Spring 2016

Small sample confidence bands for the survival functions under proportional hazards model

Emad Mohamed Abdurasul

Department: Mathematics and Statistics

SMALL SAMPLE CONFIDENCE BANDS FOR THE SURVIVAL FUNCTIONS UNDER PROPORTIONAL HAZARDS MODEL

by

EMAD MOHAMED ABDURASUL

A DISSERTATION

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

2016

Approved by:

Dr. Robert Paige, Advisor
Dr. V.A. Samaranayake
Dr. Xuerong Wen
Dr. Gayla Olbricht
Dr. Xiaoping Du

ABSTRACT

In this work, a saddlepoint-based method is developed for generating small sample confidence bands for the population survival function from the Kaplan-Meier (KM), the product limit (PL), and Abdushukurov-Cheng-Lin (ACL) survival function estimators, under the proportional hazards model. In the process the exact distribution of these estimators is derived and developed mid-population tolerance bands for said estimators. The proposed saddlepoint method depends upon the Mellin transform of the zero-truncated survival estimator which is derived for the KM, PL, and ACL estimators. These transforms are inverted via saddlepoint approximations to yield highly accurate approximations to the cumulative distribution functions of the respective cumulative hazard function estimators and these distribution functions are then inverted to produce saddlepoint confidence bands. The saddlepoint confidence bands for the KM, PL and ACL estimators is compared with those obtained from competing large sample methods as well as those obtained from the exact distribution. In the simulation studies it is found that the saddlepoint confidence bands are very close to the confidence bands derived from the exact distribution, while being much easier to compute, and outperform the competing large sample methods in terms of coverage probability.

ACKNOWLEDGMENTS

First of all , Alhamdulillah, all praises belong to Allah who has given me the health and the strength to finish this thesis. Next, I am really grateful to my advisor, Professor Robert Paige, for his support, guidance, and patience throughout my research. I also benefited greatly from his open-mindedness and his passionate encouragement for my research. I would also like to thank my committee members Dr. V.A. Samaranayake, Dr. Xuerong Wen, Dr. Gayla Olbricht and Dr. Xiaoping Du for guiding my research during the past several years and helping me to develop my background in mathematics and statistics. Furthermore, I appreciate all the help and advice given to me by the faculty members of the Department of Mathematics and Statistics, Missouri University of Science and Technology. Without Dr. Paige's guidance, help from committee members and friends, and support from my family and wife, I would never have been able to finish my dissertation

I would also like to thank my parents, brothers, sisters and all my family and all my friends. They always supported and encouraged me with their best wishes.

Finally, I certainly owe special thanks and gratitude to my wife and my children; Abdulrahim, Wijdan, Wisal, Waad, and Abdulkarim. They are always there to cheer me up and stand by me through the good and bad times.

DEDICATION

I would like to dedicate this Doctoral dissertation to my parents, my wife and my children. Without their continued support and counsel, I could not have completed this process.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. INTRODUCTION

We develop methods for making small sample inference about the survival function, in the presence of right censoring, and under the proportional hazards model. We let $X_1, X_2, ..., X_n$ denote the independent and identically distributed (i.i.d.) survival times with continuous cumulative distribution function (CDF) $F(t)$ and survival function $S(t) = 1 - F(t)$. These survival times are censored at the right by i.i.d. continuous random variables $Y_{1,} Y_{2,} ..., Y_{n,}$ which are independent of the survival times with continuous CDF $F_Y(t)$ and survival function $S_Y(t) = 1 - F_Y(t)$. The right-censored data are denoted as $(Z_1, \Delta_1), (Z_2, \Delta_2), \ldots, (Z_n, \Delta_n)$ where the time on study is $Z_i = \min\{X_i, Y_i\}$ and the survival indicator function is $\Delta_i = I(X_i \leqslant Y_i)$. The observed right-censored data are denoted as $(z_1, \delta_1), (z_2, \delta_2), ..., (z_n, \delta_n)$. The pair (X, Y) follows a proportional hazards or Koziol-Green model if there exists a real number $\beta > 0$ such that

$$S_Y(t) = S^{\beta}(t).$$

An equivalent characterization of this model in terms of cumulative hazard functions is

$$H_Y(t) = -\ln[S_Y(t)] = \beta(-\ln[S(t)]) = \beta H(t)$$

which are proportional to one another. One well-known consequence of the proportional hazards model is that Z_i and Δ_i are independent for $i = 1, \ldots, n$. Consider for instance censoring times which are Weibull with survival function

$$S_Y(t) = e^{-(\lambda t)^k}$$

then

$$H_Y(t) = -\ln\left[S_Y(t)\right] = (\lambda t)^k = \frac{\beta}{\beta}(\lambda t)^k$$
$$= \beta\left(\frac{\lambda t}{\beta^{\frac{1}{k}}}\right)^k = \beta\left(-\ln\left[S(t)\right]\right) = \beta H(t)$$

where

$$S(t) = e^{-\left(\frac{\lambda t}{\beta^{\frac{1}{k}}}\right)^k} = e^{-(\lambda t)^k/\beta}$$

and so

$$S_Y(t) = S^{\beta}(t).$$

Kaplan and Meier (1958) define an estimator for $S(t)$ as

$$\hat{S}(t) = \prod_{i=1}^{n\hat{F}_Z(t)} c_{in}^{\delta_{(i)}} \quad \text{for} \quad t \leq z_{(n)}$$

where $z_{(n)}$ denotes the largest observed time on study, $\hat{F}_Z(t)$ is the empirical CDF for the times on study;

$$\hat{F}_Z(t) = \frac{1}{n}\sum_{i=1}^{n} I(z_i \leq t),$$

where $I(z_i \leq t)$ is the indicator function of event $\{z_i \leq t\}$, survival indicators

$$\delta_{(1)}, \delta_{(2)}, \ldots, \delta_{(n)},$$

are associated with the complete set of ordered times on study which are

$$z_{(1)} < z_{(2)} < \cdots < z_{(n)}$$

and the c_{in} weights are defined as

$$c_{in} = \frac{n-i}{n-i+1} = 1 - \frac{1}{n-i+1}.$$

Estimator $\hat{S}(t)$ was not derived under the proportional hazards model, per se, and in fact is the nonparametric maximum likelihood estimate of $S(t)$ where one maximizes over the class of all piecewise constant survival curves with break points at the non-censored times on study, as shown in Kaplan and Meier (1958). One issue with estimator $\hat{S}(t)$ is that it is used only when $t \leq z_{(n)}$ and the reason for this is that estimator $\hat{S}(t)$ is undefined for $t > z_{(n)}$. A number of tail completion methods have been proposed to provide the estimator with a reasonable definition for $t > z_{(n)}$. The Product Limit (PL) or Gill estimator is simply estimator $\hat{S}(t)$ used for all $t > 0$ so that

$$\hat{S}_{PL}(t) = \prod_{i=1}^{n\hat{F}_Z(t)} c_{in}^{\delta_{(i)}} \quad \text{for} \quad t \geq 0$$

When largest time on study is censored, meaning that $\delta_{(n)} = 0$, then the PL estimator has an infinite non-zero and constant tail for $t > z_{(n)}$ where the constant is

$$\prod_{i=1}^{n-1} c_{in}^{\delta_{(i)}}.$$

Furthermore, when $\delta_{(n)} = 1$ then $\hat{S}_{PL}(t) = 0$ for $t > z_{(n)}$. In small samples the PL estimator is shown to have smaller bias than estimators with tails which always decrease to zero; see for instance Gill (1980) and Klein (1991). Note however that for large samples estimators $\hat{S}_{PL}(t)$ and $\hat{S}(t)$ are essentially equivalent. Moeschberger and Klein (1985) present a number of tail completion methods for estimator $\hat{S}(t)$ which do not result in infinite and constant tails. The methods they present fall into three general categories. First, are the expected order statistic (EOS) methods in which censored observations exceeding the largest observed failure time are replaced by their expectations under a fitted Weibull model. Next, they consider a class of methods which estimate the tail with a Weibull distribution fitted by least squares. The last class of methods involves estimating the tail with a fitted exponential distribution. Perhaps the most common tail completion method for the $\hat{S}(t)$ estimator is due to Efron (1967). Here, the $\hat{S}(t)$ estimator is set equal to zero for $t > z_{(n)}$. Its form is then given as

$$\hat{S}_{KM}(t) = \left[\prod_{i=1}^{n\hat{F}_Z(t)} c_{in}^{\delta_{(i)}} \right] I\left(t \leq z_{(n)}\right).$$

We have chosen to simply refer to Efron's tail-completed $\hat{S}(t)$ estimator as the Kaplan-Meier (KM) estimator, and denote it as $\hat{S}_{KM}(t)$ from this point forward, since $\hat{S}_{KM}(t)$ was originally proposed by Kaplan and Meier (1958) and is perhaps the most commonly used estimator of $S(t)$ with right censored data. In addition, $\hat{S}_{KM}(t)$ satisfies the self-consistency equations from Efron (1967) and is simple to implement in practice. Most studies considering distributional results for the KM estimator assume a large sample. Breslow and Crowley (1974) establish the consistency of the KM estimator and weak convergence of

$$\sqrt{n}\left[\hat{S}_{KM}(t) - S(t)\right]$$

to a mean zero Gaussian process, under fairly general conditions. Identical results hold for the PL estimator. The few studies which consider small sample settings, under the proportional hazards model, include Chang (1996) and Chen, et al. (1982). In the former, Chang (1996) shows that the exact distribution of the KM estimator under the proportional hazards model is a weighted average of permutation distributions. He however notes that exact computations with his expression are quite involved and only feasible for very small samples. Chen, et al. (1982) obtains an exact expression for the vth moment $(v > 0)$ of the KM estimator under proportional hazards and use this expression to study the bias of the KM estimator, and compare the exact variance of the KM estimator with its asymptotic variance. Abdushukurov (1984), Hollander et al. (1985), and Cheng and Lin (1987) independently proposed another estimator of survival function $S(t)$ under the proportional hazards assumption. This ACL (Abduskhurov, Cheng and Lin) survival function estimator is given as

$$\hat{S}_{ACL}(t) = \left[\hat{S}_Z(t)\right]^{\bar{\delta}}$$

where

$$\hat{S}_Z(t) = \frac{1}{n}\sum_{i=1}^{n} I(z_i > t)$$

is the empirical survival function and

$$\bar{\delta} = \frac{1}{n}\sum_{i=1}^{n} \delta_{(i)}.$$

This estimator is motivated by the fact that under the proportional hazards model

$$S_Y(t) = [S(t)]^{\beta}$$

and since $Z = \min(X, Y)$ then

$$\begin{aligned} S_Z(t) &= S(t)S_Y(t) \\ &= [S(t)]^{\beta+1} \end{aligned}$$

and so

$$S(t) = [S_Z(t)]^{\frac{1}{\beta+1}} = [S_Z(t)]^{\gamma}$$

where

$$\gamma = P(X \leq Y).$$

The ACL estimator is asymptotically more efficient that the KM and PL estimators under the proportional hazards model, as shown in Cheng and Lin (1987). They also establish, under fairly general conditions, the consistency of the ACL estimator and weak convergence of

$$\sqrt{n}\left[\hat{S}_{ACL}(t) - S(t)\right]$$

to a mean zero Gaussian process. As described above there are few studies which consider the performance of survival function estimators under the proportional hazards model in small sample settings. In this study we derive the exact distributions of the KM, PL and ACL estimators and propose novel small sample confidence intervals (CIs) for $S(t)$, for fixed t, based on saddlepoint approximations which are generated from the Mellin transform of the zero-truncated survival estimator in question. We form pointwise confidence bands from these confidence intervals and find that they outperform the classical large sample methods in terms of coverage probability. We also find that our saddlepoint CDF approximations are quite close to the exact CDFs. The remainder of this dissertation is organized as follows. The exact distributions of the KM, PL and ACL estimators, under the proportional hazards model, are derived in Section 2. The associated Mellin transforms for the zero-truncated KM, PL and ACL estimators, that provide access to saddlepoint approximations for the three estimators, are derived in Section 3. Pointwise confidence bands for $S(t)$ from the KM, PL and ACL estimators by way three methods; (i) exact distribution, (ii) saddlepoint CDF approximation, and (iii) classical large sample methods are presented in Section 4. Simulation studies comparing the performance of the various confidence bands are presented in Section 5. The exact values of the mean, bias, variance and mean squared error (MSE) of the KM, PL and ACL estimators are presented in Section 6. Finally, concluding remarks are made in Section 7.

2. EXACT DISTRIBUTIONS OF THE SURVIVAL ESTIMATORS

In this section, we derive the exact distributions of the KM, PL and ACL estimators under the proportional hazards model.

2.1. EXACT DISTRIBUTION OF THE KM ESTIMATOR

Chen, et al. (1982) obtained an exact expression for the vth moment of the KM estimator, under the proportional hazards model, for $v \geq 0$ as

$$E\left[\hat{S}_{KM}^{v}(t)\right] = \sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) \prod_{i=1}^{r} \left[\gamma c_{in}^{v} + (1-\gamma)\right] \tag{2.1}$$

where here, and in what follows,

$$b\left(r, F_Z(t)\right) = \binom{n}{r} \left[F_Z(t)\right]^r \left[S_Z(t)\right]^{n-r}.$$

Note also that this expression which was used in Chen, et al. (1982) to compute positive moments for the KM estimator is in fact valid for all $-\infty < v < \infty$. This is seen by noting that since

$$0 < c_{in}$$

for $i = 1, \ldots, n-1$ then

$$E\left[\hat{S}_{KM}^{v}(t)\right] < \infty$$

for any $-\infty < v < \infty$.

To derive the exact distribution of the KM estimator we take $v = 1$ and consider various the terms in the resulting expression for $E\left[\hat{S}_{KM}(t)\right]$. Note that this will work for all possible values of $\hat{S}_{KM}(t)$ except zero. This probability, however, is easy to compute directly by noting that

$$P\left(\hat{S}_{KM}(t)=0\right) = P\left(Z_1 \leq t, \ldots, Z_n \leq t\right)$$
$$= \left[F_Z(t)\right]^n.$$

For $r = 0$ we have

$$b\left(0, F_Z(t)\right) \prod_{i=1}^{0} \left[\gamma c_{in} + (1-\gamma)\right] = \left[S_Z(t)\right]^n$$
$$= P\left(\hat{S}_{KM}(t) = 1, r = 0\right).$$

For $r = 1$ we have one Z_i such that $Z_i \leq t$ and $(n-1)$ Z_j such that $Z_j > t$. Without lack of generality assume that $i = 1$ so then $Z_1 \leq t$ with probability (w.p.) $F_Z(t)$ and $Z_2, Z_3, \ldots, Z_n > t$ each w.p $S_Z(t)$. Furthermore, $\delta_{(1)} = 0$ w.p. $(1-\gamma)$ in which case

$$\hat{S}_{KM}(t) = c_{1n}^{\delta_{(1)}} = 1$$

or $\delta_{(1)} = 1$ w.p. γ in which case

$$\hat{S}_{KM}(t) = c_{1n}^{\delta_{(1)}} = c_{1n} = \frac{n-1}{n}.$$

There are $\binom{n}{1}$ ways for one of the Z_i to be less than t and the remainders to exceed t. As a result,

$$b\left(1, F_Z(t)\right) \prod_{i=1}^{1} \left[\gamma c_{in} + (1-\gamma)\right] = b\left(1, F_Z(t)\right)\left[\gamma c_{1n} + (1-\gamma)\right]$$
$$= \gamma b\left(1, F_Z(t)\right) c_{1n} + (1-\gamma) b\left(1, F_Z(t)\right)$$

which means that we need to consider two cases as described in Table 2.1;

Table 2.1. The Two Cases of Exact Distribution of the KM Estimator when $r = 1$.

$\delta_{(1)}$	Probability	$\hat{S}_{KM}(t) = c_{1n}^{\delta_{(1)}}$
0	$1-\gamma$	1
1	γ	c_{1n}

and which results in joint probabilities

$$P\left(\hat{S}_{KM}(t) = 1, r = 1\right) = (1-\gamma)b\left(1, F_Z(t)\right)$$

and

$$P\left(\hat{S}_{KM}(t) = c_{1n}, r = 1\right) = \gamma b\left(1, F_Z(t)\right).$$

For $r = 2$ there are two $Z_i \leq t$ and $(n-2)$ $Z_j > t$ which means that we need to consider four cases as described in Table 2.2;

Table 2.2. The Four Cases of Exact Distribution of the KM Estimator when $r = 2$.

$\left(\delta_{(1)}, \delta_{(2)}\right)$	Probability	$\hat{S}_{KM}(t) = c_{1n}^{\delta_{(1)}} c_{2n}^{\delta_{(2)}}$
$(0,0)$	$(1-\gamma)^2$	1
$(0,1)$	$\gamma(1-\gamma)$	c_{2n}
$(1,0)$	$\gamma(1-\gamma)$	c_{1n}
$(1,1)$	γ^2	$c_{1n}c_{2n}$

There are $\binom{n}{2}$ ways for two of the Z_i to be less than t and the remaining ones to exceed t. As a result,

$$P\left(\hat{S}_{KM}(t) = 1, r = 2\right) = (1-\gamma)^2 b\left(2, F_Z(t)\right),$$

$$P\left(\hat{S}_{KM}(t) = c_{1n}, r = 2\right) = \gamma\left(1-\gamma\right) b\left(2, F_Z(t)\right),$$

$$P\left(\hat{S}_{KM}(t) = c_{2n}, r = 2\right) = \gamma\left(1-\gamma\right) b\left(2, F_Z(t)\right)$$

and

$$P\left(\hat{S}_{KM}(t) = c_{1n}c_{2n}, r = 2\right) = \gamma^2 b\left(2, F_Z(t)\right).$$

In a similar fashion, for $r = 3$ there are three $Z_i \leq t$ and $(n-2)$ $Z_j > t$ which means that we need to consider eight cases which are described in Table 2.3;

Table 2.3. The Eight Cases of Exact Distribution of the KM Estimator when $r = 3$.

$\left(\delta_{(1)}, \delta_{(2)}, \delta_{(3)}\right)$	Probability	$\hat{S}_{KM}(t) = c_{1n}^{\delta_{(1)}} c_{2n}^{\delta_{(2)}} c_{3n}^{\delta_{(3)}}$
$(0,0,0)$	$(1-\gamma)^3$	1
$(1,0,0)$	$\gamma(1-\gamma)^2$	c_{1n}
$(0,1,0)$	$\gamma(1-\gamma)^2$	c_{2n}
$(0,0,1)$	$\gamma(1-\gamma)^2$	c_{2n}
$(1,1,0)$	$\gamma^2(1-\gamma)$	$c_{1n}c_{2n}$
$(1,0,1)$	$\gamma^2(1-\gamma)$	$c_{1n}c_{3n}$
$(0,1,1)$	$\gamma^2(1-\gamma)$	$c_{2n}c_{3n}$
$(1,1,1)$	γ^3	$c_{1n}c_{2n}c_{3n}$

There are $\binom{n}{3}$ ways for three of the Z_i to be less than t and the remainding ones to exceed t, meaning that

$$P\left(\hat{S}_{KM}(t) = 1, r = 3\right) = (1-\gamma)^3 b\left(3, F_Z(t)\right),$$

$$P\left(\hat{S}_{KM}(t) = c_{1n}, r = 3\right) = P\left(\hat{S}_{KM}(t) = c_{2n}, r = 3\right) = P\left(\hat{S}_{KM}(t) = c_{3n}, r = 3\right)$$
$$= \gamma\left(1-\gamma\right)^2 b\left(3, F_Z(t)\right),$$

$$\begin{aligned} P\left(\hat{S}_{KM}(t) = c_{1n}c_{2n}, r = 3\right) &= P\left(\hat{S}_{KM}(t) = c_{1n}c_{3n}, r = 3\right) \\ &= P\left(\hat{S}_{KM}(t) = c_{2n}c_{3n}, r = 3\right) \\ &= \gamma^2\left(1-\gamma\right) b\left(3, F_Z(t)\right) \end{aligned}$$

and

$$P\left(\hat{S}_{KM}(t) = c_{1n}c_{2n}c_{3n}, r = 3\right) = \gamma^3 b\left(3, F_Z(t)\right).$$

This process continues in an analogous fashion until the final case where for $r = n-1$ there are $(n-1)$ $Z_i \leq t$ and one $Z_j > t$ which means that we need to consider 2^{n-1} cases as described in Table 2.4; There are $\binom{n}{n-1} = n$ ways for $(n-1)$ of the Z_i to be

Table 2.4. The 2^{n-1} Cases of Exact Distribution of the KM Estimator when $r = n-1$.

$\left(\delta_{(1)}, \delta_{(2)}, \ldots, \delta_{(n-1)}\right)$	Probability	$\hat{S}_{KM}(t) = c_{1n}^{\delta_{(1)}} c_{2n}^{\delta_{(2)}} \ldots c_{(n-1)n}^{\delta_{(n-1)}}$
$(0, 0, \ldots, 0)$	$(1-\gamma)^{n-1}$	1
$(1, 0, \ldots, 0)$	$\gamma(1-\gamma)^{n-2}$	c_{1n}
$(0, 1, \ldots, 0)$	$\gamma(1-\gamma)^{n-2}$	c_{2n}
$\vdots$	$\vdots$	$\vdots$
$(0, 0, 0, \ldots, 1)$	$\gamma(1-\gamma)^{n-2}$	$c_{(n-1)n}$
$(1, 1, 0, \ldots, 0)$	$\gamma^2(1-\gamma)^{n-3}$	$c_{1n}c_{2n}$
$(1, 0, 1, \ldots, 0)$	$\gamma^2(1-\gamma)^{n-3}$	$c_{1n}c_{3n}$
$\vdots$	$\vdots$	$\vdots$
$(0, 0, \ldots, 1, 1)$	$\gamma^2(1-\gamma)^{n-3}$	$c_{(n-2)n}c_{(n-1)n}$
$(1, 1, 1, \ldots, 0)$	$\gamma^3(1-\gamma)^{n-4}$	$c_{1n}c_{2n}c_{3n}$
$\vdots$	$\vdots$	$\vdots$
$(0, \ldots, 1, 1, 1)$	$\gamma^3(1-\gamma)^{n-4}$	$c_{(n-3)n}c_{(n-2)n}c_{(n-1)n}$
$\vdots$	$\vdots$	$\vdots$
$(1, 1, \ldots, 1)$	γ^{n-1}	$c_{1n}c_{2n} \ldots c_{(n-1)n}$

less than t and the remainding one to exceed t. As a result we have that

$$P\left(\hat{S}_{KM}(t)=1, r=n-1\right)=(1-\gamma)^{n-1} b\left(n-1, F_Z(t)\right),$$

$$\begin{aligned} P\left(\hat{S}_{KM}(t)=c_{1n}, r=n-1\right) &= \cdots = P\left(\hat{S}_{KM}(t)=c_{(n-1)n}, r=n-1\right) \\ &= \gamma\left(1-\gamma\right)^{n-2} b\left(n-1, F_Z(t)\right), \end{aligned}$$

$$\begin{aligned} P\left(\hat{S}_{KM}(t)=c_{1n}c_{2n}, r=n-1\right) &= P\left(\hat{S}_{KM}(t)=c_{1n}c_{3n}, r=n-1\right) \\ &= \cdots = P\left(\hat{S}_{KM}(t)=c_{(n-2)n}c_{(n-1)n}, r=n-1\right) \\ &= \gamma^2\left(1-\gamma\right)^{n-3} b\left(n-1, F_Z(t)\right) \end{aligned}$$

and

$$P\left(\hat{S}_{KM}(t)=c_{1n}c_{2n}\ldots c_{(n-1)n}, r=n-1\right)=\gamma^{n-1} b\left(n-1, F_Z(t)\right).$$

Now we are able to compute marginal probabilities for $\hat{S}_{KM}(t)$. We have that

$$\begin{aligned} P\left(\hat{S}_{KM}(t)=1\right) &= \sum_{r=0}^{n-1} P\left(\hat{S}_{KM}(t)=1, r\right) \\ &= \sum_{r=0}^{n-1}\left(1-\gamma\right)^{r} b\left(r, F_Z(t)\right) \\ &= \left[\left(1-\gamma\right)F_Z(t)+S_Z(t)\right]^n-\left(1-\gamma\right)^n\left[F_Z(t)\right]^n \end{aligned}$$

$$\begin{aligned} P\left(\hat{S}_{KM}(t)=c_{1n}\right) &= \sum_{r=1}^{n-1} P\left(\hat{S}_{KM}(t)=c_{1n}, r\right) \\ &= \sum_{r=1}^{n-1}\gamma\left(1-\gamma\right)^{r-1} b\left(r, F_Z(t)\right) \end{aligned}$$

$$P\left(\hat{S}_{KM}(t) = c_{2n}\right) = \sum_{r=2}^{n-1} P\left(\hat{S}_{KM}(t) = c_{2n}, r\right)$$
$$= \sum_{r=2}^{n-1} \gamma\left(1-\gamma\right)^{r-1} b\left(r, F_Z(t)\right)$$

$$P\left(\hat{S}_{KM}(t) = c_{qn}\right) = \sum_{r=q}^{n-1} P\left(\hat{S}_{KM}(t) = c_{qn}, r\right)$$
$$= \sum_{r=q}^{n-1} \gamma\left(1-\gamma\right)^{r-1} b\left(r, F_Z(t)\right)$$

$$P\left(\hat{S}_{KM}(t) = c_{(n-1)n}\right) = \sum_{r=n-1}^{n-1} P\left(\hat{S}_{KM}(t) = c_{(n-1)n}, r\right)$$
$$= \gamma\left(1-\gamma\right)^{n-2} b\left(n-1, F_Z(t)\right)$$

and

$$P\left(\hat{S}_{KM}(t) = c_{1n}c_{2n}\right) = \sum_{r=2}^{n-1} P\left(\hat{S}_{KM}(t) = c_{1n}c_{2n}, r\right)$$
$$= \sum_{r=2}^{n-1} \gamma^2\left(1-\gamma\right)^{r-2} b\left(r, F_Z(t)\right)$$

$$P\left(\hat{S}_{KM}(t) = c_{1n}c_{qn}\right) = \sum_{r=q}^{n-1} P\left(\hat{S}_{KM}(t) = c_{1n}c_{2n}, r\right)$$
$$= \sum_{r=q}^{n-1} \gamma^2\left(1-\gamma\right)^{r-2} b\left(r, F_Z(t)\right)$$

$$P\left(\hat{S}_{KM}(t) = c_{1n}c_{(n-1)n}\right) = \sum_{r=n-1}^{n-1} P\left(\hat{S}_{KM}(t) = c_{1n}c_{2n}, r\right)$$
$$= \sum_{r=n-1}^{n-1} \gamma^2\left(1-\gamma\right)^{r-2} b\left(r, F_Z(t)\right)$$
$$= \gamma^2\left(1-\gamma\right)^{n-3} b\left(n-1, F_Z(t)\right)$$

and

$$P\left(\hat{S}_{KM}(t) = c_{1n}c_{2n}c_{3n}\right) = \sum_{r=3}^{n-1} P\left(\hat{S}_{KM}(t) = c_{1n}c_{2n}c_{3n}, r\right)$$
$$= \sum_{r=3}^{n-1} \gamma^3 (1-\gamma)^{r-3}\, b\left(r, F_Z(t)\right)$$

$$P\left(\hat{S}_{KM}(t) = c_{1n}c_{2n}\cdots c_{(n-1)n}\right) = \sum_{r=n-1}^{n-1} P\left(\hat{S}_{KM}(t) = c_{1n}c_{2n}\cdots c_{(n-1)n}, r\right)$$
$$= \gamma^{n-1} b\left(n-1, F_Z(t)\right).$$

Finally, in summary, we have that

$$P\left(\hat{S}_{KM}(t) = c_{1n}^{\delta_{(1)}} c_{2n}^{\delta_{(2)}} \cdots c_{(n-1)n}^{\delta_{(n-1)}}, r\right) = \gamma^{\sum_{i=1}^{n-1}\delta_{(i)}} (1-\gamma)^{r-\sum_{i=1}^{n-1}\delta_{(i)}}\, b\left(r, F_Z(t)\right) \quad (2.2)$$

This result is then used to obtain

$$P\left(\hat{S}_{KM}(t) = \prod_{i=1}^{n-1} c_{in}^{\delta_{(i)}}\right) = P\left(\hat{S}_{KM}(t) = c_{1n}^{\delta_{(1)}} c_{2n}^{\delta_{(2)}} \cdots c_{(n-1)n}^{\delta_{(n-1)}}\right)$$
$$= \sum_{r=\underline{r}_{KM}}^{n-1} P\left(\hat{S}_{KM}(t) = c_{1n}^{\delta_{(1)}} c_{2n}^{\delta_{(2)}} \cdots c_{(n-1)n}^{\delta_{(n-1)}}, r\right)$$
$$= \sum_{r=\underline{r}_{KM}}^{n-1} \gamma^{\sum_{i=1}^{n-1}\delta_{(i)}} (1-\gamma)^{r-\sum_{i=1}^{n-1}\delta_{(i)}}\, b\left(r, F_Z(t)\right)$$
$$\overset{\gamma<1}{=} \left(\frac{\gamma}{1-\gamma}\right)^{\sum_{i=1}^{n-1}\delta_{(i)}} \sum_{r=\underline{r}_{KM}}^{n-1} (1-\gamma)^r\, b\left(r, F_Z(t)\right) \quad (2.3)$$

where

$$\underline{r}_{KM} = \begin{cases} \max_{1\le i\le n-1}\left\{i : \delta_{(i)} = 1\right\} & \text{if} \quad \sum_{i=1}^{n-1}\delta_{(i)} > 0 \\ 0 & \text{if} \quad \sum_{i=1}^{n-1}\delta_{(i)} = 0 \end{cases}$$

and

$$P\left(\hat{S}_{KM}(t) = 0\right) = \left[F_Z(t)\right]^n.$$

In addition the KM estimator can assume at most $2^{n-1}+1$ distinct values.

2.1.1. Example. The Table 2.5 presents the exact distribution of the KM estimator for $t = 1$ and $n = 5$ when X_i and Y_i are exponentially distributed with unit rates so that $\gamma = 0.5$ and $S_Z(t) = e^{-2t}$ and there are at most $2^4 + 1 = 17$ distinct values for KM estimator $\hat{S}_{KM}(1)$;

Table 2.5. The Exact Distribution of the KM Estimator for $t = 1$ and $n = 5$.

$\hat{S}_{KM}(1)$	$P\left(\hat{S}_{KM}(1)\right)$
1	0.043844
0.8	0.043799
0.75	0.043074
0.667	0.038441
0.6	0.043074
0.533	0.038441
0.5	0.062081
0.4	0.062081
0.375	0.023640
0.333	0.023640
0.3	0.023640
0.267	0.023640
0.25	0.023640
0.2	0.023640
0	0.483324

Here the estimator takes on only 15 distinct values. This is because there are two ways to get a value of 0.4, i.e.

$$\left(\delta_{(1)}, \delta_{(2)}, \delta_{(3)}, \delta_{(4)}\right) = (1, 1, 1, 0) \text{ or } (1, 0, 0, 1)$$

and two ways to get a value of 0.5;

$$\left(\delta_{(1)}, \delta_{(2)}, \delta_{(3)}, \delta_{(4)}\right) = (0, 1, 1, 0) \text{ or } (0, 0, 0, 1)\,.$$

As such we would say that we have a redundancy value of one for the value of 0.4 and for the value of 0.5. The Table 2.6 provides various characteristics for the exact distribution of the KM estimator.

Table 2.6. Selected Characteristics of the Exact Distribution of the KM Estimator.

KM Exact Distribution: Selected Characteristics			
n	Number Binary Vectors	Distinct Values (% of Total)	Maximum Redundancy
5	$2^4 = 16$	14 (87.5%)	1
10	$2^9 = 512$	205 (40.0%)	10
15	$2^{14} = 16,384$	3,531 (21.6%)	49
20	$2^{19} = 524, 288$	33,422 (6.4%)	278
25	$2^{24} = 16,777,216$	65,839 (0.4%)	2457

Given the astronomical increase in the number of distinct values as n increases, computations involving the exact distribution are probably feasible for n values of at most 15. Note that the number of distinct point mass values for KM exact distribution is independent of t and changing the value of t will simply change the probabilities associated with the distinct mass values.

A similar phenomenon is observed in Chang (1996) where it is noted that the weighted average of permutation distributions representation for the KM estimator is computationally infeasible for large samples.

2.2. EXACT DISTRIBUTION OF THE PL ESTIMATOR

The results of the previous section can be used to derive the exact distribution of the PL estimator

$$\hat{S}_{PL}(t) = \prod_{i=1}^{n\hat{F}_Z(t)} c_{in}^{\delta_{(i)}}.$$

Note that this estimator differs from KM estimator;

$$\hat{S}_{KM}(t) = \left[\prod_{i=1}^{n\hat{F}_Z(t)} c_{in}^{\delta_{(i)}} \right] I\left(t \leq z_{(n)}\right).$$

in that the latter is, in effect computed, while ignoring the last product term

$$c_{nn}^{\delta_{(n)}}$$

due to the presence of the indicator function for the event

$$\left\{t \leq z_{(n)}\right\}.$$

Therefore, in analogy with formula (2.2) and its development in the previous section, we have that

$$P\left(\hat{S}_{PL}(t) = c_{1n}^{\delta(1)} c_{2n}^{\delta(2)} \cdots c_{nn}^{\delta(n)}, r\right) = \gamma^{\sum_{i=1}^{n} \delta_{(i)}} \left(1-\gamma\right)^{r-\sum_{i=1}^{n} \delta_{(i)}} b\left(r, F_Z(t)\right) \tag{2.4}$$

and

$$\begin{aligned} P\left(\hat{S}_{PL}(t) = \prod_{i=1}^{n} c_{in}^{\delta_{(i)}}\right) &= \sum_{r=\underline{r}_{PL}}^{n} P\left(\hat{S}_{PL}(t) = c_{1n}^{\delta_{(1)}} c_{2n}^{\delta_{(2)}} \cdots c_{nn}^{\delta_{(n)}}, r\right) \\ &= \sum_{r=\underline{r}_{PL}}^{n} \gamma^{\sum_{i=1}^{n} \delta_{(i)}} \left(1-\gamma\right)^{r-\sum_{i=1}^{n} \delta_{(i)}} b\left(r, F_Z(t)\right) \\ &\overset{\gamma \leq 1}{=} \left(\frac{\gamma}{1-\gamma}\right)^{\sum_{i=1}^{n} \delta_{(i)}} \sum_{r=\underline{r}_{PL}}^{n} \underline{r}_{PL} \left(1-\gamma\right)^{r} b\left(r, F_Z(t)\right) \end{aligned} \tag{2.5}$$

where

$$\underline{r}_{PL} = \begin{cases} \max_{1 \leq i \leq n} \left\{i : \delta_{(i)} = 1\right\} & \text{if} \quad \sum_{i=1}^{n} \delta_{(i)} > 0 \\ 0 & \text{if} \quad \sum_{i=1}^{n} \delta_{(i)} = 0 \end{cases}.$$

Note in particular that

$$\begin{aligned}
P\left(\hat{S}_{PL}(t)=1\right) &= \sum_{r=0}^{n} P\left(\hat{S}_{PL}(t)=1,r\right) \\
&= \sum_{r=0}^{n} (1-\gamma)^r\, b\left(r,F_Z(t)\right) \\
&= \left[(1-\gamma)\,F_Z(t)+S_Z(t)\right]^n
\end{aligned}$$

and

$$\begin{aligned}
P\left(\hat{S}_{PL}(t)=0\right) &= P\left(\hat{F}_Z(t)=1,\Delta_{(n)}=1\right) \\
&= \sum_{r=0}^{n-1} P\left(\hat{S}_{PL}(t)=0,q\right) \\
&= \sum_{r=0}^{n-1} \binom{n-1}{r} \gamma^r\,(1-\gamma)^{n-1-r}\,\gamma\,[F_Z(t)]^n \\
&= \gamma\,[F_Z(t)]^n \sum_{q=0}^{n-1} \binom{n-1}{r} \gamma^r\,(1-\gamma)^{n-1-r} \\
&= \gamma\,[F_Z(t)]^n
\end{aligned}$$

Computations involving the exact distribution of the PL estimator are probably feasible for $n \leq 15$ since the redundancies and distinct mass values for this distribution are identical to those for the KM estimator as shown in Table 2.6 except that now there are an additional 2^{n-1} ways to obtain a value of zero for $\hat{S}_{PL}(t)$.

Finally, as was the case for the KM estimator, the PL estimator can assume at most $2^{n-1}+1$ distinct mass values.

2.2.1. Example. We consider the same setting for the exact distribution as those for the example in Section 2.1.1 so that $t = 1$, $n = 5$ when X_i and Y_i are exponentially distributed with unit rates so that again $\gamma = 0.5$ and $S_Z(t) = e^{-2t}$. The exact distribution of the PL estimator under these settings is shown in Table 2.7.

Table 2.7. The Exact Distribution of the PL Estimator for $t = 1$ and $n = 5$.

$\hat{S}_{PL}(1)$	$P\left(\hat{S}_{PL}(1)\right)$
1	0.058948
0.8	0.058903
0.75	0.058178
0.667	0.053545
0.6	0.058178
0.533	0.053545
0.5	0.092289
0.4	0.092289
0.375	0.038744
0.333	0.038744
0.3	0.038744
0.267	0.038744
0.25	0.038744
0.2	0.038744
0.0	0.241662

2.3. EXACT DISTRIBUTION OF THE ACL ESTIMATOR

The ACL (Abduskhurov, Cheng and Lin) survival function estimator has the following form

$$\hat{S}_{ACL}(t) = \left[\hat{S}_Z(t)\right]^{\bar{\delta}}.$$

Note that when $\bar{\delta} = 0$ and $\hat{S}_Z(t) = 0$ then the value or the ACL estimator is 0^0 which is undefined. In this setting we adopt the convention that $0^0 = 0$ since this always results in a zero infinite tail for $\hat{S}_{ACL}(t)$ which seems to make more sense from a practical point of view than an infinite unit tail.

With this convention

$$\begin{aligned} P\left(\hat{S}_{ACL}(t)=0\right) &= P\left(\hat{S}_Z(t)=0\right) \\ &= P\left(Z_1 \le t, \ldots, Z_n \le t\right) \\ &= \left[F_Z(t)\right]^n . \end{aligned}$$

Note more generally that

$$P\left(\hat{S}_Z(t)=\frac{n-r}{n}\right)=b\left(r, F_Z\left(t\right)\right)$$

for $r=0,1,\ldots,n$ and

$$P\left(\bar{\Delta}=\frac{q}{n}\right)=b(q,\gamma)$$

for $q=0,1,\ldots,n$ so that

$$\begin{aligned} P\left(\hat{S}_{ACL}(t)=\left(\frac{n-r}{n}\right)^{\frac{q}{n}}\right) &= P\left(\hat{S}_Z(t)=\frac{n-r}{n}\right)P\left(\bar{\Delta}=\frac{q}{n}\right) \\ &= b\left(r, F_Z\left(t\right)\right)b(q,\gamma). \end{aligned}$$

The exact distribution of the ACL estimator has at most

$$\begin{aligned} (n+1)^2-2n-(n-1) &= (n+1)^2-3n+1 \\ &= n^2-n+2 \end{aligned} \tag{2.6}$$

distinct mass values. To see this note that among the $(n+1)^2$ possible ordered (r,q) pairs there are n pairs, with $q=0$, and an additional n (r,q) pairs, with $r=1$, which yield an estimator value of "1". This means that there is a redundancy value of $2n-1$

for "1". With regards to an estimator value of "0" there are $n+1$ (r,q) pairs, with $r=0$, yielding a redundancy value of n.

In fact,

$$\begin{aligned} P\left(\hat{S}_{ACL}(t)=1\right) &= P\left(\hat{S}_Z(t)=1\right)+P\left(\bar{\Delta}=0\right)-P\left(\hat{S}_Z(t)=0,\bar{\Delta}=0\right) \\ &= b\left(0,F_Z\left(t\right)\right)+b(0,\gamma)-b\left(n,F_Z\left(t\right)\right)b(0,\gamma) \\ &= \left[S_Z\left(t\right)\right]^n+\left(1-\gamma\right)^n-\left[\left(1-\gamma\right)F_Z\left(t\right)\right]^n. \end{aligned}$$

Furthermore, numerical experimentation showed that formula (2.6) provides the number of distinct mass values for $n=5,10$ and 15. For $n=20,25$ and 30 there are in fact very few redundancies beyond those already mentioned for estimator values "1" and "0".

Given the quadratic increase in the number of distinct mass values as n increases, exact distributional computations for the ACL estimator are feasible even for very large values of n.

2.3.1. Example. Here again we consider the setting for the exact distribution that was adopted in Section 2.1.1. As such, $t=1$, $n=5$ when X_i and Y_i are exponentially distributed with unit rate, $\gamma=0.5$ and $S_Z(t)=e^{-2t}$. These settings is shown in Table 2.8.

Exact distributional computations for any one of the survival estimators we consider are at best numerically intensive and at worse numerically infeasible. As such we obtain highly accurate saddlepoint approximations to these exact distributions as detailed in section 4. These approximations involve the inversion of Mellin transforms for the zero-truncated KM, PL and ACL estimators. In the next section we derive these Mellin transforms.

Table 2.8. The Exact Distribution of the ACL Estimator for $t = 1$ and $n = 5$.

$\hat{S}_{ACL}(t)$	$P\left(\hat{S}_{ACL}(t)\right)$
1.000	0.016190
0.956	0.000227
0.915	0.000453
0.903	0.002896
0.875	0.000453
0.837	0.000227
0.833	0.018501
0.815	0.005791
0.8	0.000045
0.736	0.005791
0.725	0.059101
0.693	0.037001
0.665	0.002896
0.6	0.000579
0.577	0.037001
0.5253	0.118200
0.480	0.018501
0.4	0.003700
0.380	0.118200
0.276	0.059101
0.2	0.011820
0	0.483326

3. MELLIN TRANSFORMS FOR THE ZERO-TRUNCATED ESTIMATORS

In this section, we briefly review Mellin transforms and derive these transforms for our zero-truncated KM, PL and ACL estimators.

3.1. MELLIN TRANSFORM FOR THE ZERO-TRUNCATED KME

Recall that the Mellin transform for a positive random variable X is defined as the finite vth moment of X, i.e.

$$\mathcal{M}_X(v) = E[X^v] = \int_0^\infty X^v dF(x) < \infty$$

where $-\varepsilon < v < \varepsilon$ for some $\varepsilon > 0$. Mellin transform $\mathcal{M}_X(v)$ is also the moment generating function (MGF) of $\ln(X)$. The Mellin transform is useful in the study products of independent random variables; see for instance Springer (1979) and Butler (2007). Suppose that X and Y are independent positive random variables with Mellin transforms $\mathcal{M}_X(v)$ and $\mathcal{M}_Y(v)$, respectively. Then the Mellin transform of product XY is simply the product of the Mellin transforms; $\mathcal{M}_X(v)\mathcal{M}_Y(v)$. Furthermore, from Fourier inversion theory the Mellin transform of random variable X uniquely determines its distribution. The moment expression for the KM estimator in (2.1) forms the basis for a Mellin transform. If $E\left[\hat{S}^v_{KM}(t)\right]$ were the Mellin transform for the zero-truncated KM estimate, which we denote as $\hat{S}_{KM+}(t)$, then we would have that

$$E\left[\hat{S}^0_{KM}(t)\right] = E\left[e^{0\ln\left(\hat{S}_{KM+}(t)\right)}\right] = M_{\ln\left(\hat{S}_{KM+}(t)\right)}(0) = 1$$

but from formula (2.1) we have that

$$\begin{aligned} E\left[\hat{S}^0_{KM}(t)\right] &= \sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) \prod_{i=1}^{r} \left[\gamma c^0_{in} + (1-\gamma)\right] \\ &= \sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) \\ &= 1 - \left[F_Z(t)\right]^n . \end{aligned}$$

In fact, this Mellin transform, which we shall denote as $\mathcal{M}^{KM+}_{Tr}(v)$, is given in terms of $E\left[\hat{S}^v_{KM}(t)\right]$ as

$$\begin{aligned} \mathcal{M}^{KM+}_{Tr}(v) = E\left[\hat{S}^v_{KM+}(t)\right] &= \frac{E\left[\hat{S}^v_{KM}(t)\right]}{E\left[\hat{S}^0_{KM}(t)\right]} = \frac{E\left[\hat{S}^v_{KM}(t)\right]}{P\left[\hat{S}_{KM}(t) > 0\right]} \\ &= \frac{\sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) \prod_{i=1}^{r} \left[\gamma c^v_{in} + (1-\gamma)\right]}{1 - \left[F_Z(t)\right]^n}. \end{aligned} \tag{3.1}$$

Note that

$$\left[F_Z\left(t\right)\right]^n = P\left(Z_1 \le t, \dots, Z_n \le t\right) = P\left[\hat{S}_{KM}(t) = 0\right]$$

so then

$$E\left[\hat{S}^0_{KM}(t)\right] = P\left[\hat{S}_{KM}(t) > 0\right].$$

Function $\mathcal{M}^{KM+}_{Tr}(v)$ is the Mellin transform of the strictly positive part of KM estimator $\hat{S}_{KM}(t)$. Conditioning upon the KM estimator being strictly positive is important for the use of saddlepoint approximations to reproduce the distribution of this estimator. This is because the inversion of $\mathcal{M}^{KM+}_{Tr}(v)$ to produce a saddlepoint density or CDF approximation requires that there exists some $\varepsilon > 0$ such that $\mathcal{M}^{KM+}_{Tr}(v)$ is finite for all $-\varepsilon < v < \varepsilon$. If one were to include the zero portion of KM estimator in the computations, then the resulting transform would only be finite for $0 \leq v$.

3.2. MELLIN TRANSFORM FOR THE ZERO-TRUNCATED PLE

We use method of derivation for $E\left[\hat{S}^v_{KM}(t)\right]$ from Chen et al. (1982) with $0 \leq v$ to derive $E\left[\hat{S}^v_{PL}(t)\right]$ for $0 \leq v$. It turns out that

$$\begin{aligned} E\left[\hat{S}^v_{PL}(t)\right] &= \sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) \prod_{i=1}^{r} \left[\gamma c^v_{in} + (1-\gamma)\right] \\ &\quad + (1-\gamma)\left[F_Z(t)\right]^n \prod_{i=1}^{n-1} \left[\gamma c^v_{in} + (1-\gamma)\right] \qquad (3.2) \\ &= E\left[\hat{S}^v_{KM}(t)\right] + (1-\gamma)\left[F_Z(t)\right]^n \prod_{i=1}^{n-1} \left[\gamma c^v_{in} + (1-\gamma)\right]. \qquad (3.3) \end{aligned}$$

To see why this is the case note that an equivalent expression for the PL estimator is

$$\hat{S}_{PL}(t) = \prod_{i=1}^{n} c_{in}^{v\delta_i I(z_i \leq t)}.$$

From the double expectation formula we have that

$$E\left[\hat{S}^v_{PL}(t)\right] = E_Z\left\{E_{\Delta|Z}\left[\prod_{i=1}^{n} c_{in}^{v\Delta_i I(Z_i \leq t)}\right]\right\}$$

where

$$Z = (Z_1, \ldots, Z_n)$$

and

$$\Delta = (\Delta_1, \ldots, \Delta_n).$$

Under the proportional hazards model, Z_i and Δ_i are independent for $i = 1, \ldots, n$. Therefore, we have that

$$E_{\Delta|Z}\left[\prod_{i=1}^{n} c_{in}^{v\Delta_i I(Z_i \le t)}\right] = \prod_{i=1}^{n} E_{\Delta_i|Z_i}\left(c_{in}^{v\Delta_i I(Z_i \le t)}\right)$$

where for $1 \le i \le n-1$

$$\begin{aligned} E_{\Delta_i|Z_i}\left[c_{in}^{v\Delta_i I(Z_i \le t)}\right] &= \begin{cases} 1 & \text{if} \quad Z_i > t \\ \gamma c_{in}^{v} + (1-\gamma) & \text{if} \quad Z_i \le t \end{cases} \\ &= \gamma c_{in}^{vI(Z_i \le t)} + (1-\gamma). \end{aligned}$$

and for $i = n$ we have that

$$E_{\Delta_n|Z_n}\left[c_{nn}^{v\Delta_n I(Z_n \le t)}\right] = (1-\gamma)^{I(Z_n \le t)}$$

since

$$c_{nn}^{v\Delta_n I(Z_n \le t)} = 1$$

if and only if $\Delta_n = 0$ or $Z_n > t$.

This then yields

$$\begin{aligned} E_{\Delta|Z}\left[\prod_{i=1}^{n} c_{in}^{v\Delta_i I(Z_i \le t)}\right] &= (1-\gamma)^{I(Z_n \le t)} \prod_{i=1}^{n-1}\left[\gamma c_{in}^{vI(Z_i \le t)} + (1-\gamma)\right] \\ &= \prod_{i=1}^{n\hat{F}_Z(t)} \left[\gamma c_{in}^{v} + (1-\gamma)\right] \end{aligned}$$

and finally

$$\begin{aligned} E\left[\hat{S}_{PL}^{v}(t)\right] = E_Z\left\{E_{\Delta|Z}\left[\prod_{i=1}^{n} c_{in}^{v\Delta_i I(Z_i \le t)}\right]\right\} &= E_Z\left\{\prod_{i=1}^{n\hat{F}_Z(t)} \left[\gamma c_{in}^{v} + (1-\gamma)\right]\right\} \\ = \sum_{r=0}^{n} b\left(r, F_Z\left(t\right)\right) \prod_{i=1}^{r} \left[\gamma c_{in}^{v} + (1-\gamma)\right]. \end{aligned}$$

An equivalent expression which removes the troublesome $c_{nn} = 0$ term is as follows

$$
\begin{aligned}
E\left[\hat{S}^v_{PL}(t)\right] &= \sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) \prod_{i=1}^{r} \left[\gamma c^v_{in} + (1-\gamma)\right] \\
&+ b\left(n, F_Z(t)\right) \prod_{i=1}^{n} \left[\gamma c^v_{in} + (1-\gamma)\right] \\
&= \sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) \prod_{i=1}^{r} \left[\gamma c^v_{in} + (1-\gamma)\right] + \left[F_Z(t)\right]^n \prod_{i=1}^{n} \left[\gamma c^v_{in} + (1-\gamma)\right] \\
&= \sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) \prod_{i=1}^{r} \left[\gamma c^v_{in} + (1-\gamma)\right] \\
&+ (1-\gamma)\left[F_Z(t)\right]^n \prod_{i=1}^{n-1} \left[\gamma c^v_{in} + (1-\gamma)\right] \\
&= E\left[\hat{S}^v_{KM}(t)\right] + (1-\gamma)\left[F_Z(t)\right]^n \prod_{i=1}^{n-1} \left[\gamma c^v_{in} + (1-\gamma)\right].
\end{aligned}
$$

This moment expression is easily seen to be valid for $-\infty < v < \infty$. When $v = 0$ we have that

$$
\begin{aligned}
E\left[\hat{S}^0_{PL}(t)\right] &= E\left[\hat{S}^0_{KM}(t)\right] + (1-\gamma)\left[F_Z(t)\right]^n \prod_{i=1}^{n-1} \left[\gamma c^0_{in} + (1-\gamma)\right] \\
&= 1 - \left[F_Z(t)\right]^n + (1-\gamma)\left[F_Z(t)\right]^n \\
&= 1 - \gamma\left[F_Z(t)\right]^n
\end{aligned}
$$

Finally, the Mellin transform for the zero-truncated PL estimate, which we denote as $\hat{S}_{PL+}(t)$, is given in terms of the $E\left[\hat{S}^v_{PL}(t)\right]$ as

$$
\begin{aligned}
\mathcal{M}^{PL+}_{Tr}(v) = E\left[\hat{S}^v_{PL+}(t)\right] &= \frac{E\left[\hat{S}^v_{PL}(t)\right]}{E\left[\hat{S}^0_{PL}(t)\right]} = \frac{E\left[\hat{S}^v_{PL}(t)\right]}{P\left[\hat{S}_{PL}(t) > 0\right]} \\
&= \frac{E\left[\hat{S}^v_{KM}(t)\right] + (1-\gamma)\left[F_Z(t)\right]^n \prod_{i=1}^{n-1} \left[\gamma c^v_{in} + (1-\gamma)\right]}{1 - \gamma\left[F_Z(t)\right]^n}.
\end{aligned}
$$

3.3. MELLIN TRANSFORM FOR THE ZERO-TRUNCATED ACLE

Here again we use method of derivation from Chen et al. (1982) to derive $E\left[\hat{S}^v_{ACL}(t)\right]$ for $0 \leq v$. From the double expectation formula we have that

$$E\left[\hat{S}^v_{ACL}(t)\right] = E_Z\left[E_{\Delta|Z}\left\{\left[\hat{S}^v_Z(t)\right]^{\bar{\Delta}}\right\}\right]$$

where

$$Z = (Z_1, \ldots, Z_n)$$

and

$$\Delta = (\Delta_1, \ldots, \Delta_n)\,.$$

Due to the independence of Z_i and Δ_i we have that

$$E_{\Delta|Z}\left\{\left[\hat{S}^v_Z(t)\right]^{\bar{\Delta}}\right\} = \prod_{i=1}^{n} E_{\Delta_i|Z}\left(\left[\hat{S}^{\frac{v}{n}}_Z(t)\right]^{\Delta_i}\right)$$

where

$$E_{\Delta_i|Z}\left(\left[\hat{S}^{\frac{v}{n}}_Z(t)\right]^{\Delta_i}\right) = \begin{cases} \gamma\hat{S}^{\frac{v}{n}}_Z(t) + (1-\gamma) & \text{if} \quad \hat{S}_Z(t) > 0 \\ 0 & \text{if} \quad \hat{S}_Z(t) = 0 \end{cases}.$$

Therefore, we have that

$$E_{\Delta|Z}\left\{\left[\hat{S}^v_Z(t)\right]^{\bar{\Delta}}\right\} = \begin{cases} \left[\gamma\hat{S}^{\frac{v}{n}}_Z(t) + (1-\gamma)\right]^n & \text{if} \quad \hat{S}_Z(t) > 0 \\ 0 & \text{if} \quad \hat{S}_Z(t) = 0 \end{cases}$$

and finally

$$E\left[\hat{S}^v_{ACL}(t)\right] = E_Z\left[E_{\Delta|Z}\left\{\left[\hat{S}^v_Z(t)\right]^{\bar{\Delta}}\right\}\right]$$
$$= \sum_{r=0}^{n-1} b\left(r, F_Z\left(t\right)\right)\left[\gamma\left(\frac{n-r}{n}\right)^{\frac{v}{n}} + (1-\gamma)\right]^n.$$

This moment expression is easily seen to be valid for $-\infty < v < \infty$. Note that Zhang et al. (2006) present an equivalent formula for $E\left[\hat{S}^v_{ACL}(t)\right]$ with $v > 0$. When $v = 0$ we have that

$$E\left[\hat{S}^0_{ACL}(t)\right] = \sum_{r=0}^{n-1} b\left(r, F_Z\left(t\right)\right)\left[\gamma\left(\frac{n-r}{n}\right)^{\frac{0}{n}} + (1-\gamma)\right]^n$$
$$= \sum_{r=0}^{n-1} b\left(r, F_Z\left(t\right)\right)$$
$$= 1 - \left[F_Z(t)\right]^n.$$

The Mellin transform for the zero-truncated ACL estimate, which we denote as $\hat{S}_{ACL+}\left(t\right)$, is of the form

$$\mathcal{M}^{ACL+}_{Tr}\left(v\right) = E\left[\hat{S}^v_{ACL+}\left(t\right)\right] = \frac{E\left[\hat{S}^v_{ACL}(t)\right]}{E\left[\hat{S}^0_{ACL}(t)\right]} = \frac{E\left[\hat{S}^v_{ACL}(t)\right]}{P\left[\hat{S}_{ACL}(t) > 0\right]}$$
$$= \frac{\sum_{r=0}^{n-1} b\left(r, F_Z\left(t\right)\right)\left[\gamma\left(\frac{n-r}{n}\right)^{\frac{v}{n}} + (1-\gamma)\right]^n}{1 - \left[F_Z(t)\right]^n}.$$

4. POINTWISE CONFIDENCE BANDS FOR SURVIVAL FUNCTIONS

In this section, we review the existing confidence band methods which we will consider in the simulations performed in section 5 and we develop novel confidence band methods for the KM, PL and ACL estimators based upon the exact distributions of these estimators and the saddlepoint approximations to these distributions.

4.1. EXISTING METHODS

The first classical method which we will compare with our confidence band methods for the KM and PL estimators is the exponential Greenwood confidence band. While this method was developed for the KM estimator, it is a large sample method, so given the asymptotic equivalence of the KM and PL estimators it will also be a good comparator for the confidence band method we develop for the PL estimator.

Greenwood (1926) provides a formula for the (approximate) asymptotic variance of $\hat{S}_{KM}(t)$ as

$$\widehat{Var}\left\{\hat{S}_{KM}(t)\right\} = \hat{S}_{KM}(t)^2 \sum_{t_{(i)} \leq t} \frac{d_{(i)}}{n_{(i^-)}\left(n_{(i^-)} - d_{(i)}\right)}$$

where $t_{(i)}$ as the ith ordered time of "death", $d_{(i)}$ is the number of deaths recorded at time $t_{(i)}$ and $n_{(i^-)}$ is the total number of individuals at risk of death an instant just before time $t_{(i)}$.

A natural 95% pointwise confidence band for survival function $S(t)$ is

$$\hat{S}_{KM}(t) \pm 1.96\sqrt{\widehat{Var}\left\{\hat{S}_{KM}(t)\right\}}.$$

This confidence band does not work well with small samples because the upper and lower bands can easily fall outside of unit interval $(0, 1)$. Therefore, we opted to use the 95% exponential Greenwood confidence band (Kalbfleisch and Prentice, 1980). This method has as its basis the following approximate variance result

$$\widehat{Var}\left\{\ln\left(-\ln\hat{S}_{KM}(t)\right)\right\} = \frac{1}{\left[\ln\hat{S}_{KM}(t)\right]^2}\sum_{t_{(i)}\leq t}\frac{d_{(i)}}{n_{(i^-)}\left(n_{(i^-)}-d_{(i)}\right)}$$

which naturally leads to the following 95% confidence band for $\ln\left(-\ln S\left(t\right)\right)$

$$\ln\left(-\ln\hat{S}_{KM}(t)\right) \pm 1.96\widehat{Var}\left\{\ln\left(-\ln\hat{S}_{KM}(t)\right)\right\} = \left(\widehat{LB}\left(t\right), \widehat{UB}\left(t\right)\right).$$

Now 95% confidence bands for $S\left(t\right)$ is gotten by applying the

$$k\left(x\right) = \exp\left\{-\exp\left(x\right)\right\}$$

tranformation to the above confidence bounds to yield

$$\left(\exp\left\{-\exp\left(\widehat{UB}\left(t\right)\right)\right\}, \exp\left\{-\exp\left(\widehat{LB}\left(t\right)\right)\right\}\right).$$

This confidence band has the advantage that its upper and lower bands are guaranteed fall inside unit interval $\left(0,1\right)$. Borgan and Leistøl (1990) found that this confidence band method performed well for $n \geq 25$ and when as much as 50% of observations have been censored.

Jeong and Cho (2002) provide a large sample 95% confidence band for $S(t)$ of the form

$$\hat{S}_{ACL}(t) \pm 1.96\sqrt{\widehat{Var}\left[\hat{S}_{ACL}(t)\right]}$$

where

$$\widehat{Var}\left[\hat{S}_{ACL}(t)\right] = \bar{\delta}^2\left[\hat{S}_Z(t)\right]^{2\bar{\delta}-1}\hat{F}_Z\left(t\right) + \bar{\delta}\left(1-\bar{\delta}\right)\left[\hat{S}_Z(t)\right]^{2\bar{\delta}} \times \left[\ln\hat{S}_Z(t)\right]^2.$$

To the best of our knowledge little is known about the performance of this confidence band in finite samples. Jeong and Cho (2002) only mention it in passing and consider the performance of their proposed confidence band for the median survival time in their simulation studies.

4.2. PROPOSED METHODS

To generate a $(1-\alpha)\,100\%$ confidence bands for survival function $S\,(t)$, which are denoted as $\left(\hat{S}_L(t), \hat{S}_U(t)\right)$, from a some survival estimator $\hat{S}(t)$, we solve equations

$$\hat{P}\left(\hat{S}(t) \leq \hat{S}_L(t)\right) = \alpha/2$$
$$\hat{P}\left(\hat{S}(t) \leq \hat{S}_U(t)\right) = 1-\alpha/2$$

where the probabilities are estimated with the sample data. As such, confidence bounds $\hat{S}_L(t)$ and $\hat{S}_U(t)$ represent $(\alpha/2)$th and $(1-\alpha/2)$th quantiles, respectively, of the boostrap distribution for survival estimator $\hat{S}(t)$.

For any fixed value of t, the KM, PL and ACL estimators have estimated distributions which are discrete and non-lattice. As a result, one typically cannot solve the above equations exactly. One approach is to use linear interpolation to obtain an approximate solution. When this is done, one is in fact computing a mid-quantile, as defined in Parzen (2008), and the resulting CI will correspond to a mid-p-value CI. Agresti (1992) and Routledge (1994) advocated the use of mid-p-values when forming CIs based on discrete distributions.

Parzen (2008) posits that a theoretical justification mid-p-value inference is that inversion formulas for a univariate discrete probability mass function (PMF) or CDF from their associated characteristic function. These inversion formulas are valid in the convex hull of a random variable's support; meaning that these formulas work at the mass points as well as all points between the mass points.

However, characteristic function inversion integrals for probability density functions (PDF) or CDFs of a random variable X can rarely be evaluated in closed-form; see for instance Billingsley (1986). When the complex-valued integrand of these inversion integrals has a single saddlepoint then one may apply the classical method of steepest descent for complex-valued integrals to the inversion intergrals to obtain highly accurate saddlepoint PDF and CDF approximations; see Butler (2007) for further elaboration. Saddlepoint approximations were first proposed in Daniels (1954) where it is also shown the characteristic function inversion integrals have a single saddlepoint exactly when $M_X(v)$, the MGF of X, exists.

In such a case, the saddlepoint PDF approximation from Daniels (1954) is of the form

$$\hat{f}(x) = \frac{1}{\sqrt{2\pi K''_X(\hat{v})}} \exp\{K_X(\hat{v}) - \hat{v}x\}$$

where $K_X(v) = \ln[M_X(v)]$ is the cumulant generating function (CGF) for X, $K'_X(v)$ and $K''_X(v)$ are the first and second derivatives of the CGF, respectively, and saddlepoint $\hat{v}$ solves the saddlepoint equation

$$K'_X(\hat{v}) = x.$$

Luganani and Rice (1980) (LR) provide a saddlepoint CDF approximation of the form

$$\hat{F}(x) = \begin{cases} \Phi(\hat{w}) + \phi(\hat{w})\,[\hat{w}^{-1} - \hat{u}^{-1}], & \text{if } x \neq E(X) \\ \frac{1}{2} + K'''_X(0)\left[72\pi K''_X(0)^3\right]^{-1/2}, & \text{if } x = E(X) \end{cases} \tag{4.1}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal CDF and normal PDF, respectively, $K'''_X(v)$ is the third derivative of the CGF $K_X(v)$,

$$\hat{w} = sgn(\hat{v})\sqrt{2(\hat{v}x - K_X(\hat{v}))}$$

and

$$\hat{u} = \hat{v}\sqrt{K''_X(\hat{v})}.$$

Davison and Wang (2002) show that for the problem in which a PMF is replaced by a saddlepoint approximation is continuous and provides a p-value which is a good approximation to the mid-p-value. In addition, Butler (2007) shows that saddlepoint approximations generally perform very in terms of accurately approximating non-normal distributions and offering substantial improvements over existing methods even in very small samples. We shall see in the simulations Section 5 that saddlepoint approximations perform remarkably well in approximating the exact distributions of KM, PL and ACL estimators and the resulting confidence bands are very close to confidence bands one would obtain from an exact distribution directly.

4.2.1. Pointwise Population Tolerance Intervals. For concreteness, we consider the classical and mid-p-value definitions of population quantiles for KM estimator $\hat{S}_{KM}(1)$, with $n = 5$, based on unit exponential survival and censoring times whose exact PMF was derived in Section 2.1 and associated exact CDF is shown in Table 4.1;

Table 4.1. The exact distribution of the KM estimator with associated exact CDF for $t = 1$ and $n = 5$.

$\hat{S}_{KM}(1)$	PMF	CDF
1	0.043844	1.000000
0.8	0.043799	0.956155
0.75	0.043074	0.912356
0.667	0.038441	0.869282
0.6	0.043074	0.830842
0.533	0.038441	0.787768
0.5	0.062081	0.749327
0.4	0.062081	0.687246
0.375	0.023640	0.625165
0.333	0.023640	0.601525
0.3	0.023640	0.577885
0.267	0.023640	0.554245
0.25	0.023640	0.530604
0.2	0.023640	0.506964
0	0.483324	0.483324

More specifically, we consider the calculation of interval $\left(\hat{S}^{pop}_{KM,L}(1), \hat{S}^{pop}_{KM,U}(1)\right)$ which satisfies equations

$$P\left(\hat{S}_{KM}(1) \leq \hat{S}^{pop}_{KM,L}(1)\right) = \alpha/2$$
$$P\left(\hat{S}_{KM}(1) \leq \hat{S}^{pop}_{KM,U}(1)\right) = 1 - \alpha/2$$

under the exact distribution for $\hat{S}_{KM}(1)$. Note that $\left(\hat{S}^{pop}_{KM,L}(1), \hat{S}^{pop}_{KM,U}(1)\right)$ is in fact the $(1-\alpha)\,100\%$ two-sided symmetric population tolerance interval for $\hat{S}_{KM}(1)$ since

$$P\left(\hat{S}^{pop}_{KM,L}(1) \leq \hat{S}_{KM}(1) \leq \hat{S}^{pop}_{KM,U}(1)\right) \geq 1 - \alpha.$$

Also, previously mentioned the sample version of this tolerance interval, in which $\hat{F}_Z\,(1)$ replaces $F_Z\,(1)$ and $\hat{\gamma}$ replaces γ in the PMF formulas for $\hat{S}_{KM}(1)$, would be denoted as $\left(\hat{S}_{KM,L}(1), \hat{S}_{KM,U}(1)\right)$ and corresponds to the $(1-\alpha)\,100\%$ bootstrap CI for $E\left[\hat{S}_{KM}(1)\right]$.

Under the classical definition of population quantile x_p, the pth quantile for discrete random variable X, is any number which satisfies equations

$$P(X \leq x_p) \geq p$$
$$P(X \geq x_p) \geq 1 - p.$$

If for instance we consider the exact distribution of the KM estimator from above and let $p = 0.025$ then

$$P\left(\hat{S}_{KM}(1) \leq 0\right) = 0.48332 \geq 0.025,$$
$$P\left(\hat{S}_{KM}(1) \geq 0\right) = 1 \geq 1 - 0.025,$$

and for $p = 0.975$ we have

$$P\left(\hat{S}_{KM}(1) \leq 1\right) = 1 \geq 0.975$$
$$P\left(\hat{S}_{KM}(1) \geq 1\right) = 0.04384 \geq 1 - 0.975$$

which yields a fairly non-informative 95% two-sided population tolerance interval for $\hat{S}_{KM}(1)$ of $(0, 1)$.

If instead we consider a tolerance interval based upon a mid-p correction we find that

$$\hat{S}^{pop}_{KM,L}(1) = 0$$

since

$$P\left(\hat{S}_{KM}(1) = 0\right) = 0.483324 > 0.025$$

and

$$\hat{S}^{pop}_{KM,U}(1) = 0.8 + (0.975 - 0.956)\,\frac{1.0 - 0.8}{1 - 0.956} = 0.886.$$

Note that with regards to 95% tolerance intervals for other values of t if

$$P\left(\hat{S}_{KM}(t) = 1\right) \geq 0.975$$

then

$$\hat{S}^{pop}_{KM,U}(t) = 1.$$

If, on the other hand,

$$P\left(\hat{S}_{KM}(t) = 0\right) \geq 0.975$$

then

$$\left(\hat{S}^{pop}_{KM,L}(t), \hat{S}^{pop}_{KM,U}(t)\right) = (0, 0) .$$

Finally, if

$$0.025 \leq P\left(\hat{S}_{KM}(t) = 0\right) \leq 0.975$$

then

$$\hat{S}^{pop}_{KM,U}(t) = 0.$$

Note also that, in lieu of working with the exact distributions of our KM, PL and ACL estimators to obtain population tolerance intervals, which can be quite cumbersome when these distributions have a large number of distinct point mass values, one could consider the highly accurate LR saddlepoint CDF approximation, instead.

The first step in this method is to note that the CDF for $\hat{S}_{KM}(t)$ maybe written as

$$\begin{aligned} P\left(\hat{S}_{KM}(t) \leq x\right) &= P\left(\hat{S}_{KM}(t) \leq x | \hat{S}_{KM}(t) > 0\right) P\left(\hat{S}_{KM}(t) > 0\right) \\ &\quad + P\left(\hat{S}_{KM}(t) \leq x | \hat{S}_{KM}(t) = 0\right) P\left(\hat{S}_{KM}(t) = 0\right) \\ &= P\left(\hat{S}_{KM}(t) \leq x | \hat{S}_{KM}(t) > 0\right) \{1 - [F_Z(t)]^n\} + [F_Z(t)]^n \\ &= P\left(\ln\left(\hat{S}_{KM}(t)\right) \leq \ln(x) | \hat{S}_{KM}(t) > 0\right) \{1 - [F_Z(t)]^n\} + [F_Z(t)]^n . \end{aligned}$$

Then a saddlepoint approximation to this CDF is obtained by replacing

$$P\left(\hat{S}_{KM}(t) \leq x | \hat{S}_{KM}(t) > 0\right) = P\left(\ln\left(\hat{S}_{KM}(t)\right) \leq \ln(x) | \hat{S}_{KM}(t) > 0\right)$$

with

$$\hat{P}\left(\ln\left(\hat{S}_{KM}(t)\right) \leq \ln(x) | \hat{S}_{KM}(t) > 0\right) = \hat{P}\left(\hat{S}_{KM}(t) \leq x | \hat{S}_{KM}(t) > 0\right),$$

its LR saddlepoint approximation generated from Mellin transform $\mathcal{M}_{Tr}^{KM+}(v)$. The resulting saddlepoint approximation is of the form

$$\hat{P}\left(\hat{S}_{KM}(t) \leq x\right) = \hat{P}\left(\hat{S}_{KM}(t) \leq x | \hat{S}_{KM}(t) > 0\right)\left\{1 - [F_Z(t)]^n\right\} + [F_Z(t)]^n .$$

When determining a $(1-\alpha)\,100\%$ two-sided symmetric population tolerance interval for $\hat{S}_{KM}(t)$ one would solve the equations

$$\hat{P}\left(\hat{S}_{KM}(t) \leq x | \hat{S}_{KM}(t) > 0\right) = \left\{\frac{\alpha/2 - [F_Z(t)]^n}{1 - [F_Z(t)]^n}\right\}_{[0,1]}$$
$$\hat{P}\left(\hat{S}_{KM}(t) \leq x | \hat{S}_{KM}(t) > 0\right) = \left\{\frac{1 - \alpha/2 - [F_Z(t)]^n}{1 - [F_Z(t)]^n}\right\}_{[0,1]}$$

where the two-sided rounding function $\{\cdot\}_{[0,1]}$ is defined as

$$\{x\}_{[0,1]} = \begin{cases} 0 & \text{if} \quad x < 0 \\ x & \text{if} \quad 0 \leq x \leq 1 \\ 1 & \text{if} \quad x > 1 \end{cases} .$$

One issue which arises in using our saddlepoint approximations is that the LR saddlepoint CDF approximation is not defined on boundary of the support for a random variable. For instance with random variable $\hat{S}_{KM}(t)|\hat{S}_{KM}(t) > 0$ saddlepoint $\hat{v} = \infty$ for $x = 1$ and $\hat{v} = -\infty$ for $x = \frac{1}{n}$. In our computations we can avoid the boundary since one can always compute the exact probabilities for $x = 1$ and $x = \frac{1}{n}$. This is because, for each of these values, there is only one product of fractions, in each case, which correspond to their values. This is also the case for $x = \frac{n-1}{n}$ and $x = \frac{1}{n-1}$, the next two support points which are closest to the support boundary points. Recall that

$$P\left(\hat{S}_{KM}(t) = 1\right) = [(1-\gamma)\,F_Z(t) + S_Z(t)]^n - [(1-\gamma)\,F_Z(t)]^n ,$$

$$P\left(\hat{S}_{KM}(t) = \frac{1}{n}\right) = \gamma^{n-1} b\,(n-1, F_Z(t))\,,$$

$$P\left(\hat{S}_{KM}(t) = \frac{n-1}{n}\right) = \sum_{r=1}^{n-1} \gamma\,(1-\gamma)^{r-1}\, b\,(r, F_Z(t))$$

and

$$P\left(\hat{S}_{KM}\left(t\right)=\frac{1}{n-1}\right)=\gamma^{n-2}\left(1-\gamma\right)b\left(n-1,F_{Z}(t)\right).$$

Now we can obtain exact (interpolated) mid-p CDF values over the intervals

$$\frac{1}{n}\leq x<\frac{1}{n-1}$$

and

$$\frac{n-1}{n}\leq x<1$$

and use the LR saddlepoint CDF approximation over the interval

$$\frac{1}{n-1}\leq x<\frac{n-1}{n}.$$

This results in an adjusted LR saddlepoint CDF approximation of the form

$$\hat{P}\left(\hat{S}_{KM}\left(t\right)\leq x|\hat{S}_{KM}\left(t\right)>0\right)$$
$$=\begin{cases} 0 & \text{if} \quad x<\frac{1}{n} \\ \frac{\left[n(n-1)\left(x-\frac{1}{n-1}\right)\right]P\left(\hat{S}_{KM}(t)=\frac{1}{n-1}\right)+P\left(0<\hat{S}_{KM}(t)\leq\frac{1}{n-1}\right)}{1-[F_{Z}(t)]^{n}} & \text{if} \quad \frac{1}{n}\leq x<\frac{1}{n-1} \\ \Phi\left(\hat{w}\right)+\phi\left(\hat{w}\right)\left(\frac{1}{\hat{w}}-\frac{1}{\hat{u}}\right) & \text{if} \quad \frac{1}{n-1}\leq x<\frac{n-1}{n} \\ \frac{n(x-1)P\left(\hat{S}_{KM}(t)=1\right)}{1-[F_{Z}(t)]^{n}}+1 & \text{if} \quad \frac{n-1}{n}\leq x<1 \\ 1 & \text{if} \quad x\geq 1 \end{cases}.$$

Note that the adjusted LR saddlepoint CDF approximation for the PL estimator will have same form as that above with exception that “KM” is replaced with “PL” throughout and the term

$$P\left(\hat{S}_{KM}\left(t\right)>0\right)=1-\left[F_{Z}\left(t\right)\right]^{n}$$

is replaced by a

$$P\left(\hat{S}_{PL}\left(t\right)>0\right)=1-\gamma\left[F_{Z}\left(t\right)\right]^{n}.$$

In a similar fashion we can define a saddlepoint CDF approximation for the ACL estimator for use in approximating population tolerance intervals. This approximation is of the form

$$\hat{P}\left(\hat{S}_{ACL}(t)\leq x\right)=\hat{P}\left(\ln\left(\hat{S}_{ACL}(t)\right)\leq\ln\left(x\right)|\hat{S}_{ACL}(t)>0\right)\{1-\left[F_{Z}(t)\right]^{n}\}+\left[F_{Z}(t)\right]^{n}.$$

where

$$\hat{P}\left(\ln\left(\hat{S}_{ACL}(t)\right)\leq\ln\left(x\right)|\hat{S}_{ACL}(t)>0\right)=\hat{P}\left(\hat{S}_{ACL}(t)\leq x|\hat{S}_{ACL}(t)>0\right)$$

its LR saddlepoint CDF approximation based on Mellin transform $\mathcal{M}_{Tr}^{ACL+}\left(v\right)$. Here, as was the case for the KM and PL estimators, we determine a $\left(1-\alpha\right)100\%$ two-sided symmetric population tolerance interval for $\hat{S}_{ACL}(t)$ by solving equations

$$\begin{aligned}\hat{P}\left(\hat{S}_{ACL}(t)\leq x|\hat{S}_{ACL}(t)>0\right)&=\left\{\frac{\alpha/2-\left[F_{Z}(t)\right]^{n}}{1-\left[F_{Z}(t)\right]^{n}}\right\}_{[0,1]}\\\hat{P}\left(\hat{S}_{ACL}(t)\leq x|\hat{S}_{ACL}(t)>0\right)&=\left\{\frac{1-\alpha/2-\left[F_{Z}(t)\right]^{n}}{1-\left[F_{Z}(t)\right]^{n}}\right\}_{[0,1]}.\end{aligned}$$

Also, as before, the LR saddlepoint CDF approximation for $\hat{S}_{ACL}(t)|\hat{S}_{ACL}(t)>0$ will not exist on the support boundary for this random variable. Fortunately, it is possible to compute, in closed-form, probabilites for the two mass points closest to the upper bound of the conditional support;

$$P\left(\hat{S}_{ACL}(t)=1\right)=\left[S_{Z}\left(t\right)\right]^{n}+\left(1-\gamma\right)^{n}-\left[\left(1-\gamma\right)F_{Z}\left(t\right)\right]^{n}$$

and

$$P\left(\hat{S}_{ACL}(t)=\left(\frac{n-1}{n}\right)^{\frac{1}{n}}\right)=b\left(1,F_Z\left(t\right)\right)b(1,\gamma),$$

and for the two mass points closest to the lower bound of the conditional support;

$$P\left(\hat{S}_{ACL}(t)=\frac{1}{n}\right)=b\left(n-1,F_Z\left(t\right)\right)b(n,\gamma)$$

and

$$P\left(\hat{S}_{ACL}(t)=\left(\frac{1}{n}\right)^{\frac{n-1}{n}}\right)=b\left(n-1,F_Z\left(t\right)\right)b(n-1,\gamma).$$

Now, in a fashion similar to what was done for the KM and PL estimators, we can obtain exact (interpolated) mid-p CDF values over the intervals

$$\frac{1}{n}\leq x<\left(\frac{1}{n}\right)^{\frac{n-1}{n}}$$

and

$$\left(\frac{n-1}{n}\right)^{\frac{1}{n}}\leq x<1$$

and use the LR saddlepoint CDF approximation, generated from Mellin transform $\mathcal{M}_{Tr}^{ACL+}(v)$, over the interval

$$\left(\frac{1}{n}\right)^{\frac{n-1}{n}}\leq x<\left(\frac{n-1}{n}\right)^{\frac{1}{n}}.$$

This results in an adjusted LR saddlepoint CDF approximation of the form

$$\hat{P}\left(\hat{S}_{ACL}(t) \leq x | \hat{S}_{ACL}(t) > 0\right)$$
$$= \begin{cases} 0 & \text{if} \quad x < \frac{1}{n} \\ \frac{n\left(x-\left(\frac{1}{n}\right)^{\frac{n-1}{n}}\right)P\left(\hat{S}_{ACL}(t)=\left(\frac{1}{n}\right)^{\frac{n-1}{n}}\right)}{\left(\sqrt[n]{n}-1\right)(1-[F_Z(t)]^n)} + \frac{P\left(0<\hat{S}_{ACL}(t)\leq\left(\frac{1}{n}\right)^{\frac{n-1}{n}}\right)}{1-[F_Z(t)]^n} & \text{if} \quad \frac{1}{n} \leq x < \left(\frac{1}{n}\right)^{\frac{n-1}{n}} \\ \Phi(\hat{w}) + \phi(\hat{w})\left(\frac{1}{\hat{w}} - \frac{1}{\hat{u}}\right) & \text{if} \quad \left(\frac{1}{n}\right)^{\frac{n-1}{n}} \leq x < \left(\frac{n-1}{n}\right)^{\frac{1}{n}} \\ \frac{\left[1-\left(\frac{n-1}{n}\right)^{\frac{1}{n}}\right]^{-1}(x-1)P\left(\hat{S}_{ACL}(t)=1\right)}{1-[F_Z(t)]^n} + 1 & \text{if} \quad \left(\frac{n-1}{n}\right)^{\frac{1}{n}} \leq x < 1 \\ 1 & \text{if} \quad x \geq 1 \end{cases}.$$

Below we have graphs of the 95% two-sided symmetric population tolerance bounds for the KM, PL and ACL estimators obtained from the exact interpolated CDFs and from the adjusted LR saddlepoint CDF approximations.

Here X_i and Y_i are exponentially distributed with rates 1 and β, respectively, so that

$$\gamma = \frac{1}{1+\beta}$$

where $\beta = 0.5, 1.0, 1.5$ and 2.0. In addition for each one of these β settings we consider samples of size $n = 5, 10, 15, 20, 25$ and 30. Note however that for the KM and PL estimators we only consider exact interpolated CDF computations when $n \leq 15$ since for larger values of n it is numerically very burdensome, if at all possible, to perform these computations. However, for the ACL estimator we were able to perform exact interpolated CDF computations for all values of n. Similarly, we are able to generate population tolerance bands from adjusted LR saddlepoint CDF approximations for any value of n. In addition, note that in our plots we restrict ourselves to time $t \in [0.05, 3.0]$. The reason for this is that for values of t less than 0.05 the tolerance intervals are very often the trivial $[1, 1]$ interval. We restrict ourselves to t values less than 3.0 since the

$$-\ln(0.05) \simeq 3.0$$

is the 95th quantile on the true survival distribution.

Finally our decision to only consider exponentially distributed survival and censoring times follows what is commonly done in studies of the KM, PL and ACL survival function, see for instance Chen, et al. (1982), Chang and Cheng (1985) and Chang (1996). The argument for only considering exponentially distributed data was originally given in Chen, et al. (1982). The idea is that calculations under exponentially distributed survival and censoring times apply to general proportional hazards models after appropriate transformation of the data. If

$$S_Y(t) = S^{\beta}(t)$$

then

$$S(X) \sim U$$

where U is a standard uniform random variable and so

$$-\ln[1 - S(X)] \sim -\ln[S(X)]$$

has an exponential distribution with unit rate. From this, one finds that $-\ln[S(Y)]$ has an exponential distribution with rate β since

$$-\ln[S(Y)] = -\frac{1}{\beta}\ln[S_Y(Y)] \sim -\frac{1}{\beta}\ln[U].$$

With this in mind, Chen, et al. (1982) define

$$R^{-1}(t) = \inf\{z : -\ln[S(z)] > t\}$$

and note that $R(X)$ and $R(Y)$ are exponentially distributed with rate parameters 1 and β, respectively. They then define the resulting right-censored data

$$Z_{i,R} = \min\{R(X_i), R(Y_i)\}$$

$$\Delta_{i,R} = \Delta_i = I(R(X_i) \leqslant R(Y_i))$$

for $i = 1, \ldots, n$ and note that ordered samples

$$\left\{(R\left(Z_{(1)}\right), \Delta_{(1)}), (R\left(Z_{(2)}\right), \Delta_{(2)}), \ldots, (R\left(Z_{(n)}\right), \Delta_{(n)})\right\}$$

and

$$\left\{(Z_{(1),R}, \Delta_{(1),R}), (Z_{(2),R}, \Delta_{(2),R}), \ldots, (Z_{(n),R}, \Delta_{(n),R})\right\}$$

have the same joint distributions. Finally, they note that if we let $\hat{S}(t)_R$ denote a survival function estimator computed from the transformed sample

$$\left\{(Z_{(1),R}, \Delta_{(1),R}), (Z_{(2),R}, \Delta_{(2),R}), \ldots, (Z_{(n),R}, \Delta_{(n),R})\right\}$$

then processes

$$\left\{\hat{S}(R(t))_R : 0 \leq t < \infty\right\}$$

and

$$\left\{\hat{S}(t) : 0 \leq t < \infty\right\}$$

have the same distributions.

The Figures 4.1 - 4.12 show that the 95% two-sided symmetric population tolerance bounds for the KM, PL and ACL estimators obtained from the exact interpolated CDFs and from the adjusted LR saddlepoint CDF approximations are nearly identical for the KM and PL estimators when $n \leq 15$ and are also nearly identical for the ACL estimator for all values of n.

Each of the Figures 4.1 - 4.12 below show the 95% population tolerance bands for the given survival estimator, with $t \in (0.05, 3.0)$, and sample size $n = 5, 10, 15, 20, 25$ and 30 where the bands from the exact CDF, when available, are shown as black dotted curves and the bands from the adjusted LR saddlepoint CDF approximation are shown as gold solid curves.

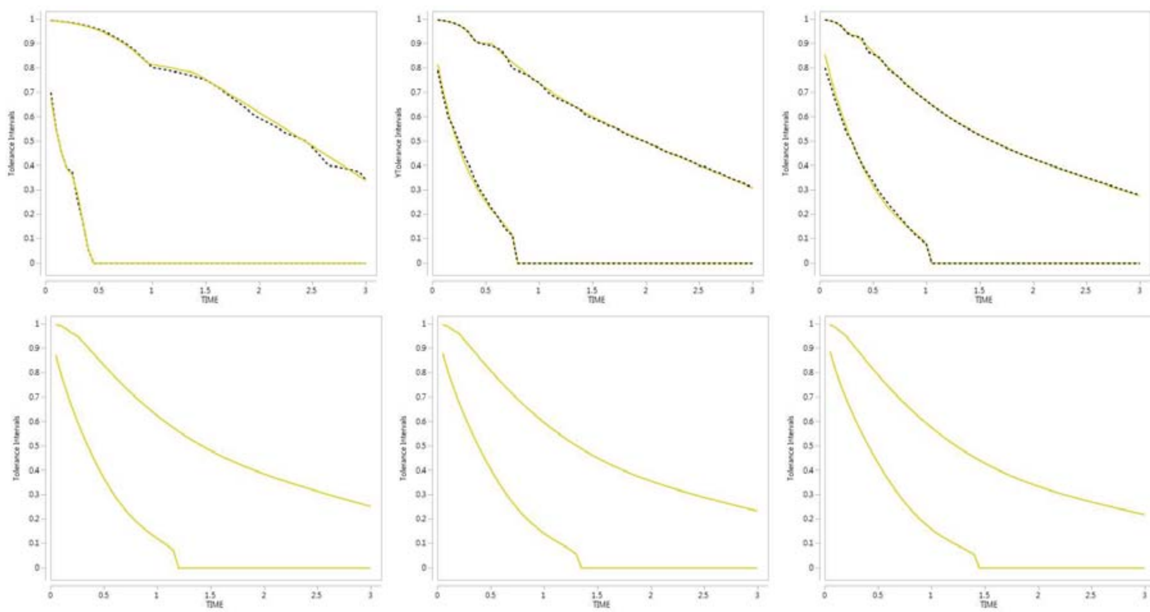

Figure 4.1. The 95% population tolerance bands for the KM estimator with $\beta = 0.5$.

4.2.2. Pointwise Bootstrap Confidence Bands.

Bootstrap confidence bands for $S(t)$ from the KM, PL and ACL estimators, with confidence $(1-\alpha)\,100\%$, are obtained by replacing γ with its $\hat{\gamma}$ and replacing $F_Z(t)$ by $\hat{F}_Z(t)$ and computing the 95% two-sided population tolerance interval for the survival estimators from either the from the exact interpolated CDFs or from the adjusted LR saddlepoint CDF approximation.

The performance of these confidence band methods is compared with that of the common competing confidence band methods in the next section.

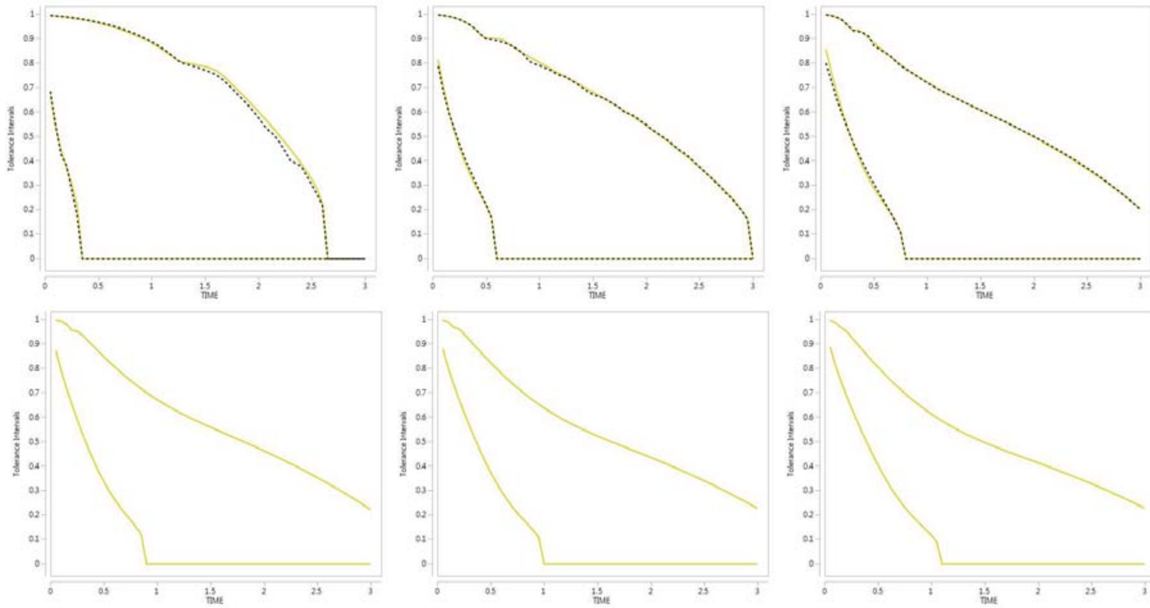

Figure 4.2. The 95% population tolerance bands for the KM estimator with $\beta = 1$.

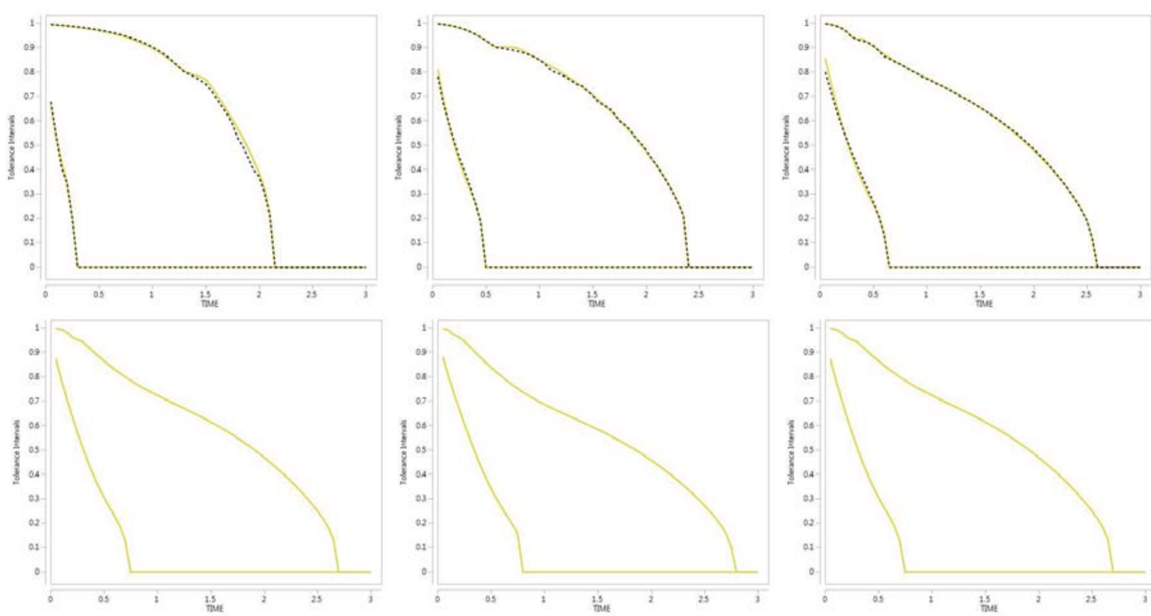

Figure 4.3. The 95% population tolerance bands for the KM estimator with $\beta = 1.5$.

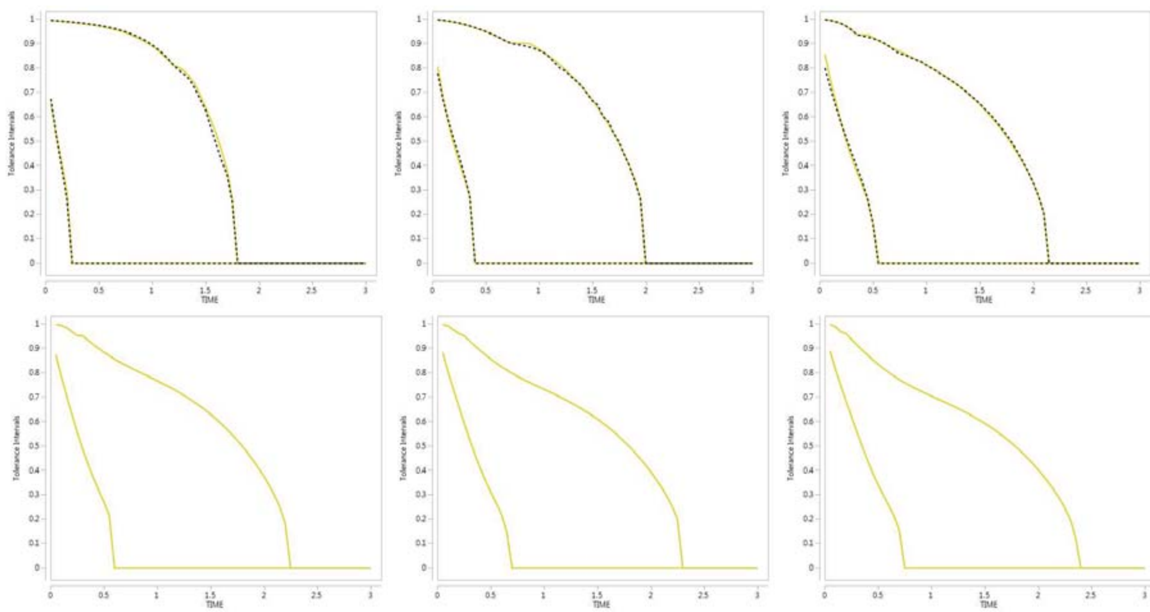

Figure 4.4. The 95% population tolerance bands for the KM estimator with $\beta = 2$.

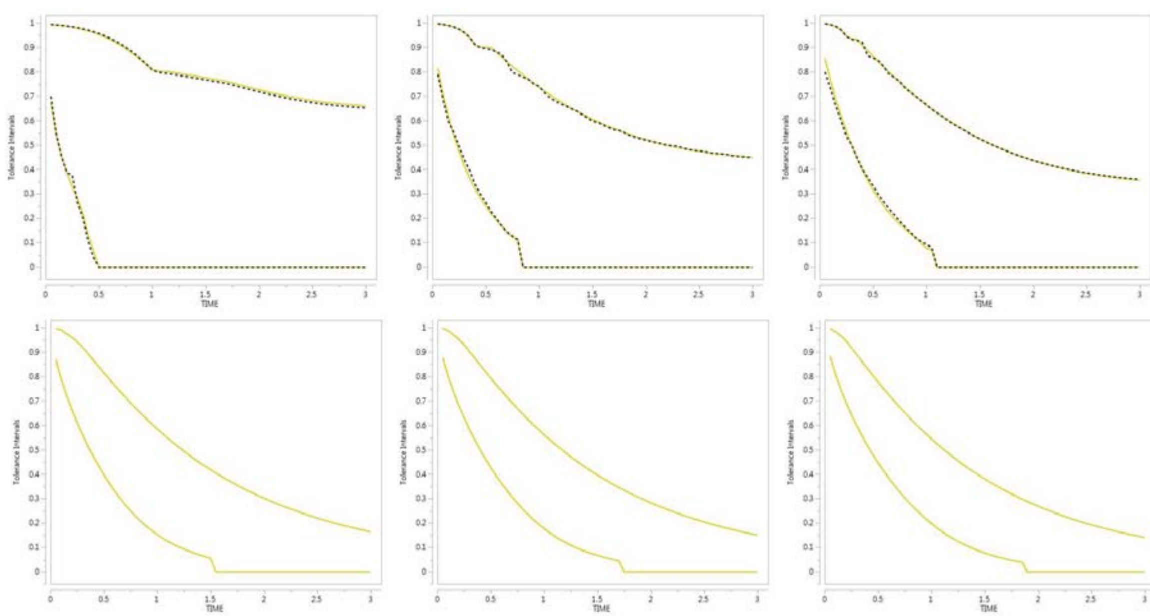

Figure 4.5. The 95% population tolerance bands for the PL estimator with $\beta = 0.5$.

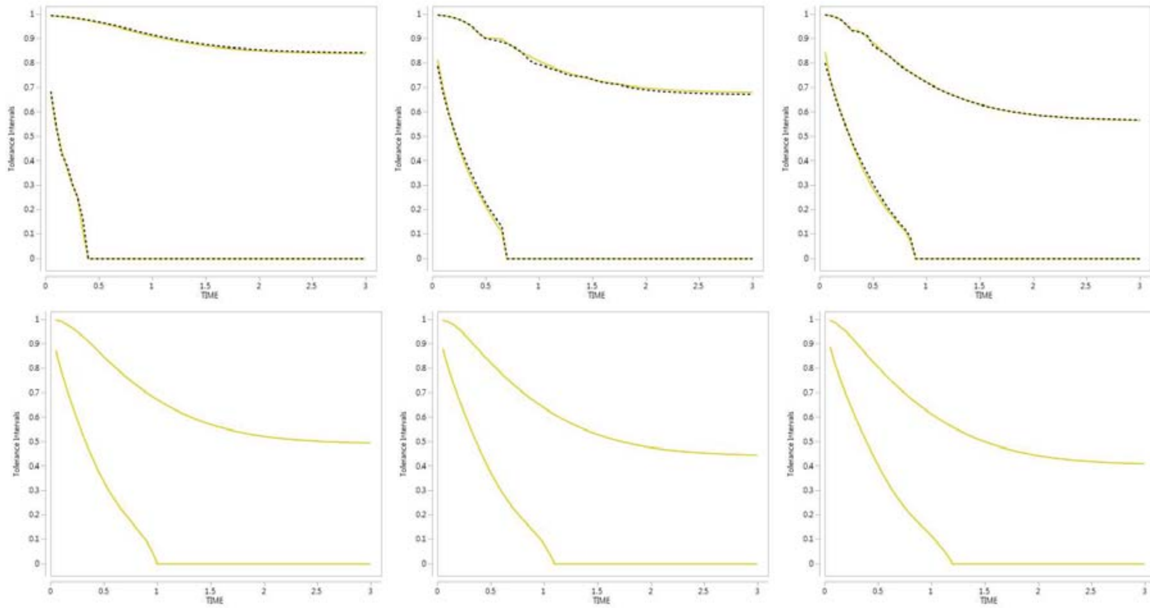

Figure 4.6. The 95% population tolerance bands for the PL estimator with $\beta = 1$.

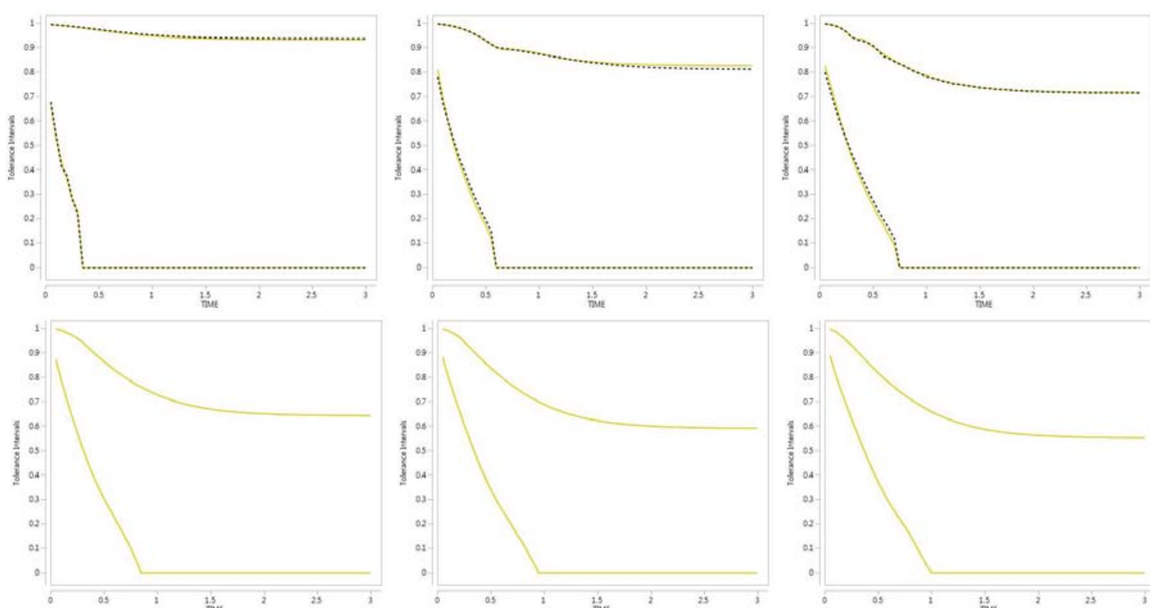

Figure 4.7. The 95% population tolerance bands for the PL estimator with $\beta = 1.5$.

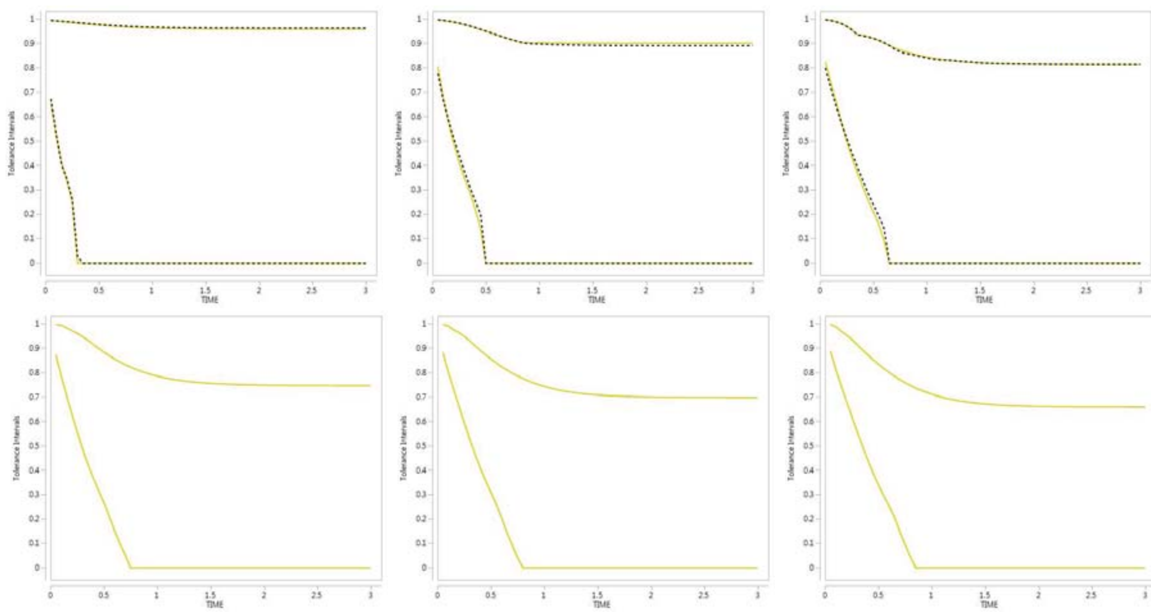

Figure 4.8. The 95% population tolerance bands for the PL estimator with $\beta = 2$.

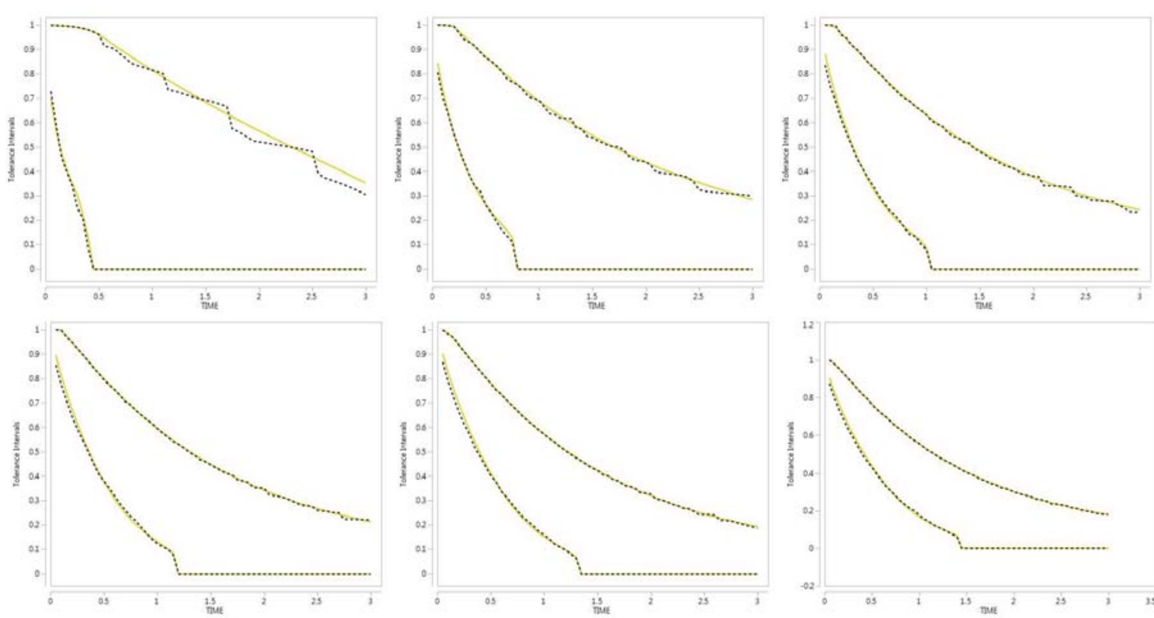

Figure 4.9. The 95% population tolerance bands for the ACL estimator with $\beta = 0.5$.

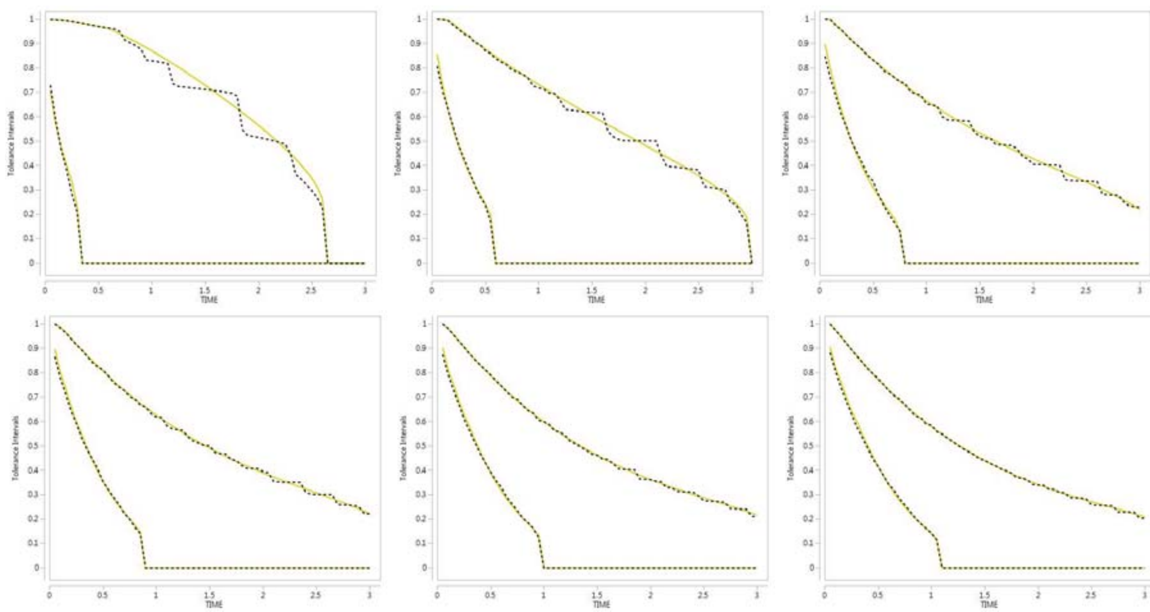

Figure 4.10. The 95% population tolerance bands for the ACL estimator with $\beta = 1$.

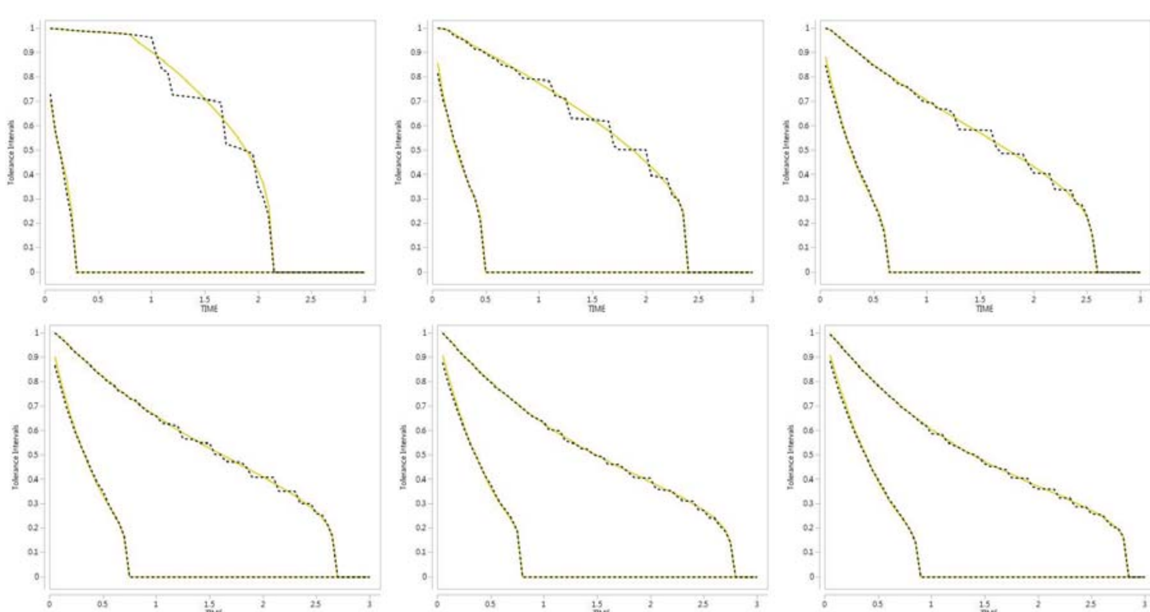

Figure 4.11. The 95% population tolerance bands for the ACL estimator with $\beta = 1.5$.

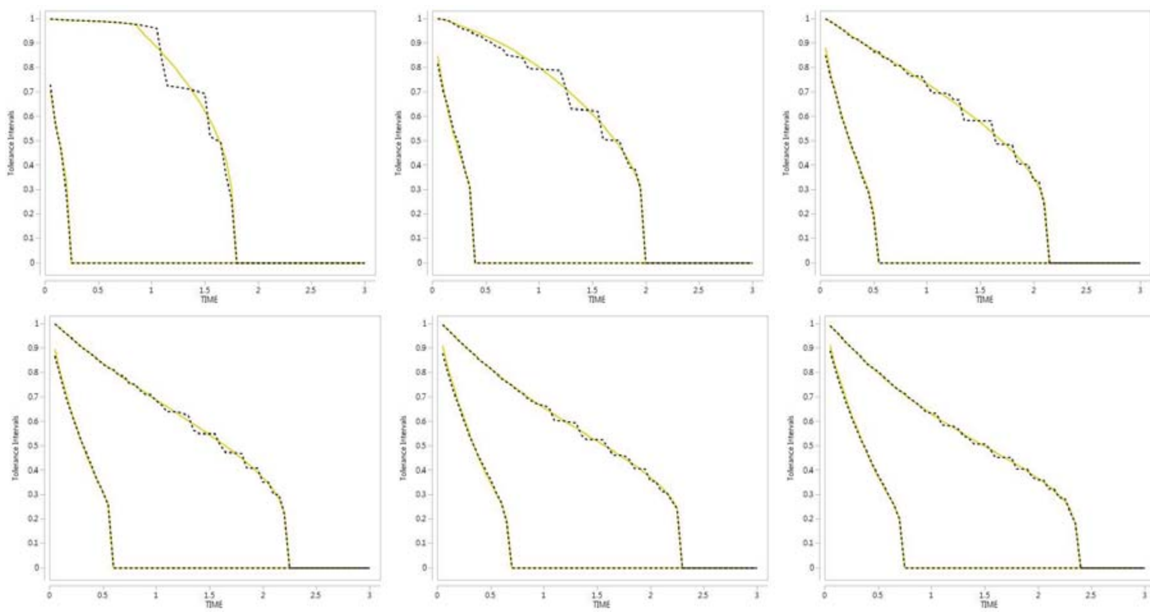

Figure 4.12. The 95% population tolerance bands for the ACL estimator with $\beta = 2$.

5. SIMULATION STUDIES

In this section, we perform simulation studies to compare the performance of our bootstrap confidence band methods for the KM, PL and ACL estimators, from the exact interpolated CDF or the adjusted LR saddlepoint CDF approximation, with that of traditional confidence band methods.

As was the case in section we restrict ourselves to exponentially distributed survival times, with unit rate, and exponential censoring times with rate $\beta = 0.5, 1.0, 1.5$ and 2.0. We also consider samples of size $n = 5, 10, 15, 20, 25$ and 30 but only consider exact interpolated CDF computations when $n \leq 15$ for the KM and PL estimators while we are able to do exact computations for the ACL estimator at all sample sizes. For each (β, n) combination we simulate $10,000$ random samples of n (X, Y) pairs, e.g.

$$(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n).$$

From each sample we compute 95% confidence bands for a survival estimator $\hat{S}(t)$ (KM, PL or ACL) from the exact interpolated CDF (when possible), the adjusted LR saddlepoint CDF approximation and a classical competing method and construct 95% confidence bands, from each method, for $S(t)$. Here we let time t range from 0.05 to 3.0 in steps of size 0.05 for a total of 60 confidence intervals in each confidence band. For each of the two or three confidence bands, we compute empirical coverage probabilities which are defined as

$$\hat{P}\left(\hat{S}_L(t) \leq S(t) \leq \hat{S}_U(t) | \hat{S}_L(t) \neq \hat{S}_U(t)\right) = \frac{\sum_{i=1}^{10,000} I\left(\hat{S}_L(t)_i \leq S(t) \leq \hat{S}_U(t)_i\right)}{\sum_{i=1}^{10,000} I\left(\hat{S}_L(t)_i \neq \hat{S}_U(t)_i\right)}$$

where $\left(\hat{S}_L(t)_i, \hat{S}_U(t)_i\right)$ denotes the confidence interval computed from the ith random at time t where $t = 0.05, 0.10, 0.15, 0.20, ...$ or 3.

The Figures 5.1 - 5.12 compare coverage probabilities for the various methods. In each graph, the coverage probabilities of a particular bootstrap confidence band method

(KM, PL or ACL) are presented where the adjusted LR saddlepoint CDF approximation coverage is shown as a gold solid curve and the exact distribution coverage (when available) is shown as a black dotted curve. Each graph also has a red dashed 95% reference line and coverage probabilities for the exponential Greenwood method (the comparator for the KM and PL estimators) or the large sample ACL confidence bands (the comparator for the ACL estimator) are shown as a green dash-dot-dotted curve.

In each instance we find that generally coverage probabilities from the adjusted LR saddlepoint CDF approximation are quite close to those computed from the exact distribution. In additional, both of these methods generally outperform the traditional confidence band methods by a wide margin.

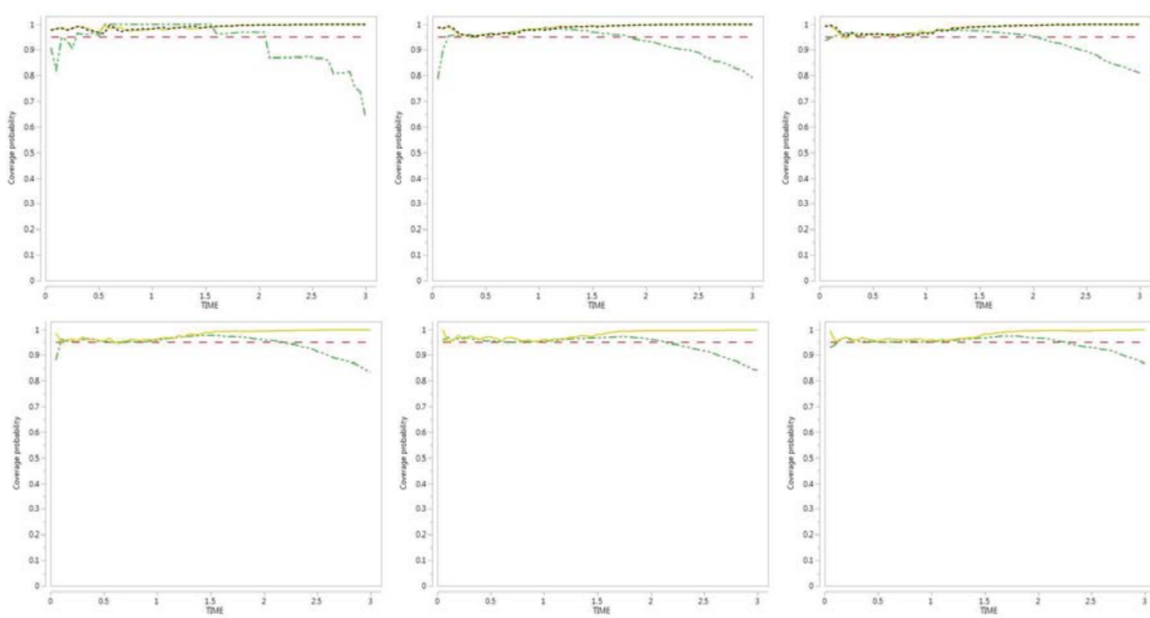

Figure 5.1. Coverage probabilities for various pointwise confidence band methods based on the KM estimator with $\beta = 0.5$.

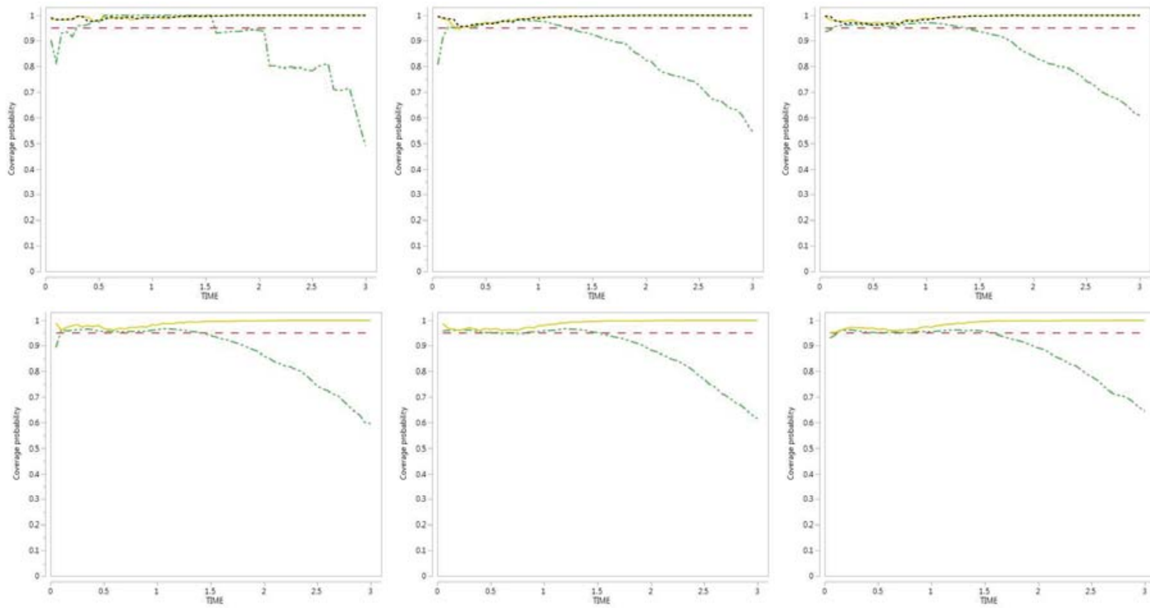

Figure 5.2. Coverage probabilities for various pointwise confidence band methods based on the KM estimator with $\beta = 1$.

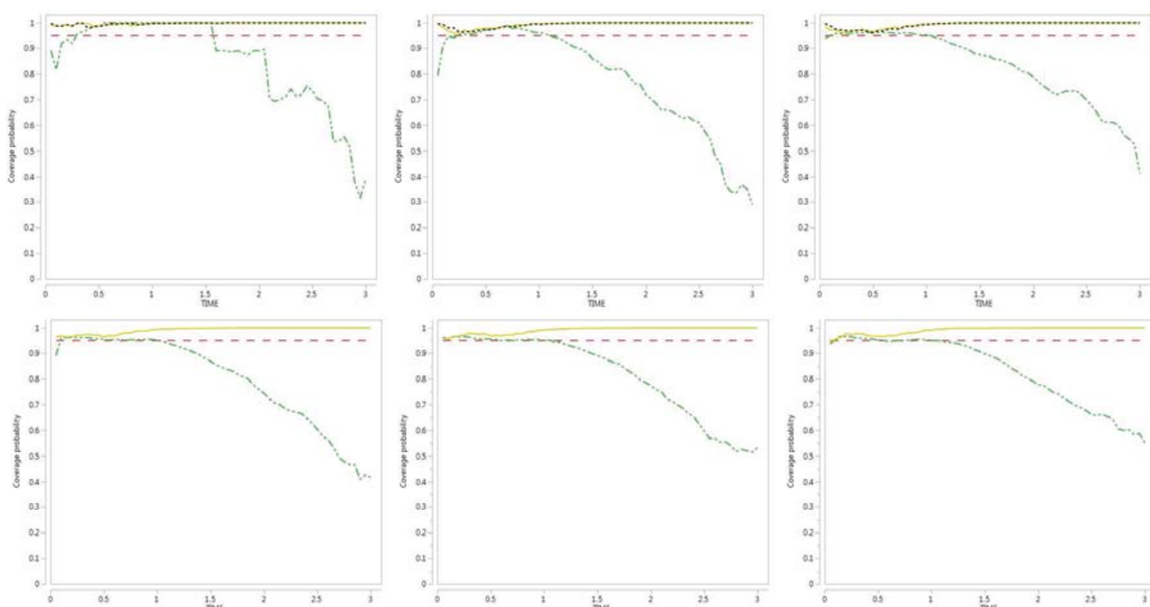

Figure 5.3. Coverage probabilities for various pointwise confidence band methods based on the KM estimator with $\beta = 1.5$.

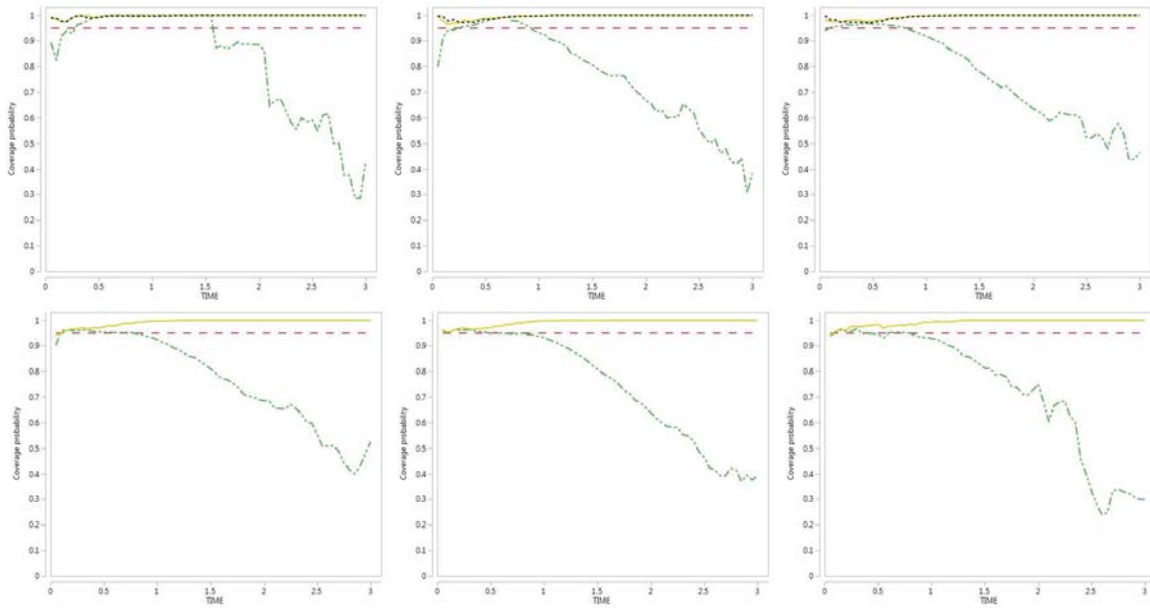

Figure 5.4. Coverage probabilities for various pointwise confidence band methods based on the KM estimator with $\beta = 2$.

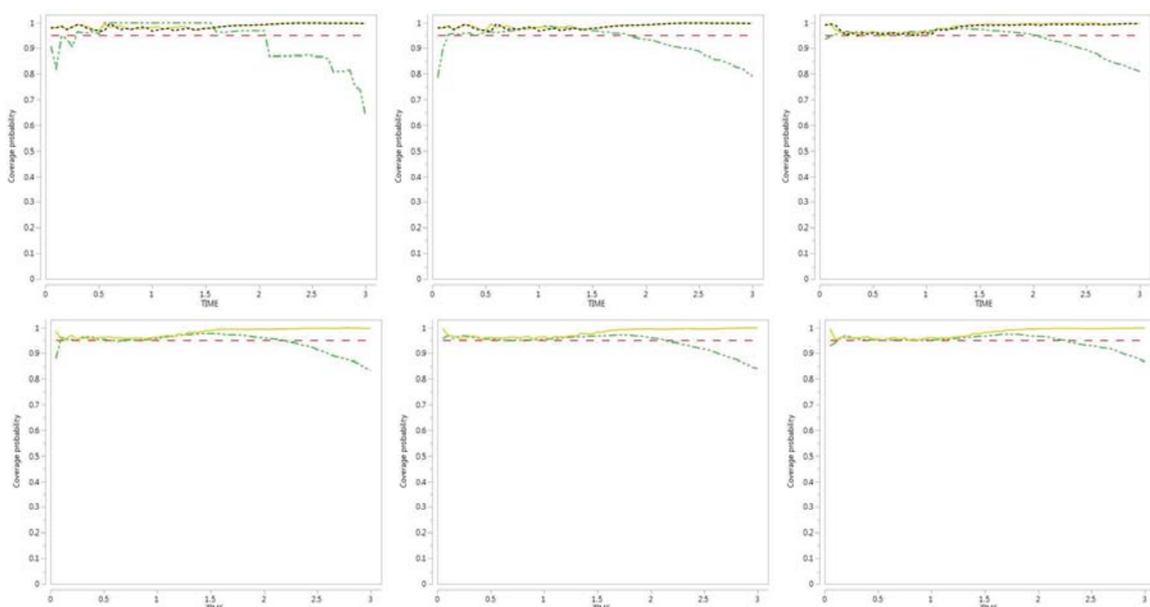

Figure 5.5. Coverage probabilities for various pointwise confidence band methods based on the PL estimator with $\beta = 0.5$.

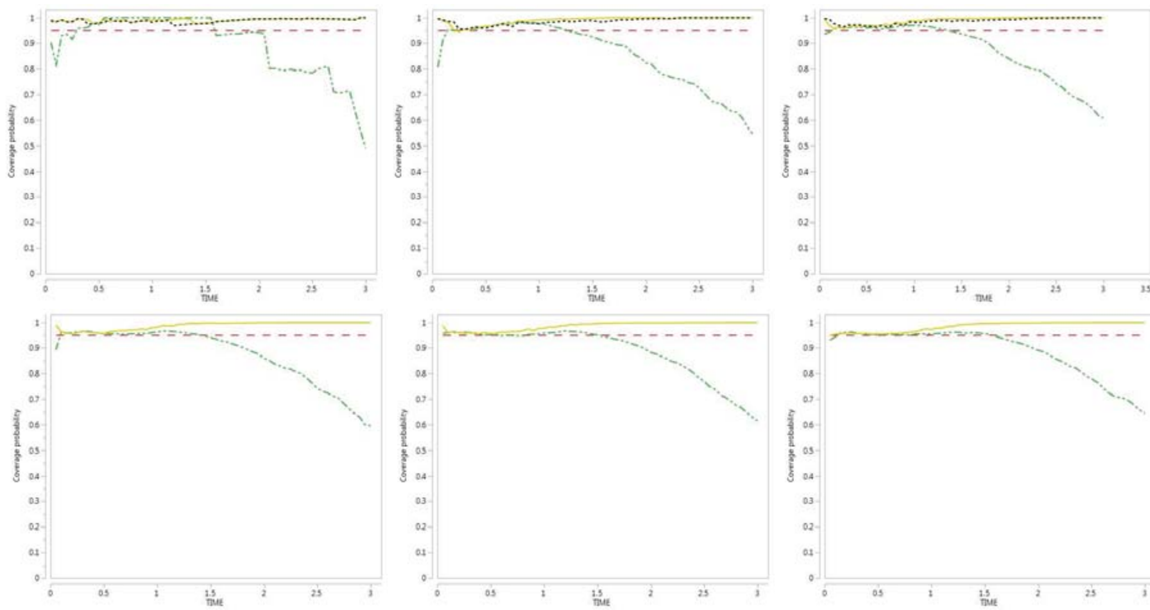

Figure 5.6. Coverage probabilities for various pointwise confidence band methods based on the PL estimator with $\beta = 1$.

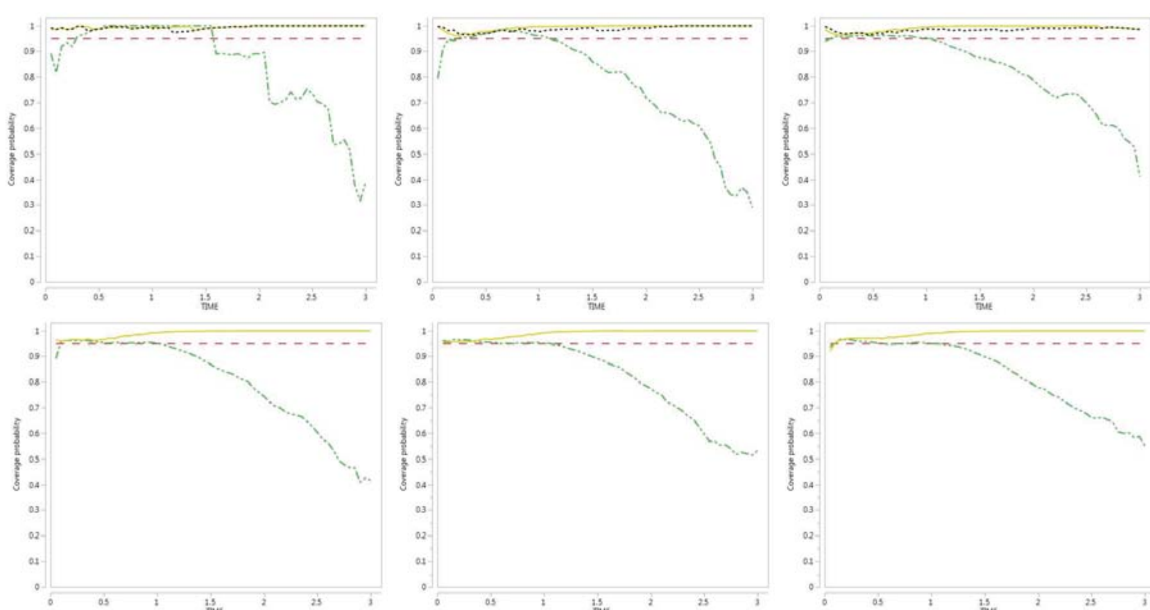

Figure 5.7. Coverage probabilities for various pointwise confidence band methods based on the PL estimator with $\beta = 1.5$.

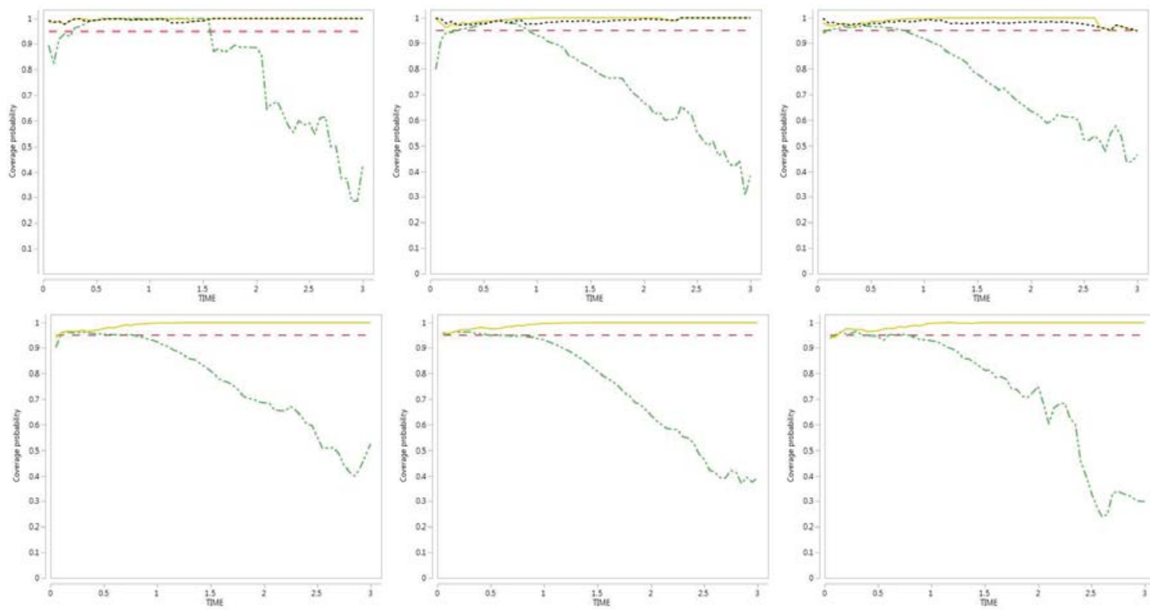

Figure 5.8. Coverage probabilities for various pointwise confidence band methods based on the PL estimator with $\beta = 2$.

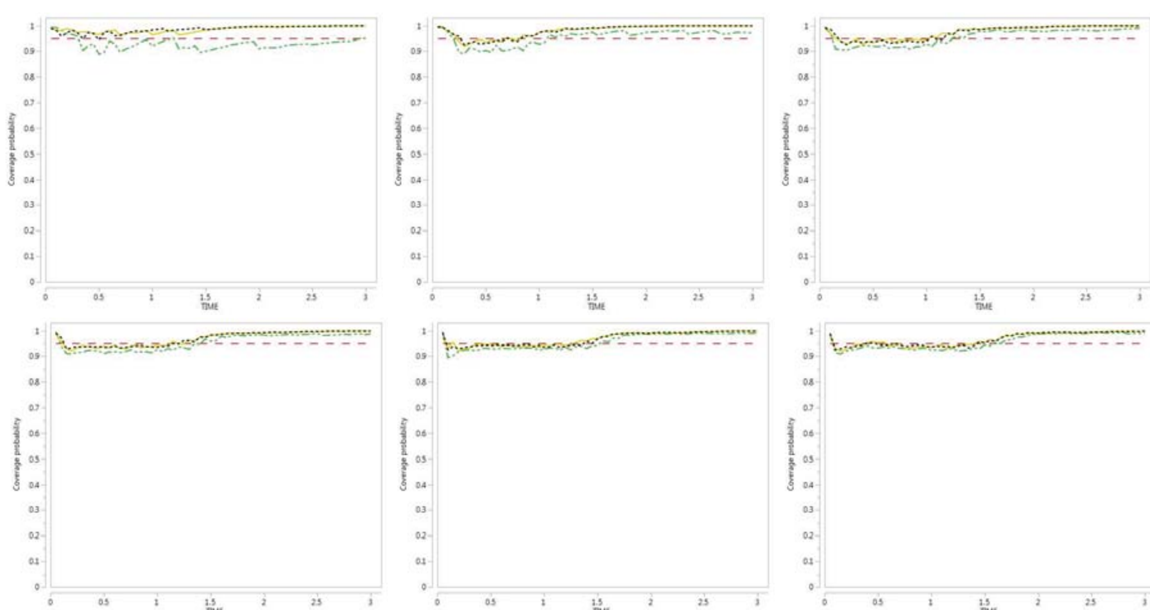

Figure 5.9. Coverage probabilities for various pointwise confidence band methods based on the ACL estimator with $\beta = 0.5$.

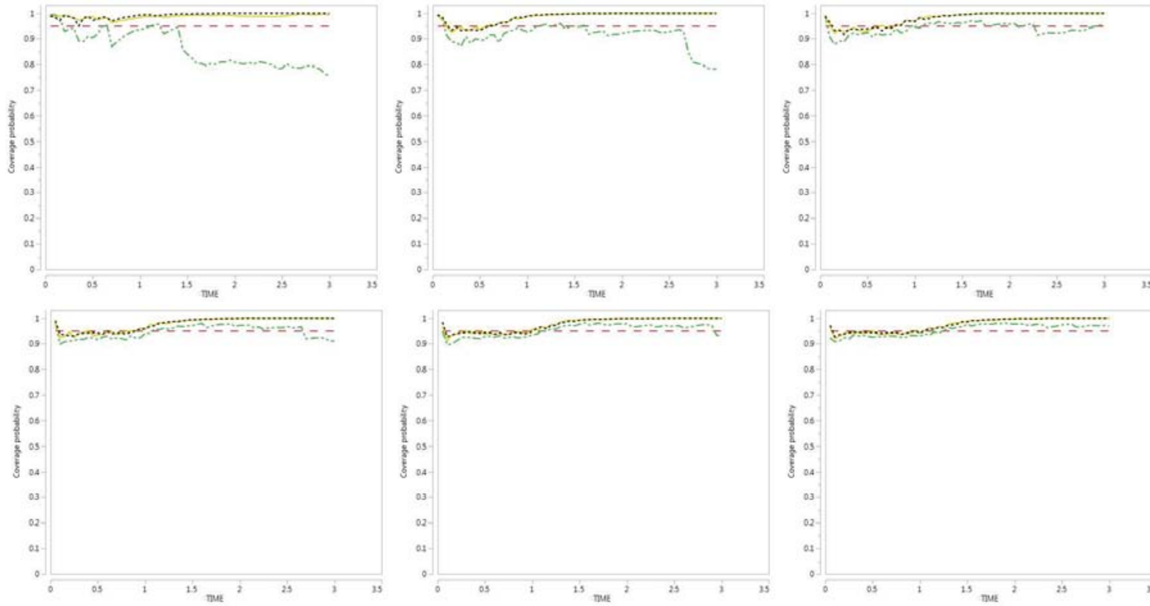

Figure 5.10. Coverage probabilities for various pointwise confidence band methods based on the ACL estimator with $\beta = 1$.

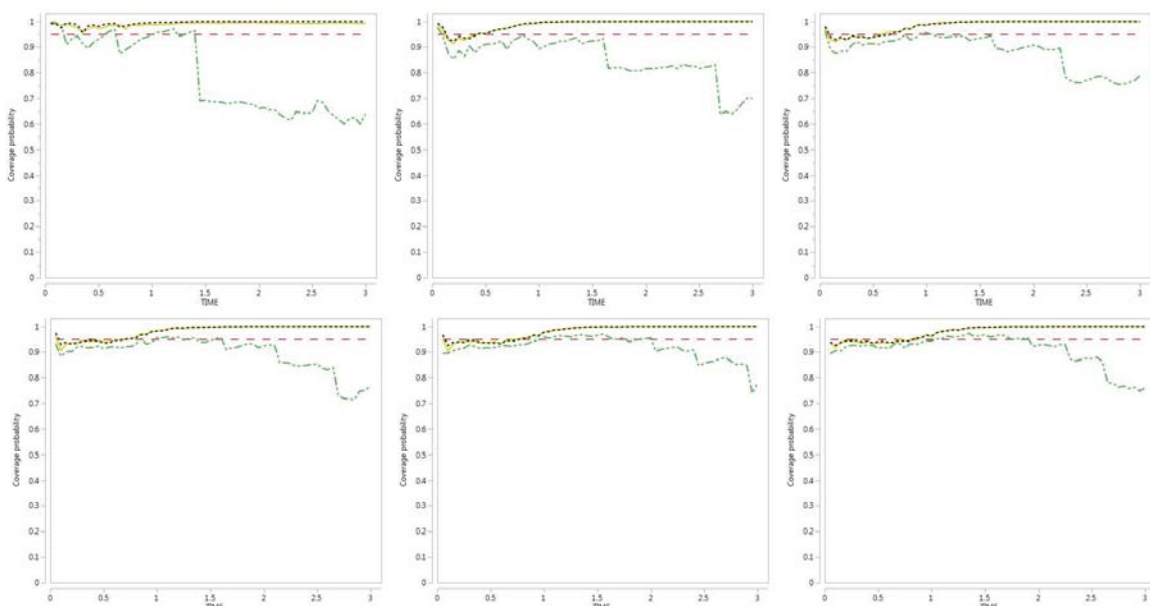

Figure 5.11. Coverage probabilities for various pointwise confidence band methods based on the ACL estimator with $\beta = 1.5$.

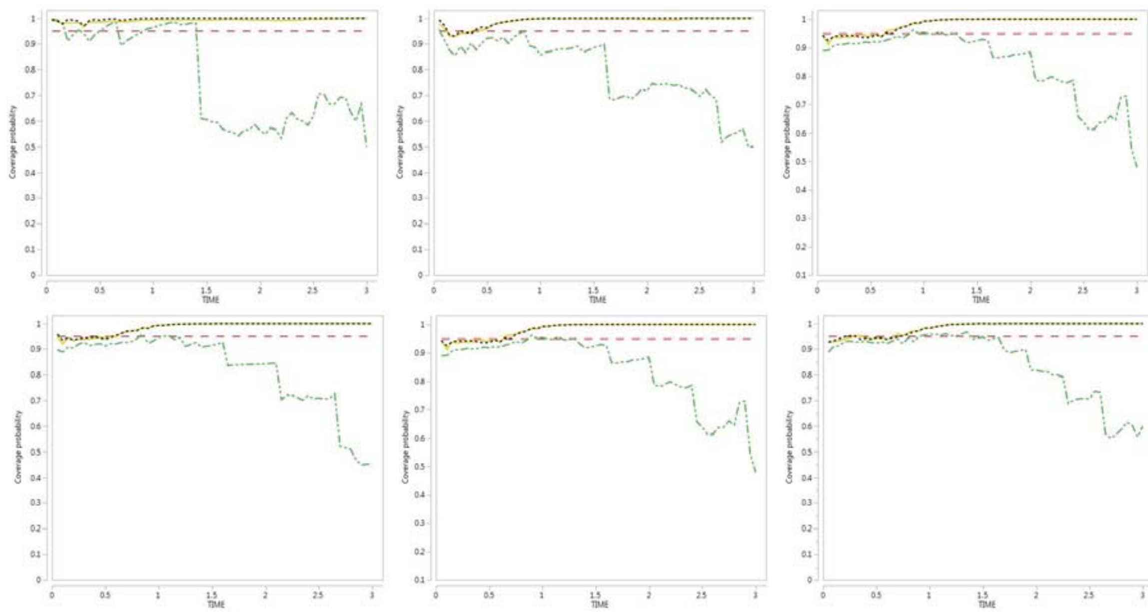

Figure 5.12. Coverage probabilities for various pointwise confidence band methods based on the ACL estimator with $\beta = 2$.

6. THE MEAN, BIAS, VARIANCE AND MSE FOR SURVIVAL ESTIMATORS

In this section we compare the exact mean, bias, variance and MSE of the KM, PL and ACL estimators, from their Mellin transforms, for $\beta = 0.5, 1, 1.5$ and 2, sample sizes $n = 5, 10, 15, 20, 25$ and 30, and times $t = 0.5, 1.0, 1.5, 2.0$ as is shown in Tables 6.1 - 6.4;

Table 6.1. The mean, bias, variance and MSE of the three survival estimators when $t = 0.5$.

		Mean, Bias, Variance and MSE for Survival Estimators with $t = 0.5$, $S(t) = 0.6065$											
		KM				PL				ACL			
β	n	$\mu(t)$	$-b(t)$	$\sigma^2(t)$	MSE	$\mu(t)$	$-b(t)$	$\sigma^2(t)$	MSE	$\mu(t)$	$-b(t)$	$\sigma^2(t)$	MSE
0.5	5	0.5784	0.0281	0.0653	0.0661	0.5835	0.0230	0.0617	0.0622	0.5946	0.0119	0.0517	0.0518
	10	0.6059	0.0007	0.0283	0.0283	0.6060	0.0005	0.0281	0.0281	0.6025	0.0040	0.0235	0.0235
	15	0.6065	0.0000	0.0185	0.0185	0.6065	0.0000	0.0185	0.0185	0.6040	0.0026	0.0154	0.0154
	20	0.6065	0.0000	0.0138	0.0138	0.6065	0.0000	0.0138	0.0138	0.6046	0.0019	0.0115	0.0115
	25	0.6065	0.0000	0.011	0.0110	0.6065	0.0000	0.0110	0.0110	0.6050	0.0015	0.0092	0.0092
	30	0.6065	0.0000	0.0092	0.0092	0.6065	0.0000	0.0092	0.0092	0.6053	0.0013	0.0077	0.0077
1	5	0.5776	0.0289	0.0843	0.0851	0.6025	0.0041	0.0696	0.0696	0.5745	0.0320	0.0701	0.0711
	10	0.6048	0.0017	0.0348	0.0348	0.6066	-0.0001	0.0334	0.0334	0.5988	0.0078	0.0279	0.0280
	15	0.6064	0.0001	0.0219	0.0219	0.6065	0.0000	0.0218	0.0218	0.6024	0.0041	0.0174	0.0174
	20	0.6065	0.0000	0.0162	0.0162	0.6065	0.0000	0.0162	0.0162	0.6036	0.0029	0.0128	0.0128
	25	0.6065	0.0000	0.0129	0.0129	0.6065	0.0000	0.0129	0.0129	0.6042	0.0023	0.0102	0.0102
	30	0.6065	0.0000	0.0107	0.0107	0.6065	0.0000	0.0107	0.0107	0.6046	0.0019	0.0085	0.0850
1.5	5	0.5483	0.0582	0.1110	0.1110	0.6120	-0.0055	0.0797	0.0797	0.5402	0.0664	0.0964	0.1008
	10	0.5981	0.0014	0.0461	0.0461	0.6072	-0.0006	0.0400	0.0400	0.6072	0.0161	0.0376	0.0379
	15	0.6052	0.0002	0.0271	0.0271	0.6066	-0.0001	0.0260	0.0260	0.5905	0.0067	0.0214	0.0214
	20	0.6063	0.0000	0.0194	0.0194	0.6065	0.0000	0.0192	0.0192	0.5999	0.0041	0.0151	0.0151
	25	0.6065	0.0000	0.0152	0.0152	0.6065	0.0000	0.0152	0.0152	0.6024	0.0030	0.0118	0.0118
	30	0.6065	0.0000	0.0126	0.0126	0.6065	0.0000	0.0126	0.0126	0.6035	0.0025	0.0098	0.0098
2	5	0.5014	0.1051	0.1371	0.1482	0.6210	-0.0144	0.0896	0.0898	0.4927	0.1138	0.1228	0.1358
	10	0.5815	0.025	0.0639	0.0645	0.6087	-0.0022	0.0478	0.0478	0.5728	0.0337	0.0540	0.0551
	15	0.6002	0.0063	0.0359	0.0359	0.6069	-0.0004	0.0313	0.0313	0.5934	0.0131	0.0288	0.0290
	20	0.6049	0.0017	0.0242	0.0242	0.6066	-0.0001	0.0229	0.0229	0.5997	0.0068	0.0189	0.0189
	25	0.6061	0.0004	0.0184	0.0184	0.6065	0.0000	0.0181	0.0181	0.6021	0.0044	0.0142	0.0142
	30	0.6064	0.0001	0.0150	0.0150	0.6065	0.0000	0.0149	0.0149	0.6032	0.0034	0.0115	0.0115

From the Tables 6.1 - 6.4 we see that for fixed β and n, the bias increases as t increases for all methods which is to be expected since it is harder to estimate the

Table 6.2. The mean, bias, variance and MSE of the three survival estimators when $t = 1$.

		Mean, Bias, Variance and MSE for Survival Estimators with $t = 1$, $S(t) = 0.3679$											
		KM				PL				ACL			
β	n	$\mu(t)$	$-b(t)$	$\sigma^2(t)$	MSE	$\mu(t)$	$-b(t)$	$\sigma^2(t)$	MSE	$\mu(t)$	$-b(t)$	$\sigma^2(t)$	MSE
0.5	5	0.3401	0.0278	0.0758	0.0766	0.3754	−0.0075	0.0665	0.0666	0.3335	0.0344	0.0662	0.0674
	10	0.3625	0.0054	0.0361	0.0361	0.3688	−0.0009	0.0334	0.0334	0.3568	0.0111	0.0313	0.0314
	15	0.3667	0.0012	0.0226	0.0226	0.3680	−0.0002	0.0219	0.021 9	0.3626	0.0053	0.0197	0.0197
	20	0.3676	0.0003	0.0164	0.016 4	0.3679	0.0000	0.0163	0.0163	0.3645	0.0033	0.0143	0.0 43
	25	0.3678	0.0001	0.013	0.0130	0.3679	0.0000	0.0129	0.0129	0.3654	0.0024	0.0113	0.0113
	30	0.3679	0.0000	0.0107	0.0107	0.3679	0.0000	0.0107	0.0107	0.3659	0.0019	0.0094	0.0940
1	5	0.2798	0.0881	0.0990	0.106 8	0.3987	−0.0308	0.0881	0.0890	0.2727	0.0952	0.0873	0.0964
	10	0.3340	0.0339	0.0589	0.0600	0.3752	−0.0073	0.0474	0.0475	0.3258	0.0421	0.0503	0.0521
	15	0.3538	0.0141	0.0378	0.0380	0.3701	−0.0022	0.0318	0.0318	0.3462	0.0217	0.0315	0.0320
	20	0.3618	0.0061	0.0265	0.0265	0.3686	−0.0008	0.0236	0.0236	0.3552	0.0127	0.0218	0.0220
	25	0.3652	0.0027	0.0200	0.0200	0.3682	−0.0003	0.0186	0.018 6	0.3597	0.0082	0.0163	0.0164
	30	0.3667	0.0012	0.0160	0.0160	0.3680	−0.0001	0.0154	0.015 4	0.3620	0.0058	0.0130	0.0130
1.5	5	0.2087	0.1592	0.1006	0.125 9	0.4332	−0.0653	0.1068	0.111 06	0.2035	0.1644	0.0907	0.117 7
	10	0.2790	0.0889	0.0794	0.0873	0.3911	−0.0232	0.0632	0.0637	0.2711	0.0968	0.0691	0.078 5
	15	0.3157	0.0522	0.0596	0.0623	0.3781	−0.0102	0.0445	0.0446	0.3067	0.0612	0.0503	0.0540
	20	0.3365	0.0314	0.0450	0.0460	0.3728	−0.0049	0.0340	0.0340	0.3273	0.0405	0.0370	0.038 6
	25	0.3487	0.0191	0.0348	0.0352	0.3704	−0.0025	0.0273	0.0273	0.3400	0.0279	0.0280	0.0288
	30	0.3561	0.0118	0.0276	0.0277	0.3692	−0.0013	0.0226	0.0226	0.3479	0.0199	0.0218	0.0222
2	5	0.1458	0.2221	0.0859	0.135 2	0.4731	−0.1052	0.1206	0.131 7	0.1424	0.2255	0.0787	0.129 6
	10	0.2130	0.1549	0.0852	0.109 2	0.4164	−0.0485	0.0779	0.0803	0.2069	0.1609	0.0756	0.101 5
	15	0.2565	0.1114	0.0755	0.0879	0.3946	−0.0268	0.0580	0.0587	0.2484	0.1195	0.0651	0.079 4
	20	0.2865	0.0814	0.0648	0.0714	0.3839	−0.0160	0.0462	0.0465	0.2772	0.0907	0.0545	0.0627
	25	0.3078	0.0601	0.0550	0.0586	0.3779	−0.0100	0.0382	0.038 3	0.2979	0.0700	0.0452	0.050 1
	30	0.3232	0.0447	0.0465	0.0485	0.3744	−0.0065	0.0324	0.0324	0.3131	0.0548	0.0375	0.0405

tails of the survival function than for smaller values of t. The bias and variance also increases, for fixed t and n, as β the increases since larger values of β correspond to greater amounts of censoring. Finally, we see that the PL estimator has smaller bias than the KM or ACL estimators for all values of t, β and n.

Table 6.3. The mean, bias, variance and MSE of the three survival estimators when $t = 1.5$.

		Mean, Bias, Variance and MSE for Survival Estimators with $t = 1.5$, $S(t) = 0.2231$											
		KM				PL				ACL			
β	n	$\mu(t)$	$-b(t)$	$\sigma^2(t)$	MSE	$\mu(t)$	$-b(t)$	$\sigma^2(t)$	MSE	$\mu(t)$	$-b(t)$	$\sigma^2(t)$	MSE
0.5	5	0.1783	0.0448	0.0578	0.0598	0.2499	−0.0267	0.0596	0.0603	0.1737	0.0494	0.0508	0.053 2
	10	0.2040	0.0191	0.0340	0.0344	0.2301	−0.0070	0.0303	0.0303	0.1987	0.0244	0.0293	0.0299
	15	0.2141	0.0090	0.0227	0.0228	0.2256	−0.0024	0.0202	0.0202	0.2090	0.0141	0.0194	0.0196
	20	0.2187	0.0045	0.0165	0.016 5	0.2241	−0.0010	0.0151	0.0151	0.2141	0.0090	0.0141	0.0142
	25	0.2209	0.0023	0.0127	0.0127	0.2235	−0.0004	0.0120	0.0120	0.2169	0.0063	0.0109	0.0109
	30	0.2219	0.0012	0.0103	0.01030	0.2233	−0.0002	0.0099	0.0099	0.2185	0.0046	0.0088	0.0882
1	5	0.1147	0.1084	0.0561	0.0679	0.3054	−0.0823	0.0880	0.0948	0.1114	0.1117	0.0499	0.0624
	10	0.1521	0.0710	0.0466	0.0516	0.2579	−0.0347	0.0488	0.0500	0.1469	0.0762	0.0399	0.0457
	15	0.1741	0.0490	0.0380	0.04040	0.2413	−0.0182	0.0340	0.0343	0.1676	0.0555	0.0316	0.034 7
	20	0.1884	0.0347	0.0311	0.0323	0.2336	−0.0104	0.0261	0.0262	0.1813	0.0418	0.0253	0.0270
	25	0.1982	0.0250	0.0257	2. 632	0.2295	−0.0064	0.0211	0.0211	0.1908	0.0323	0.0206	0.0216
	30	0.2050	0.0181	0.0216	0.0220	0.2272	−0.0040	0.0176	0.0176	0.1976	0.0255	0.0171	0.0178
1.5	5	0.0652	0.1580	0.0398	0.0648	0.3710	−0.1479	0.1112	0.133 1	0.0634	0.1597	0.0360	0.0615
	10	0.0958	0.1274	0.0421	0.0583	0.3040	−0.0809	0.0684	0.0749	0.0925	0.1307	0.0369	0.0540
	15	0.1177	0.1055	0.0410	0.052 1	0.2755	−0.0523	0.0503	0.0530	0.1129	0.1102	0.0347	0.0468
	20	0.1346	0.0886	0.0387	0.0466	0.2597	−0.0365	0.0402	0.0415	0.1287	0.0944	0.0320	0.0409
	25	0.1481	0.0750	0.0361	0.0417	0.2498	−0.0267	0.0335	0.0342	0.1413	0.0818	0.0293	0.0360
	30	0.1592	0.0639	0.0335	0.0376	0.2432	−0.0201	0.0288	0.0292	0.1518	0.0714	0.0267	0.03180
2	5	0.0346	0.1885	0.0239	0.0594	0.4341	−0.2110	0.1263	0.170 8	0.0337	0.1894	0.0220	0.0579
	10	0.0543	0.1688	0.0295	0.0580	0.3575	−0.1343	0.0847	0.102 7	0.0525	0.1706	0.0263	0.05549
	15	0.0699	0.1533	0.0322	0.0557	0.3213	−0.0981	0.0657	0.0753	0.0672	0.1560	0.0278	0.05219
	20	0.0830	0.1401	0.0335	0.0531	0.2994	−0.0763	0.0544	0.0602	0.0794	0.1437	0.0283	0.0490
	25	0.0944	0.1287	0.0340	0.0506	0.2846	−0.0615	0.0468	0.0506	0.0900	0.1331	0.0282	0.04599
	30	0.1045	0.1186	0.0340	0.0481	0.2739	−0.0507	0.0412	0.043 8	0.0994	0.1237	0.0278	0.04319

Table 6.4. The mean, bias, variance and MSE of the three survival estimators when $t = 2$.

		Mean, Bias, Variance and MSE for Survival Estimators with $t = 2$, $S(t) = 0.1353$											
		KM				PL				ACL			
β	n	$\mu(t)$	$-b(t)$	$\sigma^2(t)$	MSE	$\mu(t)$	$-b(t)$	$\sigma^2(t)$	MSE	$\mu(t)$	$-b(t)$	$\sigma^2(t)$	MSE
0.5	5	0.0888	0.0466	0.0343	0.0365	0.1855	−0.0501	0.0516	0.0541	0.0862	0.0491	0.0302	0.0326
	10	0.1069	0.0285	0.0238	0.0246	0.1546	−0.0193	0.0254	0.0258	0.1033	0.0320	0.0203	0.0213
	15	0.1165	0.0188	0.0178	0.0182	0.1449	−0.0095	0.0169	0.0170	0.1125	0.0229	0.0149	0.0154
	20	0.1225	0.0129	0.0140	0.0142	0.1406	−0.0053	0.0127	0.0127	0.1182	0.0172	0.0116	0.0119
	25	0.1263	0.0090	0.0113	0.0114	0.1384	−0.0031	0.0101	0.0101	0.1220	0.0133	0.0093	0.0095
	30	0.1289	0.0064	0.0094	0.0094	0.1373	−0.0019	0.0084	0.0084	0.1248	0.0106	0.0077	0.0078
1	5	0.0440	0.0913	0.0242	0.0325	0.2684	−0.1330	0.0859	0.1036	0.0427	0.0927	0.0215	0.0301
	10	0.0611	0.0742	0.0232	0.0287	0.2076	−0.0723	0.0470	0.0522	0.0588	0.0766	0.0199	0.0258
	15	0.0731	0.0623	0.0216	0.0255	0.1825	−0.0472	0.0327	0.0349	0.0698	0.0656	0.0179	0.0222
	20	0.0822	0.0532	0.0200	0.0228	0.1688	−0.0335	0.0252	0.0263	0.0781	0.0572	0.0162	0.0195
	25	0.0895	0.0458	0.0185	0.0206	0.1602	−0.0249	0.0205	0.0211	0.0849	0.0505	0.0146	0.0172
	30	0.0955	0.0398	0.0171	0.0187	0.1544	−0.0191	0.0174	0.0178	0.0904	0.0449	0.0133	0.0153
1.5	5	0.0191	0.1162	0.0125	0.0260	0.3521	−0.2168	0.1119	0.1589	0.0186	0.1167	0.0113	0.0249
	10	0.0290	0.1063	0.0145	0.0258	0.2759	−0.1406	0.0690	0.0888	0.0279	0.1074	0.0127	0.0242
	15	0.0367	0.0987	0.0154	0.0251	0.2404	−0.1051	0.0511	0.0621	0.0351	0.1003	0.0130	0.0231
	20	0.0431	0.0922	0.0158	0.0243	0.2191	−0.0837	0.0410	0.0480	0.0410	0.0943	0.0131	0.0220
	25	0.0488	0.0866	0.0160	0.0235	0.2045	−0.0692	0.0346	0.0394	0.0462	0.0892	0.0130	0.0210
	30	0.0538	0.0815	0.0161	0.0227	0.1939	−0.0586	0.0300	0.0334	0.0508	0.0846	0.0128	0.0200
2	5	0.0078	0.1275	0.0056	0.0219	0.4251	−0.2898	0.1274	0.2114	0.0076	0.1277	0.0052	0.0215
	10	0.0125	0.1228	0.0073	0.0224	0.3432	−0.2079	0.0861	0.1293	0.0121	0.1233	0.0065	0.0217
	15	0.0163	0.1190	0.0083	0.0225	0.3027	−0.1674	0.0672	0.0952	0.0157	0.1196	0.0072	0.0215
	20	0.0197	0.1156	0.0091	0.0225	0.2772	−0.1419	0.0560	0.0761	0.0188	0.1165	0.0077	0.0213
	25	0.0228	0.1125	0.0097	0.0224	0.2591	−0.1238	0.0485	0.0638	0.0217	0.1137	0.0081	0.0210
	30	0.0257	0.1097	0.0102	0.0222	0.2455	0.1101	0.0431	0.0552	0.0243	0.1110	0.0083	0.0206

7. CONCLUSIONS

We developed a saddlepoint-based method for generating small sample confidence bands for the $S(t)$ from KM, PL and ACL survival function estimators, under the proportional hazards model. As part of this development we derived the exact distribution of these estimators and developed mid-p population tolerance bands for said estimators. In our simulation studies, for the KM, PL and ACL estimators we compared our saddlepoint confidence bands with those obtained from competing large sample methods as well as those obtained from the exact distributions. We found that the saddlepoint confidence bands are very close to the confidence bands derived from the exact distribution, while being much easier to compute, and outperform the competing large sample methods in terms of coverage probability.

APPENDIX A

DERIVATIVES OF THE CGF FOR THE LOGARITM OF ZERO-TRUNCATED KM ESTIMATOR

The Mellin transform for the zero-truncated KM estimate, $KM+$, is given as

$$\mathcal{M}_{Tr}^{KM+}(v) = \frac{\sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) \prod_{i=1}^{r} \left[\gamma c_{in}^{v} + (1-\gamma)\right]}{1 - \left[F_Z(t)\right]^n}.$$

The CGF of $\ln(KM+)$ is given as

$$\begin{aligned} K_{\ln(KM+)}(v) &= \ln\left[\mathcal{M}_{Tr}^{KM+}(v)\right] \\ &= \ln\left\{\sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) g_r(v)\right\} \\ &- \ln\left\{1 - \left[F_Z(t)\right]^n\right\} \end{aligned}$$

where

$$g_r(v) = \prod_{i=1}^{r} \left[\gamma c_{in}^{v} + (1-\gamma)\right]$$

The first derivative of the CGF is given as

$$\begin{aligned} K'_{\ln(KM+)}(v) &= \frac{\sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) g'_r(v)}{\sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) g_r(v)} = \frac{B'(v)}{B(v)} \\ &= \frac{\left\{1 - \left[F_Z(t)\right]^n\right\} \mathcal{M}_{Tr}^{KM+}(v)'}{\left\{1 - \left[F_Z(t)\right]^n\right\} \mathcal{M}_{Tr}^{KM+}(v)} \end{aligned}$$

where $\mathcal{M}_{Tr}^{KM+}(v)'$ denotes the first derivative of Mellin transform $\mathcal{M}_{Tr}^{KM+}(v)$,

$$g'_r(v) = g_r(v) h_r(v)$$

and

$$h_r(v) = \sum_{i=1}^{r} \frac{\gamma c_{in}^{v} \ln(c_{in})}{\gamma c_{in}^{v} + (1-\gamma)}.$$

The second derivative of the CGF is

$$K''_{\ln(KM+)}(v) = \frac{B(v)B''(v) - [B'(v)]^2}{[B(v)]^2}$$

where

$$B''(v) = \sum_{r=0}^{n-1} b(r, F_Z(t))\, g''_r(v),$$

$$g''_r(v) = g(v)h'(v) + g'(v)h(v)$$

and

$$h'(v) = \sum_{i=1}^{r} \left\{ \frac{\gamma c_{in}^v [\ln(c_{in})]^2}{\gamma c_{in}^v + (1-\gamma)} - \left[\frac{\gamma c_{in}^v \ln(c_{in})}{\gamma c_{in}^v + (1-\gamma)} \right]^2 \right\}.$$

These results follow from the fact that

$$g_r(v) = \prod_{i=1}^{r} [\gamma c_{in}^v + (1-\gamma)]$$

so that

$$\ln[g_r(v)] = \ln\left\{ \prod_{i=1}^{r} [\gamma c_{in}^v + (1-\gamma)] \right\} = \sum_{i=1}^{r} \ln[\gamma c_{in}^v + (1-\gamma)].$$

Let

$$h_r(v) = \frac{d \ln[g_r(v)]}{dv} = \frac{g'_r(v)}{g_r(v)} = \sum_{i=1}^{r} \frac{\gamma c_{in}^v \ln(c_{in})}{\gamma c_{in}^v + (1-\gamma)} = \sum_{i=1}^{r} j_i(v)$$

which means that

$$\begin{aligned} g'_r(v) &= g_r(v) \sum_{i=1}^{r} \frac{\gamma c_{in}^v \ln(c_{in})}{\gamma c_{in}^v + (1-\gamma)} \\ &= \left\{ \prod_{i=1}^{r} [\gamma c_{in}^v + (1-\gamma)] \right\} \sum_{i=1}^{r} \frac{\gamma c_{in}^v \ln(c_{in})}{\gamma c_{in}^v + (1-\gamma)} = g_r(v)\, h_r(v). \end{aligned}$$

Note that

$$\ln\left[j_i\left(v\right)\right] = \ln\left[\frac{\gamma c_{in}^v \ln\left(c_{in}\right)}{\gamma c_{in}^v + \left(1-\gamma\right)}\right]$$
$$= \ln\left(\gamma\right) + v\ln\left(c_{in}\right) + \ln\left[\ln\left(c_{in}\right)\right] - \ln\left[\gamma c_{in}^v + \left(1-\gamma\right)\right]$$

so then

$$\frac{j_i'\left(v\right)}{j_i\left(v\right)} = \ln\left(c_{in}\right) - \frac{\gamma c_{in}^v \ln\left(c_{in}\right)}{\gamma c_{in}^v + \left(1-\gamma\right)}$$

which means that

$$j_i'\left(v\right) = j_i\left(v\right)\left[\ln\left(c_{in}\right) - j_i\left(v\right)\right].$$

Also since

$$g_r'(v) = g_r(v)h_r\left(v\right)$$

then

$$\begin{aligned} g_r''(v) &= g_r(v)h_r'\left(v\right) + g_r'(v)h_r\left(v\right) \\ &= g_r(v)h_r'\left(v\right) + g_r(v)\left[h_r\left(v\right)\right]^2 \\ &= g_r(v)\left[h_r'\left(v\right) + h_r^2\left(v\right)\right]. \end{aligned}$$

APPENDIX B

DERIVATIVES OF THE CGF FOR THE LOGARITM OF ZERO-TRUNCATED PL ESTIMATOR

The Mellin transform for the zero-truncated PL estimate, $PL+$, is given as

$$\begin{aligned}\mathcal{M}_{Tr}^{PL+}(v) &= \frac{E\left[\hat{S}_{KM}^{v}(t)\right] + (1-\gamma)\left[F_Z(t)\right]^n \prod_{i=1}^{n-1}\left[\gamma c_{in}^{v} + (1-\gamma)\right]}{1-\gamma\left[F_Z(t)\right]^n} \\ &= \frac{1-\left[F_Z(t)\right]^n}{1-\gamma\left[F_Z(t)\right]^n}\mathcal{M}_{Tr}^{KM+}(v) + \frac{(1-\gamma)\left[F_Z(t)\right]^n g_{n-1}(v)}{1-\gamma\left[F_Z(t)\right]^n}\end{aligned}$$

and the CGF of $\ln(PL+)$ is given as

$$\begin{aligned}K_{\ln(PL+)}(v) &= \ln\left[\left\{1-\left[F_Z(t)\right]^n\right\}\mathcal{M}_{Tr}^{KM+}(v) + (1-\gamma)\left[F_Z(t)\right]^n g_{n-1}(v)\right] \\ &\quad - \ln\left\{1-\gamma\left[F_Z(t)\right]^n\right\}\end{aligned}$$

and as such we will use results from the previous section in our derivations.

The first derivative of this CGF is given as

$$\begin{aligned}K'_{\ln(PL+)}(v) &= \frac{\left\{1-\left[F_Z(t)\right]^n\right\}\mathcal{M}_{Tr}^{KM+}(v)' + (1-\gamma)\left[F_Z(t)\right]^n g'_{n-1}(v)}{\left\{1-\left[F_Z(t)\right]^n\right\}\mathcal{M}_{Tr}^{KM+}(v) + (1-\gamma)\left[F_Z(t)\right]^n g_{n-1}(v)} \\ &= \frac{B'(v) + (1-\gamma)\left[F_Z(t)\right]^n g'_{n-1}(v)}{B(v) + (1-\gamma)\left[F_Z(t)\right]^n g_{n-1}(v)}\end{aligned}$$

and the second derivative of the CGF is

$$\begin{aligned}&K''_{\ln(KM+)}(v) \\ &= \frac{\left\{B(v)+(1-\gamma)\left[F_Z(t)\right]^n g_{n-1}(v)\right\}\left\{B''(v)+(1-\gamma)\left[F_Z(t)\right]^n g''_{n-1}(v)\right\}}{\left[B(v)+(1-\gamma)\left[F_Z(t)\right]^n g_{n-1}(v)\right]^2} \\ &\quad - \frac{\left[B'(v)+(1-\gamma)\left[F_Z(t)\right]^n g'_{n-1}(v)\right]^2}{\left[B(v)+(1-\gamma)\left[F_Z(t)\right]^n g_{n-1}(v)\right]^2}.\end{aligned}$$

APPENDIX C

DERIVATIVES OF THE CGF FOR THE LOGARITM OF ZERO-TRUNCATED ACL ESTIMATOR

The Mellin transform for the zero-truncated ACL estimate is of the form

$$\mathcal{M}_{Tr}^{ACL+}(v) = \frac{\sum_{r=0}^{n-1} b\left(r, F_Z(t)\right)\left[\gamma\left(\frac{n-r}{n}\right)^{\frac{v}{n}} + (1-\gamma)\right]^n}{1-\left[F_Z(t)\right]^n}$$
$$= \frac{\sum_{r=0}^{n-1} b\left(r, F_Z(t)\right) g_r(v)}{1-\left[F_Z(t)\right]^n}$$

which is identical in form to the Mellin transform for the zero-truncated KM estimate with the exception that in the latter

$$g_r(v) = \prod_{i=1}^{r}\left[\gamma c_{in}^{v} + (1-\gamma)\right].$$

As a result, it suffices to determine expressions for the first and second derivatives of

$$g_r(v) = \left[\gamma\left(\frac{n-r}{n}\right)^{\frac{v}{n}} + (1-\gamma)\right]^n.$$

Note that

$$g_r(v) = \left[h_r(v)\right]^n$$

where

$$h_r(v) = (1-\gamma) + \gamma\left(\frac{r}{n}\right)^{\frac{v}{n}}.$$

From this we obtain

$$g_r'(v) = n\left[h_r(v)\right]^{n-1} h_r'(v)$$

where

$$h_r'(v) = \frac{\gamma}{n}\left(\frac{r}{n}\right)^{\frac{v}{n}} \ln\left(\frac{r}{n}\right)$$

and furthermore

$$g_r''(v) = n\left[h_r(v)\right]^{n-1} h_r''(v) + n(n-1)\left[h_r(v)\right]^{n-2}\left[h_r'(v)\right]^2$$

where

$$\begin{aligned} h_r''(v) &= \left[\frac{\gamma}{n}\ln\left(\frac{r}{n}\right)\right]\left[\frac{1}{n}\left(\frac{r}{n}\right)^{\frac{v}{n}}\ln\left(\frac{r}{n}\right)\right] \\ &= \gamma\left[\frac{1}{n}\ln\left(\frac{r}{n}\right)\right]^2\left(\frac{r}{n}\right)^{\frac{v}{n}}. \end{aligned}$$

BIBLIOGRAPHY

[1] Agresti A. Survey of Exact Inference for Contingency Tables. Statistical Science 1992, 7, 131-153.

[2] Abdushukurov. On some estimates of the distribution function under random censorship. Conference of Young Scientists, VINITI, 1984. 8756-V.

[3] Billingsley. Probability and Measure. 2nd Edition, Wiley, New York 1986.

[4] Butler RW. Saddlepoint Approximations with Applicationsrt. New York 2007.

[5] Breslow, Crowley. A large sample study of the life table and product limit estimates under random censorship. The Annals of Statistics 1974, 2, 437–453.

[6] Borgan, Leistøl. A note on confidence bands for the survival curve based on transformations. Scandanavian Journal of Statistics 1990, 17, 35-41

[7] Cheng PE, Chang YC. PLE versus MLE for survivor functions under a simple proportional hazards model. The Chinese Statistical Association 1985, 23, 57-73.

[8] Chen, YY, Hollander M, Langberg NA. Small-Sample Properties of Two Survival Function Estimators Based on Incomplete Data. AFOSR Technical Report 1980, 78-112

[9] Chen YY, Hollander M, Langberg, NA. Small-Sample Results for the Kaplan-Meier Estimator. Journal of the American Statistical Association 1982, 77, 141-144.

[10] Myron Chang. Exact distribution of the Kaplan-Meier estimator under the proportional hazards model. Statistics and Probability Letters 1996, 28, 153-157

[11] Cheng PE, Lin GD. Maximum likelihood estimator of a survival function under the Koziol-Green proportional hazards model. Statistics & Probability Letters 1987, 75-80.

[12] Chris Barker. The Mean, Median, and Confidence Intervals of the Kaplan-Meier Survival Estimate Computations and Applications. American Statistician 2009, 87-80.

[13] Daniels HE. Saddlepoint Approximations in Statistics. The Annals of Mathematical Statistics 1954, 25, 631-650.

[14] Dawid AP. Conditional Independence in Statistical Theory, Royal Statistical Society, Seriesm B, 1979, 41 , 1–31.

[15] Davison AC, Hinkley DV. Bootstrap Methods and Their Application, Cambridge, New York 1997, section 3.5, 82-87.

[16] Davison and Wang. Miscellanea Saddlepoint approximations as smoothers. Biometrika 2002, 89, 933-938.

[17] Bradley Efron. The two sample problem with censored data. In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, 1967, V 4, 831-853.

[18] Bradley Efron. Censored Data and the Bootstrap. American Statistical Association 1981, 76, 312- 319.

[19] Frederic U, Alex S. Some Applications of the Bootstrap to Survival Analysis. Anuario de Psicologia 1992, 55, 155-167.

[20] Gill R. Censoring and Stochastic Integrals, Mathematical Centre Tracts 124. Mathematisch Centrum, Amsterdam.

[21] Greenwood M. The natural duration of cancer. Reports on Public Health and Medical , Her Majesty's Stationery Office, London, 1926.

[22] Hosmer, Lemeshow. Applied survival analysis: regression modeling of time to event data. John Wiley & Sons, New York, 1999.

[23] Hollander, Proschan F, Sconing J. Measures of dependence for evaluating information in censored models Technical Report No. M706, Depart ment of Statistics, Florida State University, 1985.

[24] Jeong, Cho. Confidence Intervals for the Median Survival Time under Proportional Censorship. The Korean Communications in Statistics, 2002.

[25] Klein P. Small sample moments of some estimators of the variance of the Kaplan-Meier and Nelson-Aalen estimators. Scand. J. Statistics, 1991, 18, 333-340.

[26] Kaplan L, Meier P. Nonparametric estimation from incomplete observations. American Statistical Association, 1958, 457-481.

[27] Kalbfleisch J, Prentice R. The Statistical Analysis of Failure Time Data. 1st Edition, Wiley, New York, 1980

[28] Lugannani, Rice S. Saddl Point Approximation for the Distribution of the Sum of Independent Random Variables. Advances in Applied Probability, 1980,12, 475-490.

[29] Moeschberger, Klein. A comparison of several methods of estimating the survival function when there is extreme right-censoring, Biometrics, 1985, 41, 253–259.

[30] E Parzen. United statistics, confidence quantiles, Bayesian statistics. Journal Statistical Planning and Inference, 2008, 138, 2777–2785.

[31] E Parzen. United Applicable Statistics, Mid Distribution. Mid Quantile, Mid P Confidence Intervals Proportion p, University of Maryland Kedem Celebration Lecture, 2009.

[32] R D Routledge. Practicing Safe Statistics with the Mid-p. Canadian Journal of Statistics, 1994, 22, 103-110.

[33] Melvin Springer. The algebra of random variables. Wiley, New York, 1979, 470.

[34] Zhanga, Bhaskara Raob, Mitrad. On the generalized maximum likelihood estimator of survival function under Koziol–Green model. Statistical Planning and Inference, 2006, 136, 3032 – 3051.

VITA

Emad Mohamed Abdulrasul was born on December 08, 1971 in Tripoli, Libya. He enrolled in Al-fateh University, Tripoli, Libya for his Bachelor of Science degree in 1989. Emad, subsequently received a Bachelor of Science with a major in Statistics from Al-fateh University, Tripoli, Libya after three years in May 1992. During the next ten years, he was employed as a statistician with the Documentation and Information Department at General Electricity Company, Tripoli, Libya. Emad also earned a Master of Science degree in Economic Planning from Institute of Planning in Tripoli of Libya in 2005 after which he became a staff member at the Department of Economics at the University of the Seventh of October in Misurata, Libya. In 2009, Emad entered Missouri University of Science and Technology for further graduate studies and obtained his Master's degree in applied mathematics, with emphasis in statistics in 2012. He received his PhD degree in Mathematics with emphasis in Statistics, from Missouri University of Science and Technology, in May, 2016.