental Research Grant Scheme (FRGS 2014) under vote no. 5524673.

Journal of Modern Applied Statistical Methods
November 2017, Vol. 16, No. 2, 581-603.
doi: 10.22237/jmasm/1509496320

ISSN 1538 − 9472

Inferential Procedures for Log Logistic Distribution with Doubly Interval Censored Data

Yue Fang Loh
Universiti Putra Malaysia
Seri Kembangan, Malaysia

Jayanthi Arasan
Universiti Putra Malaysia
Seri Kembangan, Malaysia

Habshah Midi
Universiti Putra Malaysia
Seri Kembangan, Malaysia

M. R. Abu Bakar
Universiti Putra Malaysia
Seri Kembangan, Malaysia

The log logistic model with doubly interval censored data is examined. Three methods of constructing confidence interval estimates for the parameter of the model were compared and discussed. The results of the coverage probability study indicated that the Wald outperformed the likelihood ratio and jackknife inferential procedures.

Keywords: doubly interval censored, jackknife, likelihood ratio, log logistic, Wald

Introduction

Doubly interval censored (DIC) data is a type of interval censored (IC) data, which often arises in disease progression studies where the survival time of interest is the elapsed time between two related events that are possibly IC (De Gruttola & Lagakos, 1989; Sun, 2004). Let A and B denote the times of the occurrences of the two events with $A \leq B$ and the survival time, $Y = B - A$. The observations in Y are DIC when A and B are observed in an interval form $A \in (A_L, A_R]$ and $B \in (B_L, B_R]$ respectively with $A_L \leq A_R$ and $B_L \leq B_R$.

A well-known example of DIC data in real life can be seen in acquired immune deficiency syndrome (AIDS) cohort studies where the A and B represent the human immunodeficiency virus (HIV) infection and AIDS diagnosis time respectively, and Y is the AIDS incubation time. The HIV infection time is often determined through periodic blood tests for which it is only known to occur between the last negative test and the first positive test and therefore observations are commonly interval censored. Also, observations on the diagnosis of AIDS could be either right censored (RC) or IC due to, for example, the end of the study

Yue Fang Loh is a PhD student in the Department of Mathematics. Email at yuefangloh@yahoo.com.

and the periodic follow up nature of the study design, thus yielding DIC data on *Y* (De Gruttola & Lagakos, 1989; Kim, et al., 1993).

Statistical analysis of DIC data was first discussed by De Gruttola & Lagakos (1989) via nonparametric approach to obtain the maximum likelihood estimator of the joint distribution of HIV infection time and AIDS incubation time without truncated data. Since then, many researchers extend the statistical analysis of DIC data, especially in the context of AIDS, to include truncation effect and covariates information in nonparametric and semiparametric approaches. Authors who have contributed include Bacchetti (1990); Bacchetti & Jewell (1991); Kim, et al. (1993); Jewell (1994); Jewell et al. (1994); Gómez & Lagakos (1994); Sun (1995, 1997); Tu (1995); Gómez & Calle (1999); Goggins, et al. (1999); Sun, et al. (1999); Fang & Sun (2001); Pan (2001); and Lim, et al. (2002). The Bayesian approach has gained some attention in analysis of DIC data in recent years for severe acute respiratory syndrome (SARS) disease incubation time (McBryde, et al., 2006) and time to caries development in children (Komárek, et al., 2005; Komárek & Lesaffre, 2006, 2008; Jara, et al., 2010).

Brookmeyer & Goedart (1989) proposed a two-stage parametric regression model for jointly estimating the effects of covariates on risk of HIV infection as well as risk of progression to AIDS disease once infected. They assumed the HIV infection time, *A*, follows the piecewise exponential distribution and the onset of AIDS disease, *B*, follows the Weibull distribution. The likelihood function was presented and maximum likelihood estimates (MLEs) were obtained via Newton Raphson iterative procedure. They considered special cases of DIC data where *A* could be only IC and *B* could be only RC or observed exactly (OE). The proposed model was later adapted by Darby, et al. (1990) and fitted to data on the development of AIDS in hemophiliacs in the United Kingdom who are seropositive for HIV.

Reich, et al. (2009) studied two procedures for estimating the incubation time distribution. The first procedure defined the likelihood function with DIC data scheme and obtained the MLEs parametrically. They proposed the following likelihood function and obtained the MLE of parameter γ affecting *Y*, while parameter λ affecting *A* is assumed to be known,

$$L(\gamma;\lambda)=\sum_{i=1}^{n}\left\{\int_{a_{L_i}}^{a_{R_i}}\int_{b_{L_i}}^{b_{R_i}} f_A(a)\, f_T(b-a)\,db\,da\right\}^{d_{DC_i}} \times\left\{S_T\left(t_{L_i}\right)-S_T\left(t_{R_i}\right)\right\}^{d_{IC_i}} f_T\left(t_i\right)^{d_{OE_i}}. \tag{1}$$

The variables δ_{DC_i} , δ_{IC_i} , and δ_{OE_i} serve as indicators to identify whether the i^{th} subject is DIC, IC or OE. The second procedure involves a data reduction technique to reduce the DIC data to IC data and obtain the MLEs parametrically. They assumed *A* follows the uniform distribution and *Y* follows the log normal distribution.

Kiani & Arasan (2012) proposed a parametric model for analyzing DIC data by assuming that both *A* and *Y* follow the exponential distribution. Following Kiani & Arasan, proposed here is a parametric model that could be used to analyze DIC data. It is assumed that the first event time *A* is uniformly distributed and the survival time *Y* follows a special case of the log logistic distribution with $\gamma = 1$. We assume independent censoring for both *A* and *Y* (Oller, et al., 2004) and independence between *A* and *Y*, which are classical assumptions for the treatment of DIC survival times. All simulation studies were performed using the R programming language (R Core Team, 2015).

The Model

Let the survival time of interest *Y* be a non-negative continuous random variable with density function $f_Y(y)$ whereas $f_A(a)$ and $f_B(b)$ denote the density function of the times to the occurrences of the first event *A* and second event *B* respectively. Following Reich, et al. (2009), the distribution of *b* could be obtained if *a* is given and $f_Y(y)$ is known. Thus,

$$f_{B|A}\left(b\,|\,a\right) = f_Y\left(b - a\,|\,a\right). \tag{2}$$

Thus, the joint density function of *A* and *B* would be,

$$f_{A,B}\left(a,b\right) = f_{B|A}\left(b\,|\,a\right) f_A\left(a\right) = f_Y\left(b - a\,|\,a\right) f_A\left(a\right) = f_Y\left(b - a\right) f_A\left(a\right) \tag{3}$$

where $Y = B - A$ and *A* is assumed to be independent of *Y*. Therefore, the likelihood for a DIC data is as follows,

$$L = \int_{a_L}^{a_R}\int_{b_L}^{b_R} f_{A,B}\left(a,b\right) db\,da = \int_{a_L}^{a_R}\int_{b_L}^{b_R} f_Y\left(b-a\right) f_A\left(a\right) db\,da \tag{4}$$

The distributional assumptions on both *A* and *Y* allow us to construct the likelihood function of all data. Here, we assume $A \sim U(u_L, u_R)$ and *Y* follows the log logistic distribution with scale parameter $-\infty < \lambda < \infty$ and known shape parameter $\gamma = 1$. The density function of A is given by

$$f_A(a) = \frac{1}{u_R - u_L}, \tag{5}$$

and the survival function is

$$S_A(a) = \frac{u_R - a}{u_R - u_L}. \tag{6}$$

Similarly, the density and survival function of Y are given respectively as follows:

$$f_Y(y) = \frac{e^{\lambda}}{\left(1 + e^{\lambda} y\right)^2}, \tag{7}$$

$$S_Y(y) = \frac{1}{1 + e^{\lambda} y}. \tag{8}$$

DIC data include IC and RC lifetime data as special cases (Kalbfleisch & Prentice, 2002; Sun, 1998), therefore a comprehensive likelihood function containing all contributions with respect to each type of data need to be defined. For the i^{th} subject, in cases where both A and B are IC, Y is DIC and the likelihood contribution is

$$\begin{aligned} L_{1_i}(\lambda) &= \int_{a_{L_i}}^{a_{R_i}} \int_{b_{L_i}}^{b_{R_i}} f_Y(b - a) f_A(a)\, db\, da \\ &= \frac{1}{e^{\lambda}\left(u_R - u_L\right)} \log\left[\frac{\left\{1 + e^{\lambda}\left(b_{R_i} - a_{R_i}\right)\right\}\left\{1 + e^{\lambda}\left(b_{L_i} - a_{L_i}\right)\right\}}{\left\{1 + e^{\lambda}\left(b_{R_i} - a_{L_i}\right)\right\}\left\{1 + e^{\lambda}\left(b_{L_i} - a_{R_i}\right)\right\}}\right]. \end{aligned} \tag{9}$$

In cases where A is IC and B is RC, the likelihood contribution is

$$L_{2_i}(\lambda) = \int_{a_{L_i}}^{a_{R_i}} \int_{b_{L_i}}^{\infty} f_Y(b - a) f_A(a)\, db\, da = \frac{1}{e^{\lambda}(u_R - u_L)} \log\left[\frac{1 + e^{\lambda}\left(b_{L_i} - a_{L_i}\right)}{1 + e^{\lambda}\left(b_{L_i} - a_{R_i}\right)}\right]. \tag{10}$$

In cases where either A or B is OE while the other is IC, Y becomes IC and the interval $(y_{L_i}, y_{R_i}]$ is equal to $(b_i - a_{R_i}, b_i - a_{R_i}]$ when A is IC and $(b_{L_i} - a_i, b_{R_i} - a_i]$ when B is IC. The likelihood contribution is

$$L_{3_i}(\lambda) = \int_{y_{L_i}}^{y_{R_i}} f_Y(y)\, dy = S_Y(y_{L_i}) - S_Y(y_{R_i}) = \frac{e^{\lambda}\left(y_{R_i} - y_{L_i}\right)}{\left(1 + e^{\lambda} y_{L_i}\right)\left(1 + e^{\lambda} y_{R_i}\right)}. \tag{11}$$

In cases where A is OE and B is RC, Y becomes RC and $y_{D_i} = b_{L_i} - a_i$, the likelihood contribution is

$$L_{4_i}(\lambda) = S_Y(y_{D_i}) = \frac{1}{1 + e^{\lambda} y_{D_i}}. \tag{12}$$

In cases where both A and B are OE, Y becomes OE and $y_i = b_i - a_i$, the likelihood contribution is

$$L_{5_i}(\lambda) = f_Y(y_i) = \frac{e^{\lambda}}{\left(1 + e^{\lambda} y_i\right)^2}. \tag{13}$$

The censoring indicators for the i^{th} subject are defined as follows,

$$\begin{aligned}
\delta_{DC_i} &= 1 \text{ if } Y \text{ is DIC, 0 otherwise;} \\
\delta_{IR_i} &= 1 \text{ if } A \text{ is IC and } B \text{ is RC, 0 otherwise;} \\
\delta_{IC_i} &= 1 \text{ if } Y \text{ is IC, 0 otherwise;} \\
\delta_{RC_i} &= 1 \text{ if } Y \text{ is RC, 0 otherwise;} \\
\delta_{OE_i} &= 1 \text{ if } Y \text{ is OE, 0 otherwise;}
\end{aligned} \tag{14}$$

where $\delta_{OE_i} = 1 - (\delta_{DC_i} + \delta_{IR_i} + \delta_{IC_i} + \delta_{RC_i})$. Following that, the likelihood function for the full sample can be written as

$$
\begin{aligned}
L(l) = \prod_{i=1}^{n} & \left(\frac{1}{e^{l}\left(u_R - u_L\right)} \log \left[\frac{\left\{1+e^{l}\left(b_{R_i} - a_{R_i}\right)\right\}\left\{1+e^{l}\left(b_{L_i} - a_{L_i}\right)\right\}}{\left\{1+e^{l}\left(b_{R_i} - a_{L_i}\right)\right\}\left\{1+e^{l}\left(b_{L_i} - a_{R_i}\right)\right\}} \right] \right)^{d_{DC_i}} \\
& \times \left[\frac{1}{e^{l}\left(u_R - u_L\right)} \times \log \left\{ \frac{1+e^{l}\left(b_{L_i} - a_{L_i}\right)}{1+e^{l}\left(b_{L_i} - a_{R_i}\right)} \right\} \right]^{d_{IR_i}} \\
& \times \left\{ \frac{e^{l}\left(y_{R_i} - y_{L_i}\right)}{\left(1+e^{l} y_{L_i}\right)\left(1+e^{l} y_{R_i}\right)} \right\}^{d_{IC_i}} \times \left(\frac{1}{1+e^{l} y_{D_i}} \right)^{d_{RC_i}} \\
& \times \left\{ \frac{e^{l}}{\left(1+e^{l} y_i\right)^2} \right\}^{d_{OE_i}},
\end{aligned} \tag{15}
$$

and the log likelihood function is

$$
(\lambda) = \sum_{i=1}^{n} \left\{
\begin{array}{l}
\delta_{DC_i} \left(-\lambda - \log\left(u_R - u_L\right) + \log \left[\begin{array}{l} \log\left\{1+e^{\lambda}\left(b_{R_i} - a_{R_i}\right)\right\} \\ +\log\left\{1+e^{\lambda}\left(b_{L_i} - a_{L_i}\right)\right\} \\ -\log\left\{1+e^{\lambda}\left(b_{R_i} - a_{L_i}\right)\right\} \\ -\log\left\{1+e^{\lambda}\left(b_{L_i} - a_{R_i}\right)\right\} \end{array} \right] \right) \\
+\delta_{IR_i} \left(-\lambda - \log\left(u_R - u_L\right) + \log \left[\begin{array}{l} \log\left\{1+e^{\lambda}\left(b_{L_i} - a_{L_i}\right)\right\} \\ -\log\left\{1+e^{\lambda}\left(b_{L_i} - a_{R_i}\right)\right\} \end{array} \right] \right) \\
+\delta_{IC_i} \left\{ \lambda + \log\left(y_{R_i} - y_{L_i}\right) - \log\left(1+e^{\lambda} y_{L_i}\right) - \log\left(1+e^{\lambda} y_{R_i}\right) \right\} \\
-\delta_{RC_i} \log\left(1+e^{\lambda} y_{D_i}\right) + \delta_{OE_i} \left\{ \lambda - 2\log\left(1+e^{\lambda} y_i\right) \right\}
\end{array}
\right\} \tag{16}
$$

Let

$$
\begin{aligned}
A_{1i} &= 1 + e^{l}\left(b_{R_i} - a_{R_i}\right), & A_{9i} &= \frac{e^{l}\left(b_{R_i} - a_{R_i}\right)}{1 + e^{l}\left(b_{R_i} - a_{R_i}\right)}, \\
A_{2i} &= 1 + e^{l}\left(b_{L_i} - a_{L_i}\right), & A_{10i} &= \frac{e^{l}\left(b_{L_i} - a_{L_i}\right)}{1 + e^{l}\left(b_{L_i} - a_{L_i}\right)}, \\
A_{3i} &= 1 + e^{l}\left(b_{R_i} - a_{L_i}\right), & A_{11i} &= \frac{e^{l}\left(b_{R_i} - a_{L_i}\right)}{1 + e^{l}\left(b_{R_i} - a_{L_i}\right)}, \\
A_{4i} &= 1 + e^{l}\left(b_{L_i} - a_{R_i}\right), & A_{12i} &= \frac{e^{l}\left(b_{L_i} - a_{R_i}\right)}{1 + e^{l}\left(b_{L_i} - a_{R_i}\right)}, \\
A_{5i} &= 1 + e^{l} y_{L_i}, & A_{13i} &= \frac{e^{l} y_{L_i}}{1 + e^{l} y_{L_i}}, \\
A_{6i} &= 1 + e^{l} y_{R_i}, & A_{14i} &= \frac{e^{l} y_{R_i}}{1 + e^{l} y_{R_i}}, \\
A_{7i} &= 1 + e^{l} y_{D_i}, & A_{15i} &= \frac{e^{l} y_{D_i}}{1 + e^{l} y_{D_i}}, \\
A_{8i} &= 1 + e^{l} y_{i}, & A_{16i} &= \frac{e^{l} y_{i}}{1 + e^{l} y_{i}},
\end{aligned}
\tag{17}
$$

The first and second partial derivatives of the log likelihood function are given as follows,

$$
\frac{\partial \ell(l)}{\partial l} = \sum_{i=1}^{n} \left[\begin{array}{l}
d_{DC_i}\left\{ -1 + \left(\log \frac{A_{1i}A_{2i}}{A_{3i}A_{4i}} \right)^{-1} \left(A_{9i} + A_{10i} - A_{11i} - A_{12i} \right) \right\} \\
+ d_{IR_i}\left\{ -1 + \left(\log \frac{A_{2i}}{A_{4i}} \right)^{-1} \left(A_{10i} - A_{12i} \right) \right\} \\
+ d_{IC_i}\left(1 - A_{13i} - A_{14i}\right) - d_{RC_i} A_{15i} + d_{OE_i}\left(1 - 2A_{16i}\right)
\end{array} \right],
\tag{18}
$$

$$\frac{\partial^2 \ell(\lambda)}{\partial \lambda^2} = \sum_{i=1}^{n} \left[\begin{array}{l} \delta_{DC_i} \left(\log \frac{A_{1i} A_{2i}}{A_{3i} A_{4i}} \right)^{-2} \left\{ \begin{array}{l} \left(\log \frac{A_{1i} A_{2i}}{A_{3i} A_{4i}} \right) \left(\frac{A_{9i}}{A_{1i}} + \frac{A_{10i}}{A_{2i}} - \frac{A_{11i}}{A_{3i}} - \frac{A_{12i}}{A_{4i}} \right) \\ -\left(A_{9i} + A_{10i} - A_{11i} - A_{12i} \right)^2 \end{array} \right\} \\ +\delta_{IR_i} \left(\log \frac{A_{2i}}{A_{4i}} \right)^{-2} \left\{ \left(\log \frac{A_{2i}}{A_{4i}} \right) \left(\frac{A_{10i}}{A_{2i}} - \frac{A_{12i}}{A_{4i}} \right) - \left(A_{10i} - A_{12i} \right)^2 \right\} \\ -\delta_{IC_i} \left(\frac{A_{13i}}{A_{5i}} + \frac{A_{14i}}{A_{6i}} \right) - \delta_{RC_i} \left(\frac{A_{15i}}{A_{7i}} \right) - 2\delta_{OE_i} \left(\frac{A_{16i}}{A_{8i}} \right) \end{array} \right]. \quad (19)$$

The observed information matrix $i(\hat{l})$ which can be obtained from the second partial derivatives of the log likelihood function evaluated at $\hat{l}$ provides us with the estimate of the variance,

$$\widehat{\text{var}}(\hat{\lambda}) = \left\{ i(\hat{\lambda}) \right\}^{-1}. \quad (20)$$

The MLE of the parameter in this paper is obtained by solving the likelihood function using Newton Raphson iterative procedure, which was implemented using `maxLik` package (Henningson & Toomet, 2011) in the R programming language.

Simulation Study

A simulation study using $N = 1000$ samples, each with sample sizes $n = 30$, 50, 100, 150, 200, 250 and 300 was conducted to examine how well the estimation procedure works for the model. The $A \sim U(0,16)$ and Y is assumed to follow the log logistic distribution (special case, $\gamma = 1$) with parameter λ. The value of −4.3 was chosen as the true parameter value of λ to simulate the survival times that mimic those seen in lung cancer data (Prentice, 1973).

DIC data mostly arise in epidemiology studies with periodic follow-ups of subjects. It is common for a subject to miss some scheduled follow up appointments. Therefore, each subject will have two sequences of time, potential inspection times and actual inspection times. Assuming all subject with the same sequence of potential inspection **PT** = (pt_1, pt_2, …, pt_g), two study period, 48 and

60 months is considered and the follow ups are scheduled to be conducted on monthly basis, therefore $g = 48$ and 60. The subject will turn up for inspection at each of the pt_j with attendance probability q where $0 \le q \le 1$ and $j = 1, 2, \ldots, g$. Therefore, each subject will have their own sequence of actual inspection times $\mathbf{AT}_i = (at_{i1}, at_{i2}, \ldots, at_{ih_i})$ where $0 \le h_i \le g$ which is simulated from the Bernoulli distribution with attendance probabilities $q = 1$, 0.8 and 0.6. It is assumed that all subjects were inspected from the beginning of the study and therefore $at_{i1} = pt_1$ and have been event free at time origin, $y = 0$.

For each subject in a sample, two random numbers u_{1i} and u_{2i} are generated from $U(0,1)$ to produce a_i and y_i where

$$a_i = u_R - \left(u_R - u_L\right)u_{1i}, \tag{21}$$

and

$$y_i = e^{-/}\left(\frac{1}{u_{2i}} - 1\right). \tag{22}$$

Then b_i is calculated from $y_i + a_i$. Following that, the intervals $(a_{L_i}, a_{R_i}]$ and $(b_{L_i}, b_{R_i}]$ are obtained for a_i and b_i respectively. The a_{L_i} will be the largest element of $\mathbf{AT}_i$ which is less than a_i, and a_{R_i} will be the smallest element of $\mathbf{AT}_i$ which is greater than a_i. Similarly, the b_{L_i} will be the largest element of $\mathbf{AT}_i$ which is less than b_i, and b_{R_i} will be the smallest element of $\mathbf{AT}_i$ which is greater than b_i. If $b_i > at_{ih_i}$, then B is RC with $(b_{L_i}, b_{R_i}] = (at_{ih_i}, \infty)$.

In order to randomly select some subjects that are OE on A or B, two time-windows are defined. The time-window for OE on A is $[G_{1i}, G_{2i}] = [a_{L_i} + (a_{R_i} - a_{L_i})u_{3i} - \varepsilon, a_{L_i} + (a_{R_i} - a_{L_i})u_{3i} + \varepsilon]$, and for OE on B is $[G_{3i}, G_{4i}] = [b_{L_i} + (b_{R_i} - b_{L_i})u_{4i} - \varepsilon, b_{L_i} + (b_{R_i} - b_{L_i})u_{4i} + \varepsilon]$ where $\varepsilon = 0.25$ and u_{3i} and u_{4i} are random numbers generated from $U(0,1)$. In cases where a_i and b_i fall in the same interval, these observations are discarded and two new values of a_i and y_i are generated to calculate b_i. This simulation procedure may yield five possible types of data where $0 < a_{L_i} < a_{R_i} \le b_{L_i} < b_{R_i} < \infty$,

1. $a_{L_i} < a_i \le a_{R_i}$ and $b_{L_i} < b_i \le a_{R_i}$ then Y is DIC;
2. $a_{L_i} < a_i \le a_{R_i}$ and $b_{L_i} < b_i < \infty$ then A is IC, B is RC;

3a. $a_{L_i} < a_i \leq a_{R_i}$ and $G_{3i} \leq b_i \leq G_{4i}$ then Y is IC;

3b. $G_{1i} \leq a_i \leq G_{2i}$ and $b_{L_i} < b_i \leq b_{R_i}$ then Y is IC;

4. $G_{1i} \leq a_i \leq G_{2i}$ and $b_{L_i} < b_i < \infty$ then Y is RC;

5. $G_{1i} \leq a_i \leq G_{2i}$ and $G_{3i} \leq b_i \leq G_{4i}$ then Y is OE.

In Table 1, the proportion of different types of data in each setting indicated.

Table 1. Average percentage of different types of data for the model at 60 and 48 months study periods.

	Study period = 60			Study period = 48		
Attendance probability	***1***	***0.8***	***0.6***	***1***	***0.8***	***0.6***
Y is DIC (%)	12.78	16.64	20.80	10.80	13.91	17.36
A is IC, B is RC (%)	33.43	38.34	43.53	36.80	42.36	48.26
Y is IC (%)	20.02	18.56	16.00	17.01	15.68	13.40
Y is RC (%)	26.02	21.33	16.59	28.75	23.63	18.38
Y is OE (%)	7.75	5.13	3.08	6.65	4.42	2.60

Simulation results

The simulation study was conducted to examine the bias, standard error (SE) and root mean square error (RMSE) of the estimate at different study periods, attendance probabilities and sample sizes.

From Table 1, more DIC data were generated at 60 months study period as compared to 48 months study period. This is due to the fact that chances of observing the event of interest either exactly or in an interval are higher for longer study period. Forty-eight months study period produced more B that is RC. Higher attendance probability produces more uncensored data and shorter width of interval for IC data.

Given in Table 2 are the bias, SE and RMSE of $\hat{l}$ at various sample sizes, n attendance probabilities, q and study periods, g. The values of bias, SE and RMSE for $\hat{l}$ decrease with an increase in n, q and g. The trend indicates that smaller censoring proportion in data, smaller sample, and shorter study period yield estimates that are less efficient and rather inaccurate.

Table 2. Bias, SE and RMSE of $\hat{\lambda}$ for the model at 60 and 48 months study period

		Study period = 60			Study period = 48		
q	***n***	**Bias**	**SE**	**RMSE**	**Bias**	**SE**	**RMSE**
	30	-0.0642	0.3633	0.3689	-0.0426	0.3921	0.3944
	50	-0.0543	0.2783	0.2836	-0.0384	0.3000	0.3024
	100	-0.0349	0.1992	0.2022	-0.0393	0.2129	0.2165
1	150	-0.0297	0.1655	0.1682	-0.0355	0.1694	0.1731
	200	-0.0286	0.1400	0.1429	-0.0280	0.1413	0.1441
	250	-0.0289	0.1248	0.1281	-0.0289	0.1293	0.1325
	300	-0.0234	0.1121	0.1145	-0.0288	0.1189	0.1223
	30	-0.0703	0.3589	0.3657	-0.0746	0.3880	0.3951
	50	-0.0587	0.2793	0.2854	-0.0542	0.2898	0.2948
	100	-0.0426	0.1918	0.1964	-0.0520	0.2165	0.2227
0.8	150	-0.0351	0.1588	0.1626	-0.0459	0.1720	0.1780
	200	-0.0461	0.1338	0.1415	-0.0431	0.1399	0.1464
	250	-0.0387	0.1179	0.1241	-0.0415	0.1254	0.1321
	300	-0.0354	0.1120	0.1175	-0.0473	0.1167	0.1259
	30	-0.0641	0.3595	0.3652	-0.0975	0.3945	0.4063
	50	-0.0607	0.2747	0.2813	-0.0780	0.2970	0.3070
	100	-0.0614	0.1961	0.2055	-0.0770	0.2057	0.2196
0.6	150	-0.0635	0.1594	0.1715	-0.0689	0.1724	0.1856
	200	-0.0634	0.1347	0.1488	-0.0708	0.1488	0.1648
	250	-0.0623	0.1223	0.1372	-0.0663	0.1273	0.1435
	300	-0.0562	0.1105	0.1240	-0.0663	0.1155	0.1332

Confidence interval estimation

The performance of three CI estimates when applied to the parameter of the proposed model is compared. The first method is based on the asymptotic normality of the MLE or Wald, followed by likelihood ratio and finally the jackknife CI estimate (see Arasan & Lunn, 2009).

Wald confidence interval estimates

Let $\hat{\lambda}$ be the MLE of parameter λ. Cox & Hinkley (1974) showed under mild regularity conditions, $\hat{\lambda}$ is asymptotically normally distributed with mean λ and variance $I(\lambda)^{-1}$ where $I(\lambda)$ is the Fisher information matrix evaluated at λ. The matrix $I(\lambda)$ can be estimated by the observed information matrix evaluated at the MLE, $i(\hat{\lambda})$. The estimate of var($\hat{\lambda}$) can be obtained from the inverse of $i(\hat{\lambda})$. If

$z_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution, then the $100(1 - \alpha)\%$ confidence interval for λ could be expressed as

$$\hat{\lambda} - z_{1-\alpha/2}\sqrt{\widehat{\mathrm{var}}\left(\hat{\lambda}\right)} < \lambda < \hat{\lambda} + z_{1-\alpha/2}\sqrt{\widehat{\mathrm{var}}\left(\hat{\lambda}\right)}. \tag{23}$$

Likelihood ratio confidence interval estimates

For a parameter of interest, λ, the likelihood ratio statistic for testing H_0: $\lambda = \lambda_0$ versus H_1: $\lambda \neq \lambda_0$ is given as

$$\psi = -2\left\{\ell\left(\lambda_0\right) - \ell\left(\hat{\lambda}\right)\right\}, \tag{24}$$

where ℓ denote the log likelihood function, λ_0 maximizes $\ell(\lambda_0)$ under H_0 or restricted model and $\hat{\lambda}$ is the MLE of λ. For large sample sizes, ψ is approximately $\chi^2_{(1,1-\alpha)}$. A $100(1 - \alpha)\%$ CI of λ is constructed by finding two values of $\hat{\lambda}$ where we fail to reject H_0 at α significance level which satisfy $\ell(\lambda_0) = \ell(\hat{\lambda}) - ½\,\chi^2_{(1,1-\alpha)}$ with $\hat{\lambda}_L < \hat{\lambda}$ and $\hat{\lambda}_R > \hat{\lambda}$.

Jackknife confidence interval estimates

The jackknife is a resampling technique where each subsample removes one observation from the original sample (Efron & Tibshirani, 1993). For a sample $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, the i^{th} jackknife sample will be $\boldsymbol{y}_{(i)} = (y_1, y_2, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n)$ for $i = 1, 2, \ldots, n$. Let $\hat{\lambda}$ be the MLE for parameter λ, then $\hat{\lambda}_{(i)}$ will be the MLE of $\hat{\lambda}$ obtained from the i^{th} jackknife sample. The jackknife estimate of the parameter λ and jackknife estimate of standard error is then calculated by using

$$\hat{\lambda}_{jack} = \hat{\lambda} - (n-1)\left\{\hat{\lambda}_{(\cdot)} - \hat{\lambda}\right\}, \tag{25}$$

$$\widehat{se}\left(\hat{\lambda}\right)_{jack} = \left[\frac{n-1}{n}\sum_{i=1}^{n}\left\{\hat{\lambda}_{(i)} - \hat{\lambda}_{(\cdot)}\right\}^2\right]^{\frac{1}{2}}, \tag{26}$$

where $\hat{\lambda}_{(\cdot)} = \sum_{i=1}^{n} \frac{\hat{\lambda}_{(i)}}{n}$.

If $t_{(1-\alpha/2,\, n-1)}$ is the $1 - \alpha/2$ quantile of the student's t distribution at $n - 1$ degrees of freedom, then the $100(1 - \alpha)\%$ jackknife confidence interval for λ could be expressed as

$$\hat{\lambda}_{jack} - t_{(1-\alpha/2,n-1)} \widehat{se}\left(\hat{\lambda}\right)_{jack} < \lambda < \hat{\lambda}_{jack} + t_{(1-\alpha/2,n-1)} \widehat{se}\left(\hat{\lambda}\right)_{jack}. \tag{27}$$

Coverage probability study

A coverage probability study was conducted using $N = 1500$ samples, each with sample sizes, $n = 30, 50, 100, 150, 200, 250$ and 300 to compare the performance of the CI estimates at different sample sizes, attendance probabilities and study periods. Other assumptions of the coverage probability study are similar to what was discussed in the simulation study.

The coverage probability error of a CI is the probability that the interval does not contains the true value of the parameter and should preferably be equal or close to the nominal error probability, α. Two nominal error probabilities were chosen as 0.05 and 0.1. The left and right error probabilities were estimated and the total error probability was calculated. Following Arasan & Lunn (2009) and Kiani & Arasan (2013), the estimated left (right) error probability was obtained by summing up the numbers for the left (right) endpoint which was more (less) than the true parameter value divided by the total number of samples, N. The estimated total error probability was calculated by summing up the number of times in which an interval did not contain the true parameter value divided by N.

The estimated error probabilities for Wald, likelihood ratio and jackknife intervals are given in Equations (28), (29) and (30) respectively as follows,

$$\begin{aligned} \text{left} &= \#\left\{\hat{\lambda} - z_{1-\alpha/2}\sqrt{\widehat{\text{var}}\left(\hat{\lambda}\right)} > \lambda\right\} / 1500, \\ \text{right} &= \#\left\{\hat{\lambda} + z_{1-\alpha/2}\sqrt{\widehat{\text{var}}\left(\hat{\lambda}\right)} > \lambda\right\} / 1500, \end{aligned} \tag{28}$$

$$\begin{aligned} \text{left} &= \#\left\{\psi > \chi^2_{(1,\alpha)} \text{ and } \hat{\lambda} > \lambda\right\} / 1500, \\ \text{right} &= \#\left\{\psi > \chi^2_{(1,\alpha)} \text{ and } \hat{\lambda} < \lambda\right\} / 1500, \end{aligned} \tag{29}$$

$$\text{left} = \#\left\{\hat{\lambda}_{jack} - t_{(1-\alpha/2,n-1)}\widehat{se}\left(\hat{\lambda}\right)_{jack} > \lambda\right\}/1500,$$
$$\text{right} = \#\left\{\hat{\lambda}_{jack} - t_{(1-\alpha/2,n-1)}\widehat{se}\left(\hat{\lambda}\right)_{jack} < \lambda\right\}/1500. \tag{30}$$

Following Doganaksoy & Schmee (1993), the interval is called anticonservative if the total error probability is more than $\alpha + 2.58\text{se}(\hat{a})$. If the total error probability is less than $\alpha - 2.58\text{se}(\hat{a})$, the interval is called conservative. The interval is called symmetric when the larger of the left or right error probability is less than 1.5 times the smaller one.

The overall performances of these CI estimates methods was evaluated based on the total numbers of anticonservative (C−), conservative (C) and asymmetrical (S−) intervals. Also, the behavior of the methods at different nominal error probabilities, sample sizes, study periods and attendance probabilities are of interest.

Coverage probability results

Summarized in Table 3 are the results obtained from the coverage probability study. Given in Tables 4 and 5 are the estimated error probabilities in detail. Figures 1 and 2 provide a graphical view of the estimated left and right error probabilities.

From Tables 4 and 5, the estimated total error probabilities of all CI estimates methods are close to the nominal error probabilities, however, most of the intervals produced are highly asymmetric, regardless of the nominal level, study period, attendance probability and sample size. Both Wald and likelihood ratio methods did not produce any conservative interval, however, the jackknife method produced some conservative intervals when sample sizes were small, $n \leq 50$. The likelihood ratio method produced more anticonservative intervals than the Wald and jackknife methods. All CI estimates methods perform poorly when $q = 0.6$. The numbers of anticonservative, conservative and asymmetrical intervals produced by all CI estimates methods are smaller at higher level of α. Also, all CI estimates methods perform slightly better at $g = 48$.

Overall, the Wald method is better than likelihood ratio and jackknife methods in constructing confidence interval for the parameter of the proposed model as it produced the least number of anticonservative and asymmetrical intervals in addition to not producing any conservative interval. From Figures 1 and 2, we can observe that all CI estimate methods work very well when $q = 1$

regardless of the nominal levels and study periods. However, they start to perform poorly when $q < 1$ especially at $q = 0.6$ by deviating far from the nominal error probability as n increases.

Table 3. Summary of the performance of Wald, likelihood ratio and jackknife methods (C− = anticonservative; C = conservative; S− = asymmetrical)

		Wald			LR			Jackknife		
	q	**C−**	**C**	**S−**	**C−**	**C**	**S−**	**C−**	**C**	**S−**
***α* = 0.05, *g* = 60**	1.0	0	0	5	1	0	7	0	1	6
	0.8	0	0	6	0	0	7	0	2	6
	0.6	2	0	6	4	0	7	3	1	6
***α* = 0.05, *g* = 48**	1.0	0	0	5	1	0	6	0	1	5
	0.8	0	0	6	0	0	7	0	2	5
	0.6	3	0	7	3	0	7	2	2	6
***α* = 0.1, *g* = 60**	1.0	0	0	5	0	0	5	0	1	6
	0.8	0	0	6	0	0	7	0	1	6
	0.6	1	0	7	3	0	7	2	1	5
***α* = 0.1, *g* = 48**	1.0	0	0	5	0	0	5	0	1	5
	0.8	0	0	5	0	0	7	0	2	5
	0.6	3	0	7	3	0	7	3	0	7

Table 4. Estimated error probabilities of Wald, likelihood ratio and jackknife methods for the model when α = 0.05 (C− = anticonservative; C = conservative)

		Wald			**Likelihood Ratio**			**Jackknife**		
	n	**Left Error**	**Right Error**	**Total Error**	**Left Error**	**Right Error**	**Total Error**	**Left Error**	**Right Error**	**Total Error**
***q* = 1, *g* = 60**	30	0.0193	0.0220	0.0413	0.0167	0.0300	0.0467	0.0187	0.0053	0.0240^{C}
	50	0.0247	0.0333	0.0580	0.0227	0.0427	0.0653^{C-}	0.0253	0.0200	0.0453
	100	0.0167	0.0307	0.0473	0.0153	0.0360	0.0513	0.0173	0.0267	0.0440
	150	0.0180	0.0353	0.0533	0.0167	0.0393	0.0560	0.0193	0.0313	0.0507
	200	0.0167	0.0360	0.0527	0.0160	0.0380	0.0540	0.0193	0.0353	0.0547
	250	0.0160	0.0340	0.0500	0.0153	0.0353	0.0507	0.0173	0.0333	0.0507
	300	0.0133	0.0313	0.0447	0.0127	0.0320	0.0447	0.0140	0.0280	0.0420
***q* = 0.8, *g* = 60**	30	0.0167	0.0227	0.0393	0.0153	0.0307	0.0460	0.0173	0.0080	0.0253^{C}
	50	0.0147	0.0360	0.0507	0.0133	0.0433	0.0567	0.0133	0.0213	0.0347^{C}
	100	0.0127	0.0287	0.0413	0.0113	0.0327	0.0440	0.0167	0.0253	0.0420
	150	0.0160	0.0287	0.0447	0.0153	0.0340	0.0493	0.0180	0.0253	0.0433
	200	0.0127	0.0367	0.0493	0.0120	0.0413	0.0533	0.0107	0.0380	0.0487
	250	0.0127	0.0300	0.0427	0.0120	0.0333	0.0453	0.0120	0.0293	0.0413
	300	0.0060	0.0467	0.0527	0.0060	0.0487	0.0547	0.0067	0.0440	0.0507
***q* = 0.6, *g* = 60**	30	0.0180	0.0193	0.0373	0.0153	0.0333	0.0487	0.0193	0.0093	0.0287^{C}
	50	0.0160	0.0253	0.0413	0.0160	0.0313	0.0473	0.0200	0.0160	0.0360
	100	0.0160	0.0440	0.0600	0.0147	0.0507	0.0653^{C-}	0.0160	0.0387	0.0547
	150	0.0113	0.0460	0.0573	0.0100	0.0493	0.0593	0.0133	0.0447	0.0580
	200	0.0080	0.0560	0.0640	0.0073	0.0607	0.0680^{C-}	0.0087	0.0527	0.0613^{C-}
	250	0.0073	0.0660	0.0733^{C-}	0.0067	0.0700	0.0767^{C-}	0.0067	0.0627	0.0693^{C-}
	300	0.0060	0.0593	0.0653^{C-}	0.0060	0.0653	0.0713^{C-}	0.0067	0.0593	0.0660^{C-}
***q* = 1, *g* = 48**	30	0.0253	0.0180	0.0433	0.0213	0.0293	0.0507	0.0207	0.0007	0.0213^{C}
	50	0.0240	0.0287	0.0527	0.0193	0.0340	0.0533	0.0227	0.0167	0.0393
	100	0.0247	0.0380	0.0627	0.0233	0.0420	0.0653^{C-}	0.0260	0.0280	0.0540
	150	0.0133	0.0347	0.0480	0.0133	0.0373	0.0507	0.0147	0.0333	0.0480
	200	0.0140	0.0300	0.0440	0.0127	0.0333	0.0460	0.0153	0.0267	0.0420
	250	0.0147	0.0413	0.0560	0.0140	0.0433	0.0573	0.0153	0.0360	0.0513
	300	0.0120	0.0373	0.0493	0.0107	0.0420	0.0527	0.0127	0.0373	0.0500
***q* = 0.8, *g* = 48**	30	0.0207	0.0160	0.0367	0.0193	0.0300	0.0493	0.0200	0.0007	0.0207^{C}
	50	0.0133	0.0287	0.0420	0.0113	0.0340	0.0453	0.0147	0.0153	0.0300^{C}
	100	0.0207	0.0367	0.0573	0.0187	0.0433	0.0620	0.0227	0.0287	0.0513
	150	0.0140	0.0387	0.0527	0.0127	0.0453	0.0580	0.0173	0.0373	0.0547
	200	0.0067	0.0360	0.0427	0.0047	0.0393	0.0440	0.0120	0.0327	0.0447
	250	0.0100	0.0407	0.0507	0.0100	0.0440	0.0540	0.0107	0.0367	0.0473
	300	0.0100	0.0440	0.0540	0.0100	0.0473	0.0573	0.0107	0.0433	0.0540
***q* = 0.6, *g* = 48**	30	0.0120	0.0267	0.0387	0.0120	0.0460	0.0580	0.0173	0.0013	0.0187^{C}
	50	0.0120	0.0347	0.0467	0.0093	0.0460	0.0553	0.0147	0.0160	0.0307^{C}
	100	0.0147	0.0367	0.0513	0.0120	0.0433	0.0553	0.0180	0.0307	0.0487
	150	0.0087	0.0493	0.0580	0.0073	0.0560	0.0633	0.0107	0.0447	0.0553
	200	0.0073	0.0593	0.0667^{C-}	0.0060	0.0640	0.0700^{C-}	0.0073	0.0547	0.0620
	250	0.0067	0.0633	0.0700^{C-}	0.0060	0.0687	0.0747^{C-}	0.0067	0.0620	0.0687^{C-}
	300	0.0080	0.0660	0.0740^{C-}	0.0080	0.0740	0.0820^{C-}	0.0087	0.0673	0.0760^{C-}

Table 5. Estimated error probabilities of Wald, likelihood ratio and jackknife methods for the model when α = 0.1 (C− = anticonservative; C = conservative)

		Wald			**Likelihood Ratio**			**Jackknife**		
	n	**Left Error**	**Right Error**	**Total Error**	**Left Error**	**Right Error**	**Total Error**	**Left Error**	**Right Error**	**Total Error**
***q* = 1, *g* = 60**	30	0.0427	0.0493	0.0920	0.0400	0.0593	0.0993	0.0473	0.0280	0.0753^{C}
	50	0.0433	0.0567	0.1000	0.0427	0.0640	0.1067	0.0507	0.0493	0.1000
	100	0.0333	0.0653	0.0987	0.0327	0.0700	0.1027	0.0373	0.0580	0.0953
	150	0.0327	0.0667	0.0993	0.0320	0.0687	0.1007	0.0353	0.0647	0.1000
	200	0.0387	0.0707	0.1093	0.0360	0.0720	0.1080	0.0427	0.0653	0.1080
	250	0.0333	0.0640	0.0973	0.0327	0.0687	0.1013	0.0347	0.0633	0.0980
	300	0.0327	0.0727	0.1053	0.0313	0.0760	0.1073	0.0353	0.0693	0.1047
***q* = 0.8, *g* = 60**	30	0.0407	0.0500	0.0907	0.0387	0.0587	0.0973	0.0413	0.0267	0.0680^{C}
	50	0.0347	0.0613	0.0960	0.0287	0.0680	0.0967	0.0440	0.0547	0.0987
	100	0.0307	0.0633	0.0940	0.0293	0.0693	0.0987	0.0353	0.0560	0.0913
	150	0.0253	0.0680	0.0933	0.0240	0.0727	0.0967	0.0287	0.0633	0.0920
	200	0.0273	0.0793	0.1067	0.0253	0.0827	0.1080	0.0293	0.0740	0.1033
	250	0.0240	0.0707	0.0947	0.0233	0.0753	0.0987	0.0280	0.0687	0.0967
	300	0.0220	0.0833	0.1053	0.0220	0.0880	0.1100	0.0233	0.0833	0.1067
***q* = 0.6, *g* = 60**	30	0.0360	0.0540	0.0900	0.0347	0.0660	0.1007	0.0353	0.0287	0.0640^{C}
	50	0.0353	0.0613	0.0967	0.0347	0.0660	0.1007	0.0413	0.0440	0.0853
	100	0.0273	0.0787	0.1060	0.0267	0.0873	0.1140	0.0327	0.0733	0.1060
	150	0.0247	0.0867	0.1113	0.0240	0.0920	0.1160	0.0267	0.0807	0.1073
	200	0.0187	0.1033	0.1220	0.0173	0.1067	0.1240^{C-}	0.0193	0.1020	0.1213^{C-}
	250	0.0133	0.1053	0.1187	0.0120	0.1080	0.1200^{C-}	0.0133	0.1033	0.1167
	300	0.0133	0.1133	0.1267^{C-}	0.0127	0.1227	0.1353^{C-}	0.0167	0.1167	0.1333C-
***q* = 1, *g* = 48**	30	0.0433	0.0440	0.0873	0.0393	0.0553	0.0947	0.0427	0.0160	0.0587^{C}
	50	0.0453	0.0507	0.0960	0.0440	0.0560	0.1000	0.0500	0.0360	0.0860
	100	0.0413	0.0740	0.1153	0.0380	0.0807	0.1187	0.0433	0.0607	0.1040
	150	0.0313	0.0700	0.1013	0.0313	0.0753	0.1067	0.0340	0.0600	0.0940
	200	0.0293	0.0600	0.0893	0.0267	0.0647	0.0913	0.0307	0.0567	0.0873
	250	0.0320	0.0767	0.1087	0.0300	0.0827	0.1127	0.0353	0.0653	0.1007
	300	0.0273	0.0707	0.0980	0.0267	0.0693	0.0960	0.0287	0.0680	0.0967
***q* = 0.8, *g* = 48**	30	0.0413	0.0467	0.0880	0.0387	0.0613	0.1000	0.0433	0.0127	0.0560^{C}
	50	0.0360	0.0513	0.0873	0.0320	0.0640	0.0960	0.0380	0.0347	0.0727^{C}
	100	0.0373	0.0653	0.1027	0.0367	0.0740	0.1107	0.0407	0.0560	0.0967
	150	0.0280	0.0740	0.1020	0.0273	0.0827	0.1100	0.0313	0.0680	0.0993
	200	0.0247	0.0780	0.1027	0.0220	0.0873	0.1093	0.0253	0.0687	0.0940
	250	0.0227	0.0767	0.0993	0.0213	0.0807	0.1020	0.0240	0.0753	0.0993
	300	0.0227	0.0840	0.1067	0.0220	0.0873	0.1093	0.0260	0.0807	0.1067
***q* = 0.6, *g* = 48**	30	0.0293	0.0640	0.0933	0.0267	0.0753	0.1020	0.0353	0.0307	0.0660^{C}
	50	0.0253	0.0673	0.0927	0.0233	0.0787	0.1020	0.0293	0.0513	0.0807
	100	0.0307	0.0827	0.1133	0.0273	0.0880	0.1153	0.0327	0.0653	0.0980
	150	0.0207	0.0913	0.1120	0.0193	0.0987	0.1180	0.0207	0.0833	0.1040
	200	0.0207	0.1093	0.1300^{C-}	0.0180	0.1153	0.1333^{C-}	0.0267	0.1047	0.1313^{C-}
	250	0.0173	0.1127	0.1300^{C-}	0.0160	0.1227	0.1387^{C-}	0.0187	0.1087	0.1273^{C-}
	300	0.0153	0.1220	0.1373^{C-}	0.0147	0.1273	0.1420^{C-}	0.0153	0.1167	0.1320^{C-}

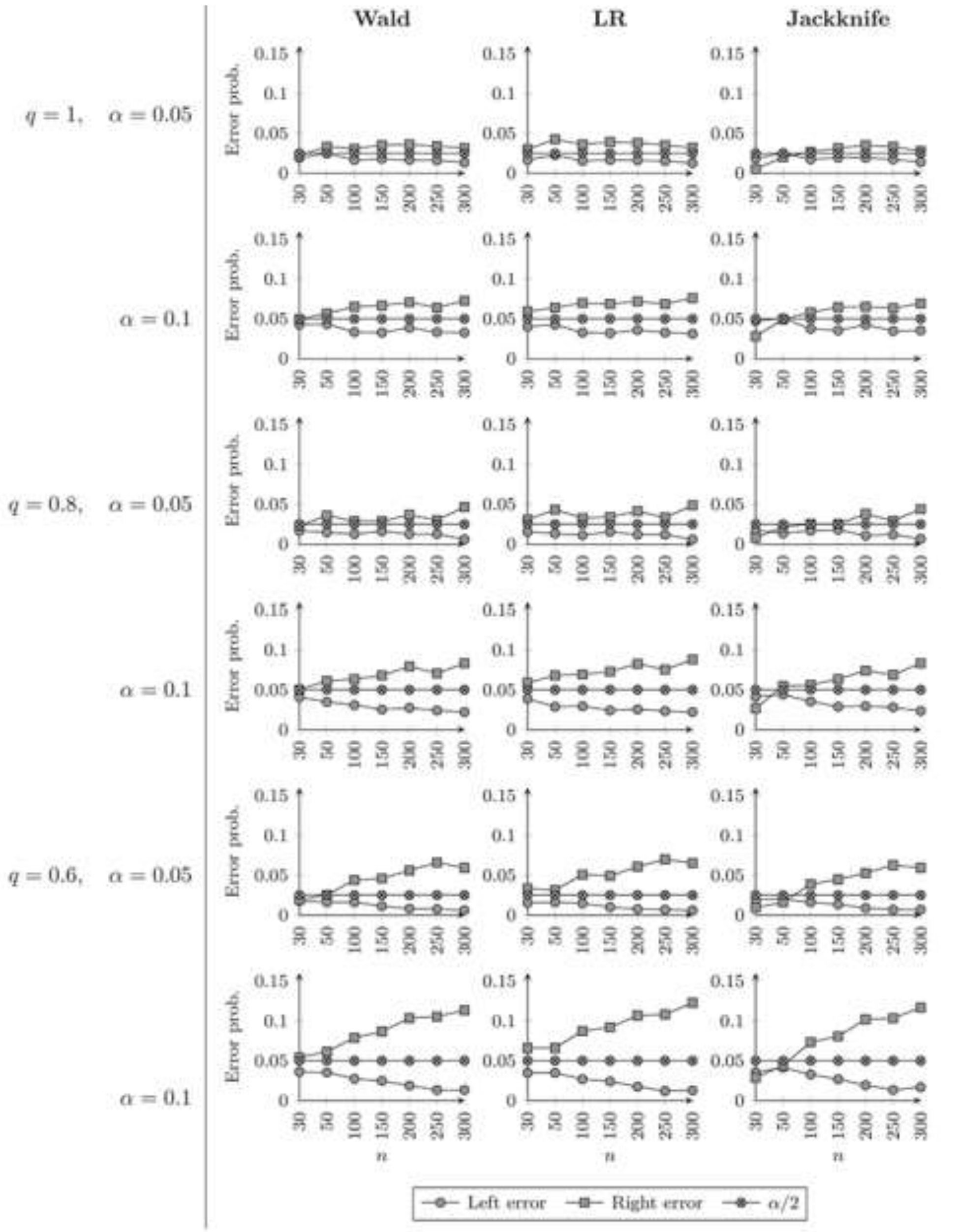

Figure 1. Estimated error probabilities of interval estimates methods when g = 60

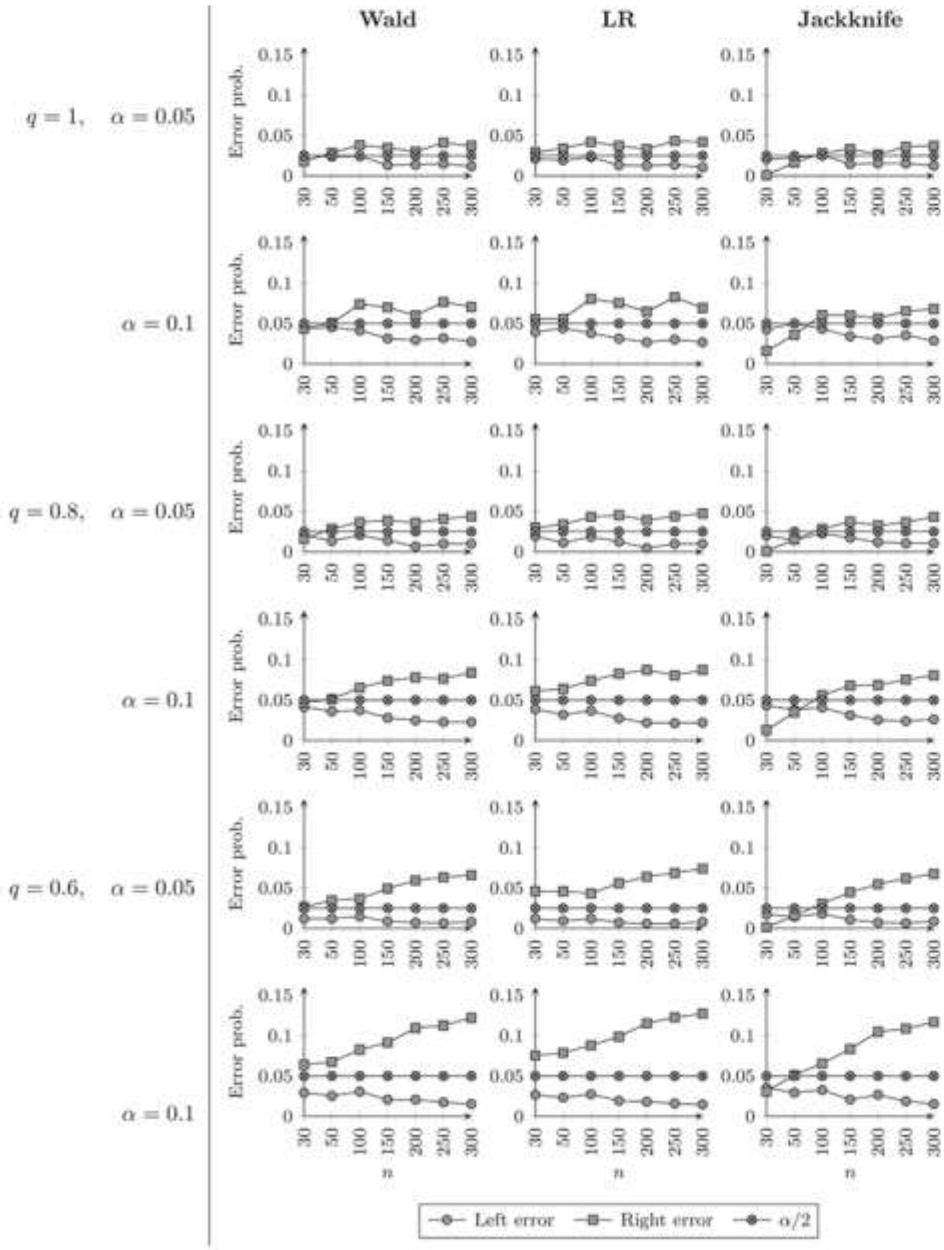

Figure 2. Estimated error probabilities of interval estimates methods when $g = 48$

Conclusion

The estimation procedure worked well for the log logistic distribution with doubly interval censored data where values of bias, standard error and root mean square error are all reasonably low. The Wald confidence interval estimates performed better than the likelihood ratio and jackknife confidence interval when dealing with doubly interval censored data. The jackknife method required more computational effort than the other two. The finite-difference gradient and Hessian which are included in the `maxLik` package in R programming language could not be applied as the derivatives become unreliable due to the complexity of the model.

Acknowledgements

We gratefully acknowledge financial support from the Ministry of Education Malaysia. The research leading to these results has received funding from the Fundamental Research Grant Scheme (FRGS 2014) under vote no. 5524673.

References

Arasan, J., & Lunn, M. (2009). Survival model of a parallel system with dependent failures and time varying covariates. *Journal of Statistical Planning and Inference, 139*(3), 944−951. doi: 10.1016/j.jspi.2008.06.007

Bacchetti, P. (1990). Estimating the incubation period of AIDS by comparing population infection and diagnosis patterns. *Journal of the American Statistical Association, 85*(412), 1002−1008. 10.2307/2289596

Bacchetti, P., & Jewell, N. P. (1991). Nonparametric estimation of the incubation period of AIDS based on a prevalent cohort with unknown infection times. *Biometrics*, *47*(3), 947−960. doi: 10.2307/2532651

Brookmeyer, R., & Goedart, J. J. (1989). Censoring in an epidemic with an application to hemophilia-associated AIDS. *Biometrics, 45*(1), 325−335. doi: 10.2307/2532057

Cox, D. R., & Hinkley, D. V. (1974). *Theoretical statistics*. CRC Press. doi: 10.1007/978-1-4899-2887-0

Darby, S. C., Doll, R., Thakrar, B., Rizza, C. R., & Cox, D. R. (1990). Time from infection with HIV to onset of AIDS in patients with haemophilia in the UK. *Statistics in Medicine, 9*(6), 681−689. doi: 10.1002/sim.4780090615

De Gruttola, V., & Lagakos, S. W. (1989). Analysis of doubly-censored survival data, with application to AIDS. *Biometrics, 45*(1), 1−11. doi: 10.2307/2532030

Doganaksoy, N., & Schmee, J. (1993). Comparisons of approximate confidence intervals for distributions used in life-data analysis. *Technometrics, 35*(2), 175−184. doi: 10.2307/1269662

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap (Monographs on Statistics and Applied Probability, no. 57)*. London: Chapman and Hall.

Fang, H. B., & Sun, J. (2001). Consistency of nonparametric maximum likelihood estimation of a distribution function based on doubly interval-censored failure time data. *Statistics & Probability Letters*, *55*(3), 311−318. doi: 10.1016/s0167-7152(01)00160-2

Goggins, W. B., Finkelstein, D. M., & Zaslavsky, A. M. (1999). Applying the Cox proportional hazards model when the change time of a binary time-varying covariate is interval censored. *Biometrics, 55*(2), 445−451. doi: 10.1111/j.0006-341x.1999.00445.x

Gómez, G., & Calle, M. L. (1999). Non-parametric estimation with doubly censored data. *Journal of Applied Statistics, 26*(1), 45−58. doi: 10.1080/02664769922647

Gómez, G., & Lagakos, S. W. (1994). Estimation of the infection time and latency distribution of AIDS with doubly censored data. *Biometrics, 50*(1), 204−212. doi: 10.2307/2533210

Henningson, A., & Toomet, O. (2011). maxLik: A package for maximum likelihood estimation in R. *Computational Statistics*, *26*(3), 443−458. doi: 10.1007/s00180-010-0217-1

Jara, A., Lesaffre, E., De Iorio, M., & Quintana, F. (2010). Bayesian semiparametric inference for multivariate doubly-interval-censored data. *The Annals of Applied Statistics, 4*(4), 2126−2149. doi: 10.1214/10-aoas368

Jewell, N. P. (1994). Non-parametric estimation and doubly-censored data: General ideas and applications to AIDS. *Statistics in Medicine*, *13*(19-20), 2081−2095. doi: 10.1002/sim.4780131917

Jewell, N. P., Malani, H. M., & Vittinghoff, E. (1994). Nonparametric estimation for a form of doubly censored data, with application to two problems in AIDS. *Journal of the American Statistical Association, 89*(425), 7−18. doi: 10.2307/2291196

Kalbfleisch, J. D., & Prentice, R. L. (2002). *The statistical analysis of failure time data*. (2nd Ed.). NY: John Wiley & Sons. doi: 10.1002/9781118032985

Kiani, K., & Arasan, J. (2012). Interval estimations for parameters of exponential model with doubly interval-censored survival time data. *Proceedings of 2nd Regional Conference on Applied and Engineering Mathematics (RCAEM-II) 2012*, pp. 653−660. Arau, Malaysia: Universiti Malaysia Perlis.

Kiani, K., & Arasan, J. (2013). Gompertz model with time-dependent covariate in the presence of interval-, right-and left-censored data. *Journal of Statistical Computation and Simulation, 83*(8), 1472−1490. doi: 10.1080/00949655.2012.662979

Kim, M. Y., De Gruttola, V. G., & Lagakos, S. W. (1993). Analyzing doubly censored data with covariates, with application to AIDS. *Biometrics, 49*(1), 13−22. doi: 10.2307/2532598

Komárek, A., & Lesaffre, E. (2006). Bayesian semi-parametric accelerated failure time model for paired doubly interval-censored data. *Statistical Modelling, 6(*1), 3−22. doi: 10.1191/1471082x06st107oa

Komárek, A., & Lesaffre, E. (2008). Bayesian accelerated failure time model with multivariate doubly interval-censored data and flexible distributional assumptions. *Journal of the American Statistical Association*, *103*(482), 523-533. doi: 10.1198/016214507000000563

Komárek, A., Lesaffre, E., Härkänen, Tomni, Declerck, D., & Virtanen, J. I. (2005). A Bayesian analysis of multivariate doubly-interval-censored dental data. *Biostatistics, 6*(1), 145−155. doi: 10.1093/biostatistics/kxh023

Lim, H., Sun, J., & Matthews, D. E. (2002). Maximum likelihood estimation of a survival function with a change point for truncated and interval-censored data. *Statistics in Medicine, 21*(5), 743−752. doi: 10.1002/sim.986

McBryde, E. S., Gibson, G., Pettitt, A. N., Zhang, Y., Zhao, B., & McElwain, D. L. S. (2006). Bayesian modelling of an epidemic of severe acute respiratory syndrome. *Bulletin of Mathematical Biology, 68*(4), 889−917. doi: 10.1007/s11538-005-9005-4

Oller, R., Gómez, G., & Calle, M. L. (2004). Interval censoring: model characterizations for the validity of the simplified likelihood. *Canadian Journal of Statistics, 32*(3), 315−326. doi: 10.2307/3315932

Pan, W. (2001). A multiple imputation approach to regression analysis for doubly censored data with application to AIDS studies. *Biometrics, 57*(4), 1245–1250. doi: 10.1111/j.0006-341x.2001.01245.x

Prentice, R. L. (1973). Exponential survivals with censoring and explanatory variables. *Biometrika*, *60*(2), 279–288. doi: 10.1093/biomet/60.2.279

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reich, N. G., Lessler, J., Cummings, D. A. T., & Brookmeyer, R. (2009). Estimating incubation period distributions with coarse data. *Statistics in Medicine, 28*(22), 2769–2784. doi: 10.1002/sim.3659

Sun, J. (1995). Empirical estimation of a distribution function with truncated and doubly interval-censored data and its application to AIDS studies. *Biometrics, 51*(3), 1096–1104. doi: 10.2307/2533008

Sun, J. (1997). Self-consistency estimation of distributions based on truncated and doubly censored survival data with applications to AIDS cohort studies. *Lifetime Data Analysis, 3*(4), 305–313. doi: 10.1023/A:1009609227969

Sun, J. (1998). Interval censoring. In *Encyclopedia of biostatistics*. (pp. 2090–2095). New York : John Wiley. doi: 10.1002/0470011815.b2a11039

Sun, J. (2004). Statistical analysis of doubly interval-censored failure time data. *Handbook of Statistics: Advances in Survival Analysis, 23*, 105–122. doi: 10.1016/s0169-7161(03)23006-6

Sun, J., Liao, Q., & Pagano, M. (1999). Regression analysis of doubly censored failure time data with applications to AIDS studies. *Biometrics, 55*(3), 909–914. doi: 10.1111/j.0006-341x.1999.00909.x

Tu, X. M. (1995). Nonparametric estimation of survival distributions with censored initiating time, and censored and truncated terminating time: Application to transfusion data for acquired immune deficiency syndrome. *Journal of the Royal Statistical Society, Series C (Applied Statistics), 44*(1), 3–16. doi: 10.2307/2986191