Likelihood Ratio Confidence Bands in Survival Analysis

M.M. SANCHEZ and S.A. MURPHY

July 21, 1995

ABSTRACT

Confidence bands, which are formed via an inversion of a likelihood ratio test will be presented. Confidence bands based on the estimator alone are known to have poor coverage probabilities due to nonnormal small sample behavior of the Kaplan-Meier estimator. This has resulted in the use of various transformations such as the arcsine and log-minus-log. However, inversion of a likelihood ratio test, eliminates the need to search for a best transformation. These bands are compared to the equal precision, the Hall-Wellner confidence bands and with their various transformations.

Key words: Likelihood Ratio Test; Semi-Parametric Model

M. M. Sanchez is a PhD Student and S. A. Murphy is Assistant Professor, Department of Statistics, Pennsylvania State University, 326 Classroom Building, University Park, PA 16802.

1 INTRODUCTION

We consider a class of confidence bands for the survival function, $S(\cdot)$, based on the combination of the likelihood ratio confidence intervals. In particular confidence bands can be formed by combining confidence intervals with equal or unequal confidence coefficients across time points. A survival function, $S_0(\cdot)$, belongs to the confidence band if and only if at each time point, t, $S_0(t)$ belongs to the confidence interval at time t.

Confidence bands for the survival function have received much attention. In all cases (Gillespie and Fisher, 1979; Hall and Wellner, 1980; Nair, 1984; Csorgo and Horvath, 1986 and Hollander and Pena, 1989), each of the combined confidence intervals is based on the Kaplan-Meier estimator or transformations thereof and estimators of its asymptotic variance. In particular, Nair's (1984) confidence band is formed by combining the confidence intervals for $S(t)$ of equal confidence coefficients across t. Each confidence interval has confidence level, $2\Phi\left(d_\alpha\right) - 1$ where Φ is the standard normal distribution function, and d_α is the cutoff point so that the resulting band has confidence level, $1 - \alpha$. For this reason, this band is called the equal precision band. (Andersen et.al. 1992). On the other hand, Hall and Wellner's (1980) confidence band is composed of the same confidence intervals but with differing confidence coefficients. In particular the confidence interval at time t has confidence level, $2\Phi\left(e_\alpha \frac{(1+\sigma^2(t))}{\sigma(t)}\right) - 1$ where $\sigma^2(t)$ is the asymptotic variance of $\hat{S}(t)/S(t)$ and e_α is the appropriate cutoff point. A more general version of the Hall-Wellner confidence bands are proposed by Gillespie and Fisher (1979). For each time t, their confidence interval has a confidence level, $\Phi\left(\frac{b_\alpha^2 + c_\alpha^2 \sigma^2(t)}{\sigma(t)}\right) - \Phi\left(\frac{b_\alpha^1 + c_\alpha^1 \sigma^2(t)}{\sigma(t)}\right)$ where $b_\alpha^1, b_\alpha^2, c_\alpha^1$ and c_α^2 are chosen such that the resulting confidence band has confidence level, $1 - \alpha$

Bands based on the Kaplan-Meier estimator, $\hat{S}$ may not perform well in small samples due to non-normality of the small sample distribution of $\hat{S}$. As a result various transformations have been suggested to improve small sample performance (Kalbfleisch and Prentice, 1980 and Thomas and Grunkemeier, 1975). Small sample properties of these confidence intervals and confidence bands and their transformations were studied by Bie, et.al. (1987) and Borgan and Liestol, (1990).

The likelihood ratio confidence intervals were first proposed by Thomas and Grunkemeier (1975). These authors showed via simulation results that the likelihood ratio confidence interval compared favorably to confidence intervals based on the asymptotic normality of the Kaplan-Meier estimator, $\hat{S}$. A theoretical justification for the use of the chi-squared percentiles in the confidence interval was later given in Murphy (1995), Gang (1995).

In the next section we derive the two likelihood ratio confidence bands, one with equal confidence coefficients across t (like EP bands) and the other with varying confidence coefficients across t (like HW bands). In section 3., the small sample properties of the confidence bands are compared with those of existing confidence bands via simulation studies. Lastly, an example from the melanoma study conducted by K.T. Drzewiecki is used to illustrate the proposed likelihood based confidence bands.

2 Likelihood-based Confidence bands

First, we describe the failure time setting under which the confidence bands are constructed and compared. The observations are $(X_1, \delta_1), (X_2, \delta_2), \ldots, (X_n, \delta_n)$ where $X_i = \min(T_i, C_i)$, T_i is the failure time, C_i is the censoring time and δ_i is an indicator variable on whether $X_i = T_i$. Assume that $T_1, T_2, \ldots, T_n$ is a random sample of failure times from a distribution F with survival function $S = 1 - F$. Similarly, $C_1, C_2, \ldots, C_n$ is a random sample from a distribution G. The failure times, T_i and the censoring times, C_i are also assumed to be independent.

Let $\alpha = f/S$, the hazard rate function of T_i,

$A(t) = \int_0^t \alpha(s)ds$, the cumulative or integrated hazard function,

$N(t) = \sum_{i=1}^n I\{X_i \leq t, \delta_i = 1\}$, the number of observed failures up to time t and

$Y(t) = \sum_{i=1}^n I\{X_i > t\}$, the number at risk just before time t.

We consider the binomial extension of the likelihood which is proportional to

$$\prod_{t>0} (\triangle A(t))^{\triangle N(t)} (1 - \triangle A(t))^{Y(t) - \triangle N(t)}$$

where the product is over t which are observed failure times, $\triangle A(t) = A(t) - A(t-)$ and $\triangle N(t) = N(t) - N(t-)$. This extension may be interpreted as $\triangle N(t)$ having a Binomial distribution with parameters $Y(t)$ and $\triangle A(t)$ given the past up to time t. Notice that maximizing the above function is equivalent to maximizing

$$\prod_{i=1}^n [\triangle A(T_i)]^{\delta_i} [1 - \triangle A(T_i)]^{1-\delta_i} \prod_{s<T_i} [1 - \triangle A(s)]. \tag{1}$$

Now, equation (1) can be expressed in terms of $F(t)$ by using the relationship $1 - F(t) = \prod_{s \leq t} [1 - \triangle A(s)]$. We get

$$\prod_{i=1}^n [F(X_i) - F(X_i-)]^{\delta_i} [1 - F(X_i)]^{1-\delta_i} \tag{2}$$

which is a likelihood function for discrete failure time data as mentioned by Andersen, et. al. (1993). Note that this is also the likelihood used by Kaplan and Meier (1958) in deriving the estimator for the survival function S. It is also used by Murphy (1995), Li (1995) and Thomas and Grunkemeier (1975). In this sequel, confidence bands for the survival function are constructed based on this binomial extension of the likelihood.

The likelihood ratio test statistic (LRT_t) of $H_0 : S(t) = \theta_0$ is $2(\ln\, L(\hat{S}) - \ln\, L(\hat{S}_0))$ where $\hat{S}$ is the maximum likelihood estimator (MLE) under the whole space and $\hat{S}_0(t)$ is the restricted MLE under H_0. The estimators $\hat{S}$ and $\hat{S}_0$ are given by

$$\hat{S}(t) = \prod_{s\leq t}(1 - d\hat{A}(s)) = \prod_{s\leq t}(1 - \triangle\hat{A}(s))$$

$$\hat{S}_0(t) = \prod_{s\leq t}(1 - d\hat{A}_0(s)) = \prod_{s\leq t}(1 - \triangle\hat{A}_0(s))$$

where

$$\hat{A}(t) = \int_0^t \frac{1}{Y(s)} dN(s)$$

is the Nelson-Aalen estimator and

$$\hat{A}_0(u) = \int_0^u \frac{1}{Y(s) + \lambda I\{s \leq t\}n} dN(s).$$

The constrained maximum likelihood estimator, $\hat{A}_0$ is derived using the Lagrange multiplier method with $n\lambda$ as the Lagrange multiplier. We maximize the function,

$$\sum_s \triangle N(s) \ln\,[\triangle A(s)] + [Y(s) - \triangle N(s)] \ln\,[1 - \triangle A(s)] + n\lambda\{\ln[\prod_{s\leq t}(1 - \triangle A(s))] - \ln[\theta_0]\}$$

by taking the derivative with respect to $\triangle A(t)$, and equating it to 0 which yields the restricted MLE,

$$\triangle\hat{A}_0(s) = \frac{\triangle N(s)}{Y(s) + n\lambda I\{s \leq t\}}$$

under $H_0 : S(t) = \theta_0$. Note that $\hat{S}_0(t)$ is uniquely determined by the Lagrangian multiplier $n\lambda$. It is monotone in $n\lambda$, *i.e.* $\hat{S}_0(t)$ increases to 1 as $n\lambda$ increases while it goes to 0 as $n\lambda$ decreases to $1 - Y(t)$. The Lagrangian multiplier, λ is chosen so that $\theta(t) = \hat{S}_0(t) = \prod_{s\leq t}(1 - \triangle\hat{A}_0(s))$. $\hat{A}_0$ is equal to $\hat{A}$ for $s > t$ so the LRT_t involves only time points less than t since beyond t it is just equal to 0. It is given by

$$\begin{aligned} LRT_t &= 2\sum_{s\leq t}\left\{\triangle N(s)\ln\,\left[\frac{Y(s) + \lambda n}{Y(s)}\right] + (Y(s) - \triangle N(s))\ln\,\left[\frac{1 - \frac{1}{Y(s)}}{1 - \frac{1}{Y(s)+\lambda n}}\right]\right\} \qquad (3) \\ &= 2\int_0^t \ln\,[1 + \bar{Y}(s)^{-1}\lambda] + (Y(s) - 1)\ln\,\left[\frac{1 - \frac{1}{Y(s)}}{1 - \frac{1}{Y(s)+\lambda n}}\right] dN(s) \end{aligned}$$

where $\bar{Y}(s) = Y(s)/n$. Note that the LRT_t as a function of t is a step function with steps at observed failure times. It can be shown that this LRT statistic is equivalent to

$$\left(\frac{\sqrt{n}(\hat{S}(t) - \theta(t))}{S(t)\sqrt{\int_0^t \bar{Y}^{-1} d\hat{A}}}\right)^2 + o_p(1)$$

uniformly in $t \in [t_1, t_2]$ where $0 < t_1 \leq t_2 \leq \tau$, $\tau = \sup\{t : S(t) > 0\}$ and $G(\tau-) < 1$ (see the Appendix). For each t, the asymptotic distribution of

$$\frac{\sqrt{n}(\hat{S}(t) - \theta(t))}{\hat{S}(t)\sqrt{\int_0^t \bar{Y}(s)^{-1} d\hat{A}(s)}}$$

is a standard normal distribution so that LRT_t is asymptotically chi-square with 1 degree of freedom (Andersen, et.al., 1993).

A $100(1-\alpha)\%$ likelihood ratio based confidence interval for $S(t)$ is the set $\{\theta(t) : LRT_t \leq z^2\}$ where z is the $1 - \alpha/2$ percentile of the standard normal distribution. For the case of equal confidence coefficient across time points, we form the LRT confidence band for $S(t)$ by combining the confidence intervals for $S(t)$ for all $t \in [t_1, t_2]$. The set $\{\theta(t) : t \in [t_1, t_2]\}$ is in all confidence intervals if $H_0 : S(t) = \theta(t)$ is accepted for all $t \in [t_1, t_2]$. This is equivalent to $LRT_t \leq (z^*)^2$, where z^* is the $1 - \alpha/2$ percentile of the asymptotic distribution of

$$\sup_{t \in [t_1, t_2]} \left| \frac{\hat{S}(t) - \theta(t)}{\hat{S}(t)\sqrt{\int_0^t \bar{Y}(s)^{-1} d\hat{A}(s)}} \right|$$

(see Andersen, et.al. 1992).

We also consider a Hall-Wellner type of band to construct a likelihood ratio based confidence band with unequal confidence coefficient across time points. For each t, we replace z^* above with $e_\alpha(\hat{c}_1, \hat{c}_2)(1 + n\hat{\sigma}^2(t))/n\hat{\sigma}(t)$ where $e_\alpha(\hat{c}_1, \hat{c}_2)$ is the upper α fractile used in the Hall-Wellner confidence bands and $\hat{\sigma}^2(t) = \int_0^t \{Y(s)(Y(s) - \triangle N(s))\}^{-1} dN(s)$.

It is not obvious that the resulting set of θ forms an interval nor that we get a band of points when they are combined over all $t \in [t_1, t_2]$. Thomas and Grunkemeier (1975) show that the resulting confidence sets from the likelihood ratio tests are indeed intervals.

In the appendix, we show that we get a band of points when the intervals for all $t \in [t_1, t_2]$ are combined. In addition, the resultant confidence band has the appropriate coverage (see Appendix for the proof). In the next section, we give a detailed discussion on the construction of the likelihood ratio confidence bands.

3 Comparison of the different confidence bands

The confidence bands considered for comparison are the log-minus-log and the arcsine transformation of the equal precision (EP) bands and the Hall-Wellner (HW) band. Borgan and Liestol (1990) show substantial improvement is observed in applying the transformations for EP bands while no substantial improvement is observed when the transformations are used on the HW bands. Each of the confidence bands considered in this sequel are described below.

The variance of $\hat{S}$ used in all the confidence bands is estimated by Greenwood's formula (Greenwood, 1926),

$$\widehat{\text{var}}\ \hat{S}(t) = (\hat{S}(t))^2 \hat{\sigma}^2(t)$$

where

$$\hat{\sigma}^2(t) = \int_0^t \{Y(s)(Y(s) - \triangle N(s))\}^{-1} dN(s) \tag{4}$$

Note that when $Y(s) = \triangle N(s)$ for some time point s, $\hat{\sigma}^2(t) = \infty$ for $t \geq s$ but at the same time $\hat{S}(t) = 0$ for $t \geq s$ so $\hat{\sigma}^2(t)$ is defined to be 0 in this situation.

3.1 EP bands

A $100(1-\alpha)\%$ asymptotic confidence band for the survival function S on $[t_1, t_2]$ has the form

$$\hat{S}(s) \pm d_\alpha(\hat{c}_1, \hat{c}_2)\hat{S}(s)\hat{\sigma}(s)$$

where $\hat{\sigma}^2(t)$ is given in eq(3) and $d_\alpha(\hat{c}_1, \hat{c}_2)$ is the upper α fractile in the distribution of

$$\sup_{c_1 \leq x \leq c_2} |W_0(x)[x(1-x)]^{-1/2}|$$

where W_0 is a tied-down Wiener process on $(0,1)$ and

$$\hat{c}_i = \frac{n\hat{\sigma}^2(t_i)}{1 + n\hat{\sigma}^2(t_i)}$$

for $i = 1, 2$ and $0 \leq t_1 \leq t_2 \leq t$.

For the log-minus-log transformation, the confidence bands has the form

$$\hat{S}(s)^{exp\{\pm c_{\alpha/2}\hat{\sigma}(s)/\log \hat{S}(s)\}}$$

while for the arcsine transformation, the confidence bands has the form

$$\sin^2\left\{\max\left(0, \arcsin(\sqrt{\hat{S}(s)}) - \frac{1}{2}d_\alpha(\hat{c_1}, \hat{c_2})\hat{\sigma}(s)\left\{\frac{\hat{S}(s)}{1-\hat{S}(s)}\right\}^{0.5}\right)\right\} \leq S(s)$$
$$\leq \sin^2\left\{\min\left(\frac{\pi}{2}, \arcsin(\hat{S}(s)^{1/2}) + \frac{1}{2}d_\alpha(\hat{c_1}, \hat{c_2})\hat{\sigma}(s)\left\{\frac{\hat{S}(s)}{1-\hat{S}(s)}\right\}^{0.5}\right)\right\}.$$

3.2 HW bands

The $100(1-\alpha)\%$ asymptotic Hall-Wellner confidence bands for S (Hall and Wellner, 1980) is given by

$$\hat{S}(s) \pm n^{-1/2}e_\alpha(\hat{c}_1, \hat{c}_2)(1 + n\hat{\sigma}^2(s))\hat{S}(s)$$

where $e_\alpha(\hat{c}_1, \hat{c}_2)$ is the upper α fractile in the distribution of

$$\sup_{c_1 \leq x \leq c_2} |W^0(x)|$$

3.3 LR bands

The main idea in the construction of the LR bands is to first construct the corresponding confidence intervals for each time point, $t \in [t_1, t_2]$ and then combine them to create the confidence band. Confidence intervals need to be computed only for the observed failures times between $[t_1, t_2]$ since the LRT_t is constant between observed failure times Note that the form of the interval is implicitly defined, (i.e. all $\theta(t) \ni H_0 : S(t) = \theta(t)$ is not rejected) so we need to use iterative techniques to find both the upper limit and lower limit of the confidence interval.

First, we explain how to compute for the LRT_t for a constraint $\hat{S}_0(t) = \theta(t)$. The initial step is to determine the value of λ (say λ_L) that satisfies the constraint, $\hat{S}_0(t) = b_L$ where

$$\hat{S}_0(t) = \prod_{s \leq t} \left(1 - \frac{1}{Y(s) + n\lambda}\right)$$

We can easily find the root λ_L of the function $\hat{S}_0(t) - b_L$ since $\hat{S}_0$ is monotone in λ. After determining λ_L, we find the value of the corresponding LRT_t (given in eq (3)).

Steps: For a fixed t,

1. Find an initial guess for the lower limit and upper limit, say b_L and b_U respectively. The initial guesses for the lower and upper limits are usually obtained from an alternative method (e.g. EP bands or HW bands). In the simulation, we use the limits computed using the log-minus-log transformed EP bands.

2. To obtain the lower limit LL, evaluate the LRT_t at $\theta(t) = b_L$ and if $LRT_t > (z^*)^2$ then increment LL up towards $\hat{S}(t)$ otherwise decrease LL down towards 0. This is repeated until $|LRT_t - (z^*)^2|$ is small.

3. To obtain the upper limit UL, evaluate the LRT_t at $\theta(t) = b_U$ (similar to step 2) and if $LRT_t > (z^*)^2$ then decrement UL from b_U towards $\hat{S}(t)$ otherwise increase UL up towards 1. Again this is repeated until $|LRT_t - (z^*)^2|$ is small.

Note that we follow the same procedure above to construct the unequal confidence coefficient likelihood ratio confidence band. We only need to replace z^* with $e_\alpha(\hat{c}_1, \hat{c}_2)(1 +$

Table 1: Achieved Error Rates and Area Covered for Equal Confidence Coefficients Bands

	Std Exp/ Std Exp			Std Exp/Unif(0,b)		
n	LR	EP log(-log)	EP arcsine	LR	EP log(-log)	EP arcsine
	$\hat{\alpha}$ (area)	$\hat{\alpha}$ (area)	$\hat{\alpha}$ (area)	$\hat{\alpha}$ (area)	$\hat{\alpha}$ (area)	$\hat{\alpha}$ (area)
25	0.04 (0.69)	0.06 (0.72)	0.04 (0.71)	0.01 (0.05)	0.07 (0.06)	0.003 (0.04)
50	0.05 (0.65)	0.07 (0.66)	0.05 (0.66)	0.01 (0.04)	0.04 (0.05)	0.005 (0.04)
200	0.04 (0.42)	0.05 (0.42)	0.05 (0.43)	0.02 (0.03)	0.03 (0.03)	0.002 (0.03)

$n\hat{\sigma}^2(t))/n\hat{\sigma}(t)$ where $e_\alpha(\hat{c}_1, \hat{c}_2)$ is the upper α fractile used in the Hall-Wellner confidence bands.

3.4 Numerical/Simulation Results

The factors that affect the simulation are the following F, the failure time distribution, G, the censoring time distribution and n, the sample size. In our study, we consider the following combination of survival and censoring distribution, (a) standard exponential/standard exponential, and (b) standard exponential/uniform. The sample sizes used range from small ($n = 25$) to large ($n = 200$).

The simulation results for the confidence bands with a nominal level of confidence, 95% are based on 10000 replications. This corresponds to a standard error on the estimated error rates $\hat{\alpha}$ of 0.002 so a 95% confidence interval for α is $\hat{\alpha} \pm 0.004$. Table 1 shows the achieved error rates and area covered for the bands with equal confidence coefficients across time points like the LR band and the transformed versions of the EP band.

Table 2: Achieved Error Rates and Area Covered for Unequal Confidence Coefficients Bands

	Std Exp/ Std Exp		Std Exp/Unif(0,b)	
n	LRU	HW	LRU	HW
	$\hat{\alpha}$ (area)	$\hat{\alpha}$ (area)	$\hat{\alpha}$ (area)	$\hat{\alpha}$ (area)
25	0.05 (0.69)	0.07 (0.78)	0.02 (0.06)	0.002 (0.06)
50	0.06 (0.66)	0.07 (0.72)	0.01 (0.05)	0.002 (0.05)
200	0.05 (0.47)	0.05 (0.49)	0.01 (0.04)	0.01 (0.04)

Overall the likelihood ratio based confidence band performs better than the EP bands when both the survival and censoring distributions are standard exponential. The achieved error rates for the LR bands are close to the nominal value of 0.05 and they cover a smaller area than the EP bands. Thus, the LR bands are more precise than the EP bands.

However, both the LR band and the EP bands don't perform well when the censoring distribution is a uniform distribution. The EP arcsine band gives too low error rates even with sample size $n = 200$. While the EP log(-log) band is better than the LR bands in terms of the achieved error rates although they both have similar area covered.

On the other hand, Table 2 contains the achieved error rates and area covered for the unequal confidence coefficient confidence bands namely, the unequal likelihood ratio (LRU) and the Hall-Wellner (HW). It shows that the LRU bands performed better than the HW bands in the case of small ($n = 25$) and moderate ($n = 50$) samples for both combinations. However, both the LRU bands and the HW bands tend to underestimate the error rates when the survival distribution is standard exponential and the censoring distribution is uniform.

4 An Example

To illustrate the likelihood ratio based confidence bands we use the Melanoma survival data (1962-1977) collected by K.T. Drzewiecki at Odense University Hospital (see Appendix of Andersen, et.al., 1992). 225 patients with malignant melanoma (cancer of the skin) had a radical operation performed which involves completely removing the tumor together with the skin within a distance of about 2.5 cm around it. All patients were followed until the end of 1977. The time variable is time since operation. Patients still alive at the conclusion of the study and those who died of causes unrelated to cancer were considered censored.

Figure 1 shows the equal confidence coefficient confidence bands valid for the interval $[0.6, 7.6]$ years. The EP log-log transformed bands were slightly wider than both the LR and EP arcine confidence bands. There is not much difference between the LR and the EP arcsine confidence bands. Figure 2 shows the unequal confidence coefficient confidence bands valid for the same time interval. The HW confidence bands include values greater than 1 so they are not that good. On the other hand, the LRU confidence bands stays within the valid range of $[0, 1]$. At the initial years, the LRU seems to be just a shifted version of the HW bands while they are similar for the later years.

5 Conclusion

The likelihood ratio confidence bands performed as well if not better than the existing confidence bands, namely the Hall-Wellner bands, EP bands and its transformed versions. It has additional advantages like respecting the range of the parameter so its lower and upper bounderies stay within the valid range $[0, 1]$ and is transformation respecting.

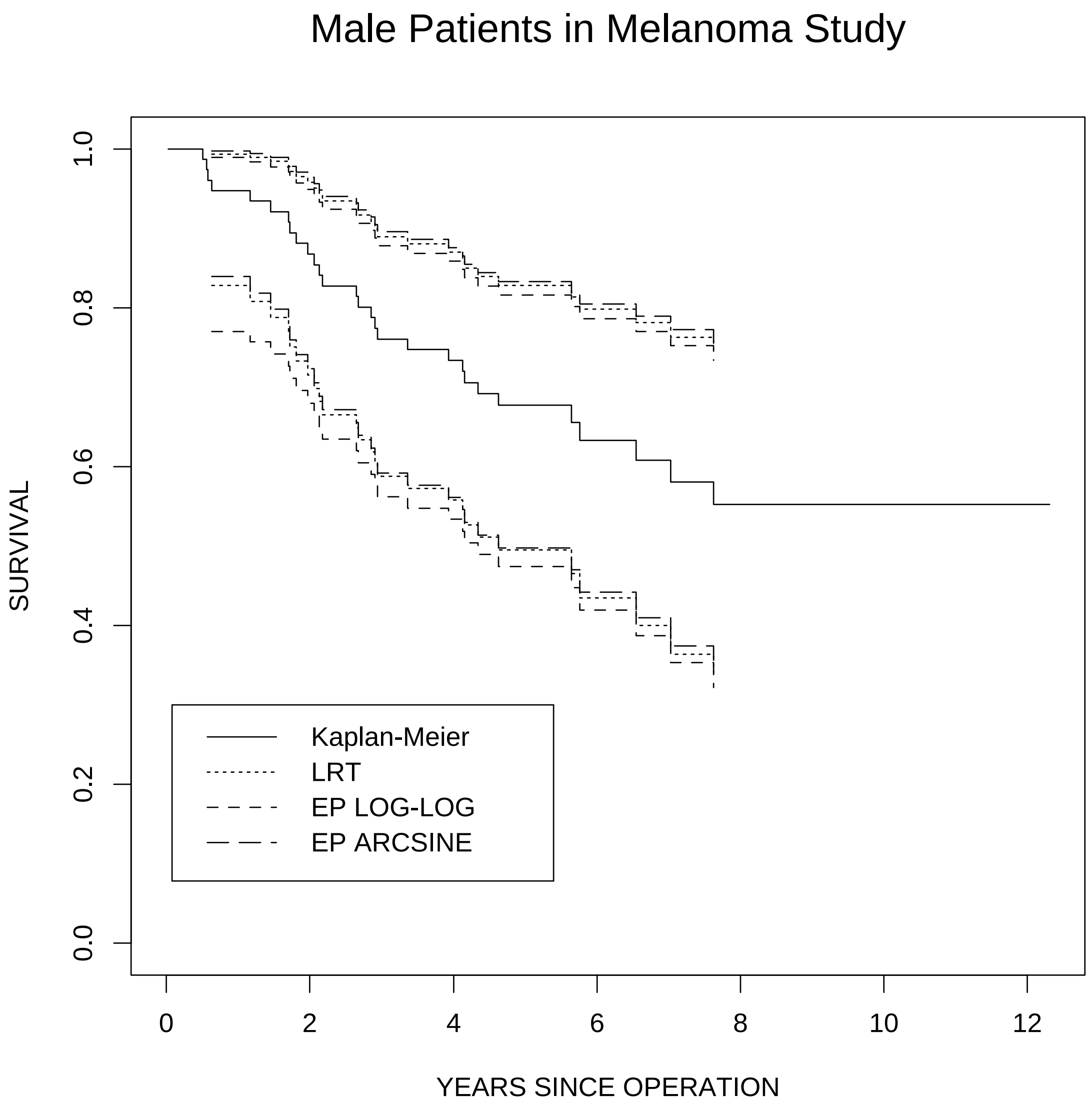

Figure 1: Equal Confidence Coefficient Confidence Bands for the interval [.6, 7.6] years

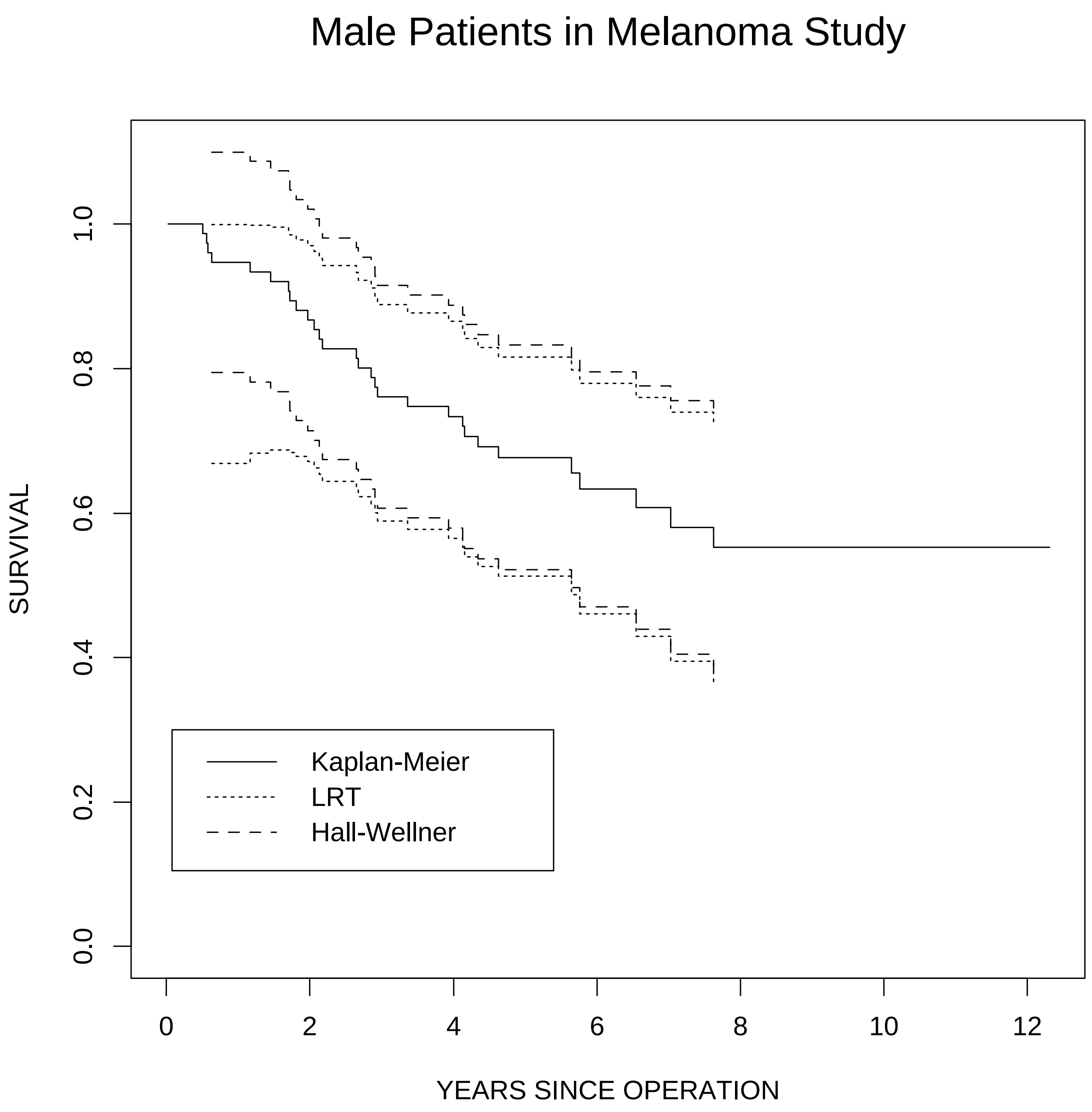

Figure 2: Unequal Confidence Coefficient Confidence Bands for the interval [.6, 7.6] years

6 Appendix

(Proofs can be found in the paper by Hollander, McKeague and Yang, 1995)

7 References

Andersen, P.K., Borgan, Ø., Gill, R.D. and Keiding, N. (1993), *Statistical Models Based on Counting Processes*, New York: Springer Verlag.

Bie, O., Borgan, Ø. and Liestøl, K. (1987), "Confidence Intervals and Confidence Bands for the Cumulative Hazard Rate Function and their Small Sample Properties," *Scandinavian Journal of Statistics*, **14**, 221-233.

Borgan, Ø. and Liestøl, K. (1990), "A Note on Confidence Intervals and Bands for the Survival Function Based on Transformations," *Scandinavian Journal of Statistics*, **17**, 35-42.

Brookmeyer, R. and Crowley, J.J. (1982), "A Confidence Interval for the Median Survival Time," *Biometrics*, **38**, 29-41.

DiCiccio, T.J., Hall, P. and Romano, J.P. (1991), "Empirical Likelihood is Bartlett Correctable," *Annals of Statistics*, **19**, 1053-1061.

Hall, P. and La Scala, B. (1990), "Methodology and Algorithms of Empirical Likelihood," *International Statistical Review*, **58**, 109-127.

Johansen, S. (1978), "The Product Limit Estimator as Maximum Likelihood Estimator," *Scandinavian Journal of Statistics*, **5**, 195-199.

Johansen, S. (1983), "An Extension of Cox's Regression Model," *International Statistical Review*, **51**, 258-262.

Kalbfleisch, J.D. and Prentice, R. L. (1980), *The Statistical Analysis Failure Time Data*, New York: Wiley.

Kaplan, E.L. and Meier, P. (1958), "Non-parametric Estimation from Incomplete Observations," Journal of the American Statistical Association, **53**, 457-481,562-563.

Kiefer, J. and Wolfowitz, J. (1956), "Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Nuisance Parameters," *Annals of Mathematical Statistics*, **27**, 887-906.

Laska, E.M. and Meisner, M.J. (1992), "Nonparametric Estimation and Testing in a Cure Model," *Biometrics*, **48**, 1223-1234.

Li, G. (1995), "On Nonparametric Likelihood Ratio Estimation of Survival Probabilities for Censored Data. *Statistics and Probability Letters* (to appear).

Murphy, S. (1995), "Likelihood Ratio Based Confidence Intervals in Survival Analysis," *Journal of the american Statistical Association* (to appear).

Owen, A.B. (1988), "Empirical Likelihood Ratio Confidence Intervals for a Single Functional," *Biometrika*, **75**, 237-249.

Thomas D.R. and Grunkemeier, G.L.(1975), "Confidence Interval Estimation of Survival Probabilities for Censored Data," Journal of the American Statistical Association, **70**, 865-871.