Empirical Likelihood in Survival Analysis

Gang Li[1], Runze Li[2], and Mai Zhou[3]

[1] Department of Biostatistics, University of California, Los Angeles, CA 90095
vli@ucla.edu
[2] Department of Statistics, The Pennsylvania State University, University Park, PA 16802-2111, rli@stat.psu.edu
[3] Department of Statistics, University of Kentucky, Lexington, KY 40506
mai@ms.uky.edu

Summary. Since the pioneering work of Thomas and Grunkemeier (1975) and Owen (1988), the empirical likelihood has been developed as a powerful nonparametric inference approach and become popular in statistical literature. There are many applications of empirical likelihood in survival analysis. In this paper, we present an overview of some recent developments of the empirical likelihood for survival data. We will focus our attentions on the two regression models: the Cox proportional hazards model and the accelerated failure time model.

Key words: Accelerated failure time model; Censored data; Cox proportional hazards model; Wilks theorem.

2000 Mathematics Subject Classification: 62N02, 62N03, 62G20.

1 Introduction

Empirical likelihood (EL) was proposed by Thomas and Grunkemeier (1975) in order to obtain better confidence intervals involving the Kaplan-Meier estimator in survival analysis. Based on the idea of Thomas and Grunkemeier, Owen (1988) established a general framework of EL for nonparametric inference. Since then, the EL method has become popular in the statistical literature. EL has many desirable properties. For instance, the EL based confidence interval is range preserving and transform respecting. There are many recent works of EL for survival analysis. These works have demonstrated the power of EL approach for analyzing censored data. Empirical likelihood has many nice properties, one of which is its ability to carry out a hypothesis testing and construct confidence intervals without the need of estimating the variance – because the EL ratio do not involve the unknown variances and the limiting distribution of EL is pivotal (chi square). This advantage of the EL method has been appreciated particularly in the survival analysis, since those variances can be very difficult to estimate, especially with censored data in survival analysis. Because of this difficulty, many estimation procedures saw limited action in the application. EL therefore can provide a way to circumvent the complicated variances and make many inference procedures practical.

Owen (1991) have investigated the use of EL in the regression models and obtained parallel results to the EL inference for the mean. Unfortunately his results are for the uncensored data only and the generalization to handle censored data is nontrivial. Thus, many authors have worked on EL for survival data. Li (1995) demonstrated that the likelihood ratio used by Thomas and Grunkemeier (1975) is a "genuine" nonparametric likelihood ratio. That is, it can be derived by considering the parameter space of all survival functions. This property is not shared by many existing EL. Li, Qin and Tiwari (1997) derived likelihood ratio-based confidence intervals for survival probabilities and for the truncation proportion under statistical setting in which the truncation distribution is either known or belong to a parametric family. Hollander, McKeague and Yang (1997) constructed simultaneous confidence bands for survival probability based on right-censored data using EL. Pan and Zhou (1999) illustrated the use of a particular kind of one-parameter sub-family of distribution in the analysis of EL. Einmahl and McKeague (1999) constructed simultaneous confidence tubes for multiple quantile plots based on multiple independent samples using the EL approach. Wang and Jing (2001) investigated how to apply the EL method to a class of functional of survival function in the presence of censoring by using an adjusted EL. Pan and Zhou (2002) studied the EL ratios for right censored data and with parameters that are linear functionals of the cumulative hazard function. Li and van Keilegom (2002) constructed confidence intervals and bands for the conditional survival and quantile functions using an EL ratio approach. McKeague and Zhao (2002) derived a simultaneous confidence band for the ratio of two survival functions based on independent right-censored data. Chen and Zhou (2003) extended the self consistent algorithm (Turnbull, 1974) to include a constraint on the nonparametric maximum likelihood estimator of the distribution function with doubly censored data. They further show how to construct confidence intervals and test hypothesis based on the nonparametric maximum likelihood estimator via the EL ratio.

The EL ratio has also been used to construct confidence intervals for other parameters or functionals of population in addition to survival probability. For instance, Ren (2001) used weighted EL ratio to derive confidence intervals for the mean with censored data. Adimari (1997) suggested a simple way to obtain EL type confidence intervals for the mean under random censorship. Li, Hollander, McKeague and Yang (1996) derived confidence bands for quantile functions using the EL ratio approach. The EL method has also been applied for linear regression with censored data (Qin and Jing, 2001, Li and Wang, 2003, Qin and Tsao, 2003). Furthermore, the EL method has been adapted for semiparametric regression models, including partial linear models

(Leblanc and Crowley, 1995, Shen, Shi and Wong, 1999, Qin and Jing, 2001, Lu, Chen and Gan, 2002, Wang and Li, 2002). Naiknimbalkar and Rajarshi (1997) proposed the EL ratio test for equality of k-medians in censored data. Cheb, Leung and Qin (2003) extend the EL method for censored data with surrogate endpoints.

This paper is organized as follows. In Section 2, we introduce in detail the EL results for mean with the right censored data, while in Section 3 we discuss the EL results for hazard with the right censored data. Section 4 discuss computation issue for the censored EL. The EL results for the Cox proportional hazards regression model is studied in Section 5. Section 6 present the EL method for the accelerated failure time models. Finally Section 7 gives a brief discussion on the EL results with other types of censored data.

2 Empirical Likelihood for the mean

The *mean* is referred to as all the statistics that can be defined by

$$\int g(t,\theta)dF(t) = K \ ; \tag{1}$$

where F is the unknown cumulative distribution function (CDF), g is a given function. Either K is a known constant then θ is an implicit parameter (eg. quantile); or θ is a known constant then K is a parameter (eg. mean, probability).

Suppose that $X_1, X_2, \cdots, X_n$ are independent and identically distributed lifetimes with CDF $F(t) = P(X_i \le t)$. Let $C_1, C_2, \cdots, C_n$ be censoring times with CDF $G(t) = P(C_i \le t)$. Further assume that the life times and the censoring times are independent. Due to censoring, we observe only

$$T_i = \min(X_i, C_i), \quad \delta_i = I_{[X_i \le C_i]}. \tag{2}$$

The EL of the censored data pertaining to F is

$$EL(F) = \prod_{i=1}^{n} [\varDelta F(T_i)]^{\delta_i} [1 - F(T_i)]^{1-\delta_i} \ , \tag{3}$$

where $\varDelta F(s) = F(s+) - F(s-)$.

It is well known that among all the cumulative distribution functions, the Kaplan-Meier (1958) estimator maximizes (3). Let us denote the Kaplan-Meier estimator by $\hat{F}_n(t)$. The maximization of the censored EL under an extra constraint (1) do not in general have an explicit expression. But we have the following results.

Theorem 1. *Suppose the true distribution of lifetimes satisfies the constraint (1) with the given g, θ and K. Assume also the asymptotic variance of $\sqrt{n} \int g(t,\theta) d\hat{F}_n(t)$ is positive and finite. Then, as $n \to \infty$*

$$-2\log\frac{\sup_F EL(F)}{EL(\hat{F}_n)} \xrightarrow{\mathcal{D}} \chi^2_{(1)} \ , \tag{4}$$

where the sup *is taken over all the CDFs that satisfy (1) and $F \ll \hat{F}_n$.*

The proof of Theorem 1 can be found in Murphy and van der Vaart (1997) or Pan and Zhou (1999).

Counting process martingale techniques now become a standard tool in the literature of survival analysis. Given the censored data (2), it is well known that we can define a filtration $\mathcal{F}_t$ such that

$$M_n(t) = \frac{\hat{F}_n(t) - F(t)}{1 - F(t)}$$

is a (local) martingale with respect to the filtration $\mathcal{F}_t$, see Fleming and Harrington (1991) for details. It is also well known that under mild regularity conditions $\sqrt{n}M_n(t)$ converges weakly to a time changed Brownian motion.

To extend the EL ratio theorem for regression models (e.g. the Cox model and the accelerated failure time model), in which the mean function used in defining the parameter, $g(t)$, is random but predictable with respect to $\mathcal{F}_t$, we need to impose the following two conditions on $g_n(t)$:

(i) $g_n(t)$ are predictable with respect to $\mathcal{F}_t$ and $g_n(t) \xrightarrow{\mathrm{P}} g(t)$ as $n \to \infty$.
(ii) $\sqrt{n}\int_{-\infty}^{\infty}\{g_n(t)[1-F_X(t)] + \int_{-\infty}^{t} g_n(s)dF_X(s)\}dM_n(t)$ converges in distribution to a zero mean normal random variable with a finite and non-zero variance.

It is worth noting that the integrand inside the curly bracket in (ii) is predictable. If we put a variable upper limit in the outside integration in (ii), then it is also a martingale. It is not difficult to give a set of sufficient conditions that will imply asymptotic normality. Usually a Lindeberg type condition is needed.

Theorem 2. *Suppose $g_n(t)$ is a random function but satisfies the above two conditions and assume that for each n, we have*

$$\int_{-\infty}^{\infty} g_n(t)dF_X(t) = 0 \ , \tag{5}$$

then

$$-2\log\frac{\sup_F EL(F)}{EL(\hat{F}_n)} \xrightarrow{\mathcal{D}} \chi^2_{(1)}$$

where the sup *in the numerator EL is taken over those F that $F \ll \hat{F}_n$ and satisfy the constraint*

$$\int_{-\infty}^{\infty} g_n(t)dF(t) = 0 \ . \tag{6}$$

The proof of Theorem 2 is given in Zhou and Li (2004). In Theorem 2, it is assumed that the true distribution of X_i satisfies (5). However, the Kaplan-Meier estimator may not satisfy this condition. Generalizations of the above two Theorems for multiple constraints of the type similar to (1) or (6) are seen to hold, but a formal proof is tedious and not available in the published works. The limiting distribution there will be $\chi^2_{(q)}$.

3 Empirical Likelihood for the hazard

Hazard function or risk function is a quantity often of interest in survival analysis, and it is often more convenient to model the survival data in terms of hazard function. For a random variable X with cumulative distribution function $F(t)$, define its cumulative hazard function to be

$$H(t) = \int_{(-\infty,t]} \frac{dF(s)}{1 - F(s-)} .$$

Given the censored data (T_i, δ_i) as in (2), one natural way to define the EL in terms of hazard is:

$$EL(H) = \prod_{i=1}^{n} [\Delta H(T_i)]^{\delta_i} \exp(-H(T_i)). \tag{7}$$

It can be easily verified that the maximization of the $EL(H)$ is obtained when the hazard is the Nelson-Aalen estimator, denoted by $\hat{H}_n(t)$. The statistics of interest is defined by

$$\int g_n(t, \theta) dH(t) = K, \tag{8}$$

where the meaning of the parameter is similar to the (1) above, except the integration is now with respect to H, and the g function is stochastic.

Theorem 3. *Suppose $g_n(t)$ is a sequence of predictable functions with respect to the filtration $\mathcal{F}_t$, and $g_n \xrightarrow{\text{P}} g(t)$ with*

$$0 < \int \frac{|g(x)|^m dH(x)}{(1 - F_X(x))(1 - G(x))} < \infty, \qquad m = 1, 2.$$

If the true underlying cumulative hazard function satisfies the condition (8) with given g_n, θ and K, then we have

$$-2 \log \frac{\sup_H EL(H)}{EL(\hat{H}_n)} \xrightarrow{\mathcal{D}} \chi^2_{(1)} \qquad \text{as } n \to \infty \ ,$$

where the sup *is taken over those H that satisfy (8) and $H \ll \hat{H}_n$.*

One nice feature here is that we can use stochastic functions to define the statistics, i.e. $g(t) = g_n(t)$, as long as $g_n(t)$ is predictable with respect to $\mathcal{F}_t$. See Pan and Zhou (2001). For example $g_n(t) =$ size of risk set at time t, will produce a statistic corresponding to the one sample log-rank test, etc.

Multivariate version of the Theorem 3 can similarly be obtained with a limiting distribution of $\chi^2_{(q)}$. The q parameters are defined through q equations similar to (8) with different g, θ and K.

4 Computation of the censored Empirical Likelihood

The computation of the EL ratios can sometimes be reduced by the Lagrange multiplier method to the dual problem. When it does, the computation of EL is relatively easy, and is equivalent to the problem of finding the root of q nonlinear but monotone equations with q unknowns (q being the degrees of freedom in the limiting chi square distribution). But more often than not the censored EL problem cannot be simplified by the Lagrange multiplier. A case in point is the right censored data with a mean constraint. No reduction to the dual problem is available there.

When the maximization of the censored empirical likelihood cannot be simplified by the Lagrange multiplier, alternative methods are available. We introduce two here.

Sequential quadratic programming (SQP) is a general optimization procedure and a lot of related literature and software are available from the optimization field. It repeatedly approximates the target function locally by a quadratic function. SQP can be used to find the maximum of the censored EL under a (linear) constraint and the constrained NPMLE. This in turn enables us to obtain the censored empirical likelihood ratio. The drawback is that without the Lagrange multiplier reduction, the memory and computation requirement increases dramatically with sample size. It needs to invert matrices of size $n \times n$. With today's computing hardware, SQP works well for small to medium sample sizes. In our own experience, it is quite fast for samples of size under 1000, but for larger samples sizes difficulty may rise.

EM algorithm has long been used to compute the NPMLE for censored data. Turnbull (1976) showed how to find NPMLE with arbitrary censored, grouped or truncated data. Zhou (2002) generalized the Turnbull's EM algorithm to obtain the maximum of the censored empirical likelihood under mean constraints, and thus the censored empirical likelihood ratio can be computed. Compare to the SQP, the generalized EM algorithm can handle much larger sample sizes, up to 10,000 and beyond. The memory requirement of the generalized EM algorithm increases linearly with sample size. The computation time is comparable to the sum of two computation problems: the same EL problem but with uncensored data (which has a Lagrange multiplier dual reduction), and the Turnbull's EM for censored data NPMLE.

Many EL papers include examples and simulation results and thus various software are developed. There are two sources of publicly available software for EL: there are Splus codes and Matlab code on the EL web site maintained by Owen but it cannot handle censored data. And there is a package *emplik* for the statistical software R (Gentalman and Ihaka, 1996) written by Mai Zhou, available from CRAN. This package includes several functions that can handle EL computations for right censored data, left censored data, doubly censored data, and right censored and left truncated data.

The package *emplik* also includes functions for computing EL in the regression models discussed in section six. No special code is needed to compute EL in the Cox proportional hazards model, since the EL coincide with the partial likelihood, and many software are available to compute this.

5 Cox proportional hazards regression model

For survival data, the most popular model is the Cox model. It is known for some time that the partial likelihood ratio of Cox (1972, 1975) can be interpreted also as the (profile) empirical likelihood ratio. See Pan (1997), Murphy and van der Vaart (2000).

Let X_i, $i = 1, \cdots, n$ be independent lifetimes and z_i, $i = 1, \cdots, n$ be its associated covariate. The Cox model assumes that

$$h(t|z_i) = h_0(t) \exp(\beta z_i),$$

where $h_0(t)$ is the unspecified baseline hazard function, β is a parameter.

The contribution to the empirical likelihood function from the i^{th} observation, (T_i, δ_i) is

$$(\Delta H_i(T_i))^{\delta_i} \exp\{-H_i(T_i)\},$$

where $H_i(t) = H_0(t) \exp(z_i\beta)$, by the model assumption of Cox model. The empirical likelihood function is then the product of the above over i:

$$EL^c(H_0, \beta) = \prod_{i=1}^{n} [\Delta H_0(T_i) \exp(z_i\beta)]^{\delta_i} \exp\{-H_0(T_i) \exp(z_i\beta)\} \ .$$

It can be verified that for any given β the EL^c is maximized at the so called Breslow estimator, $H_0 = \hat{H}_n^\beta$. Also, by definition, the maximum of $EL^c(\hat{H}_n^\beta, \beta)$ with respect to β is obtained at the Cox partial likelihood estimator of the regression parameter. Let us denote the partial likelihood estimator of Cox by $\hat{\beta}_c$.

Theorem 4. *Under conditions that will guarantee the asymptotic normality of the Cox maximum partial likelihood estimator as in Andersen and Gill (1991), we have the following empirical likelihood ratio result:*

$$-2\log \frac{EL^c(\hat{H}_n^{\beta_0}, \beta_0)}{\sup\limits_{\{\beta,\ H_0\}} EL^c(H_0, \beta)} = I(\xi)(\beta_0 - \hat{\beta}_c)^2 \ , \tag{9}$$

where $I(\cdot)$ is the information matrix as define in Andersen and Gill(1991), ξ is between β_0 and $\hat{\beta}_c$, and the sup *in the denominator is over all β and hazard $H_0 \ll \hat{H}_n$. It then follows easily that (9) $\xrightarrow{\mathcal{D}} \chi^2_{(1)}$ as $n \to \infty$.*

The proof of Theorem 4 was given in Pan (1997). Zhou (2003) further studied the EL inference of Cox model along the lines of the above discussion. He obtained the Wilks theorem of the EL for estimating/testing β when some partial information for the baseline hazard is available. The (maximum EL) estimator of β is more efficient than $\hat{\beta}_c$ due to the extra information on the baseline hazard.

6 Accelerated Failure Time model

The semiparametric accelerated failure time (AFT) model basically is a linear regression model where the responses are the logarithm of the survival times and the error term distribution is unspecified. It provides a useful alternative model to the popular Cox proportional hazards model for analyzing censored survival data. See Wei (1992). Sometimes the AFT models are seen to be even more natural than the Cox model, see Reid (1994).

For simplicity and with a slight abuse of notation, we denote X_i to be the logarithm of the lifetimes. Suppose

$$X_i = \beta^t z_i + \epsilon_i \qquad i = 1, \ldots, n;$$

where ϵ_i is independent random error, β is the regression parameter to be estimated and z_i consists of covariates. Let C_i be the censoring times, and assume that C_i and X_i are independent. Due to censoring, we observe only

$$T_i = \min(X_i, C_i); \quad \delta_i = I_{[X_i \le C_i]}; \quad z_i, \quad i = 1, \ldots, n. \tag{10}$$

For any candidate, b, of estimator of β, we define

$$e_i(b) = T_i - b^t z_i \ .$$

When $b = \beta$, the $e_i(\beta)$ are the censored ϵ_i.

Two different approaches of the EL analysis of the censored data AFT model are available in the literature. The first approach is characterized by its definition of the EL as

$$EL(AFT) = \prod_{i=1}^{n} p_i \ . \tag{11}$$

However, this is a bona fide EL only for iid uncensored data. Similar to Owen (1991), this $EL(AFT)$ is then coupled with the least squares type estimating equations

$$\sum_{i=1}^{n} z_i(T_i^* - \beta^t z_i) = 0$$

where T_i^* is defined by one of the two approaches below.

Synthetic data approach:

$$T_i^* = \frac{\delta_i T_i}{1 - G(T_i)} \ ;$$

or the Buckley-James approach:

$$T_i^* = \delta_i T_i + (1 - \delta_i) E(X_i | T_i, \beta) \ .$$

Both definition of T_i^* are based on the observation that $E(T_i^*) = E(X_i)$. Unfortunately, the censoring distribution function G in the synthetic data approach is unknown and is typically replaced by a Kaplan-Meier type estimator. In the Buckley-James approach the conditional expectation depends on the unknown error distribution and also need to be estimated. These substitution, however, makes the T_i^* dependent on each other and careful analysis show that the log EL(AFT) ratio has a limiting distribution characterized by linear combinations of chi squares, with the coefficients need to be estimated. See Qin and Jing (2001) and Li and Wang (2003) for details.

The second approach of EL for the censored AFT model defines the EL as

$$EL(error) = \prod_{i=1}^{n} p_i^{\delta_i} [1 - \sum_{e_j \le e_i} p_j]^{1-\delta_i} \ . \tag{12}$$

This EL may be viewed as the censored EL for the iid errors in the AFT model. In our opinion, the EL should reflect the censoring and thus this definition of EL is more natural.

Zhou and Li (2004) first note that the least squares estimation equation with the Buckley-James approach can be written as

$$0 = \sum_i \delta_i e_i(b) \left[z_i + \sum_{k<i,\ \delta_k=0} z_k \frac{\Delta \hat{F}_n(e_i)}{1 - \hat{F}_n(e_k)} \right] \tag{13}$$

where $\hat{F}_n$ is the Kaplan-Meier estimator computed from $(e_i,\ \delta_i)$.

They then propose to use the estimation equations

$$0 = \sum_i e_i(b) \frac{z_i + \sum_{k<i,\ \delta_k=0} z_k \frac{\Delta \hat{F}_n(e_i)}{1 - \hat{F}_n(e_k)}}{n \Delta \hat{F}_n(e_i)} \ \delta_i p_i \tag{14}$$

with the $EL(error)$ defined above.

With the aid of Theorem 2, Zhou and Li proved the following theorem.

Theorem 5. *Suppose that in the censored AFT model ϵ_i are iid with a finite second moment. Under mild regularity conditions on the censoring, we have, as $n \to \infty$*

$$-2 \log \frac{\sup EL(error)}{\sup EL(error)} \xrightarrow{\mathcal{D}} \chi^2_{(1)} ,$$

where the sup *in the numerator is taken over $b = \beta$ and all probabilities p_i that satisfy the estimating equations (14); the* sup *in the denominator is taken over $b =$ Buckley-James estimator and all probabilities p_i.*

A multivariate version of this theorem obviously also holds. Simulation study also confirms this. For details please see Zhou and Li (2004). The M-estimator of β for the censored AFT model is also discussed there.

7 Other Applications

Empirical likelihood method is applicable to many other types of censored data. But due to technical difficulties, fewer results are available. Li (1996) studied the EL with left truncated data. Similar results for the left truncated and right censored data should also hold. Murphy and Van der Vaart (1997) describe a general framework in which one possible way of studying the EL is illustrated. In particular, they showed that for doubly censored data, where the lifetimes are subject to censoring from above and below, EL results similar to Theorem 1 also hold. Huang (1996) studied the current status data, also known as case one interval censoring. He demonstrated how to use EL in a proportional hazards model where the responses are current status data, and obtained the Wilks theorem for the EL ratio for testing the regression parameters.

References

Adimari, G. (1997). Empirical likelihood type confidence intervals under random censorship. *Annals of the Institute of Statistical Mathematics*, **49**, 447-466.

Andersen, P. K. and Gill, R.D. (1982). Cox's Regression Model for Counting Processes: a Large Sample Study. *Annals of Statistics*, **10**, 1100-1200.

Buckley, J.J. and James, I.R. (1979). Linear regression with censored data. *Biometrika*, **66** 429-36.

Chen, K. and Zhou, M. (2003). Non-parametric hypothesis testing and confidence intervals with doubly censored data. *Lifetime Data Analysis*, **9**, 71-91.

Chen, S.X., Leung, D.H.Y. and Qin, J. (2003). Information recovery in a study with surrogate endpoints. *Journal of American Statistical Association*, **98**, 1052-1062.

Cox, D. R. (1972). Regression Models and Life Tables (with discussion). *J. Roy. Statist. Soc B.*, **34**, 187-220.

Cox, D. R. (1975). Partial Likelihood. *Biometrika*, **62**, 269-276.

Einmahl, J.H.J. and McKeague, I.W. (1999). Confidence tubes for multiple quantile plots via empirical likelihood. *Annals of Statistics*, **27**, 1348-1367.

Fleming, T. and Harrington, D. (1991). *Counting Processes and Survival Analysis.* Wiley, New York.

Gentleman, R. and Ihaka, R. (1996). R: A Language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5**, 299-314.

Hollander, M., McKeague, I.W. and Yang, J. (1997). Likelihood ratio-based confidence bands for survival functions. *Journal of American Statistical Association*, **92**, 215-226.

Huang, J. (1996). Efficient estimation for the proportional hazards model with interval censoring. *Annals of Statistics*, **24**, 540-568.

Kaplan, E. and Meier, P. (1958). Non-parametric Estimator From Incomplete Observations. *Journal of American Statistical Association*, **53**, 457-481.

Leblanc, M. and Crowley, J. (1995). Semiparametric regression functions. *Journal of American Statistical Association*, **90**, 95-105.

Li, G. (1995). On nonparametric likelihood ratio estimation of survival probabilities for censored data. *Statist. & Prob. Letters*, **25** , 95-104.

Li, G. (1996). Nonparametric likelihood ratio estimation of probabilities for truncated data. *Journal of American Statistical Association*, **90**, 997-1003.

Li, G., Hollander, M., McKeague I.W. and Yang, J. (1996). Nonparametric likelihood ratio confidence bands for quantile functions from incomplete survival data *Annals of Statistics*, **24**, 628-640.

Li, G., Qin, J. and Tiwari, R.C. (1997). Semiparametric likelihood ratio-based inferences for truncated data. *Journal of American Statistical Association*, **92**, 236-245.

Li, G. and van Keilegom, I. (2002). Likelihood ratio confidence bands in non-parametric regression with censored data. *Scandinavian Journal of Statistics*, **29**, 547-562.

Li, G. and Wang, Q.H. (2003). Empirical likelihood regression analysis for right censored data. *Statistica Sinica* **13**, 51-68.

Lu, J.C., Chen, D. and Gan, N.C. (2002). Semi-parametric modelling and likelihood estimation with estimating equations *Australian & New Zealand Journal of Statistics*, **44** 193-212.

McKeague, I.W. and Zhao, Y.C. (2002). Simultaneous confidence bands for ratios of survival functions via empirical likelihood. *Statistics & Probability Letters*, **60** 405-415.

Murphy, S. and van der Vaart, A. (1997). Semiparametric likelihood ratio inference. *Annals of Statistics* **25**, 1471-509.

Murphy, S. and van der Vaart, A. (2000). On profile likelihood. *Journal of American Statistical Association*, **90**, 1399-1405.

Naiknimbalkar, U.V. and Rajarshi, M.B. (1997). Empirical likelihood ratio test for equality of k medians in censored data. *Statistics & Probability Letters*, **34**, 267-273.

Owen, A. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, **75**, 237-49.

Owen, A. (1991). Empirical Likelihood for Linear Models. *Annals of Statistics*, **19** 1725-1747.

Pan, X.R. (1997). *Empirical Likelihood Ratio Method for Censored Data. Ph.D. Dissertation*, University of Kentucky, Department of Statistics.

Pan, X.R. and Zhou, M. (1999). Using one parameter sub-family of distributions in empirical likelihood with censored data. *Journal Statistical Planning and Inference* **75**, 379-392.

Pan, X.R. and Zhou, M. (2002). Empirical likelihood in terms of cumulative hazard function for censored data. *Journal of Multivariate Analysis* **80**, 166-188.

Qin, G.S. and Jing, B.Y. (2001). Empirical likelihood for censored linear regression. *Scand J of Statist*, **28**, 661-673.

Qin, G.S. and Jing, B.Y. (2001). Censored partial linear models and empirical likelihood. *Journal of Multivariate Analysis*, **78**, 37-61.

Qin, G.S. and Tsao, M. (2003). Empirical likelihood inference for median regression models for censored survival data. *Journal of Multivariate Analysis*, **85**, 416-430.

Reid, N. (1994). A conversation with Sir David Cox. *Statist. Sci.* **9**, 439-455.

Ren, J.J. (2001). Weighted empirical likelihood ratio confidence intervals for the mean with censored data. *Annals of the Institute of Statistical Mathematics*, **53**, 498-516.

Shen, X.T., Shi, J. and Wong, W.H. (1999). Random sieve likelihood and general regression models. *Journal of American Statistical Association*, **94**, 835-846.

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.

Turnbull, B. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of Royal Statistical Society, Series, B* **38**, 290-295.

Wang, Q.H. and Jing, B.Y. (2001). Empirical likelihood for a class of functionals of survival distribution with censored data. *Annals of the Institute of Statistical Mathematics*, **53**, 517-527.

Wang, Q.H. and Li, G. (2002). Empirical likelihood semiparametric regression analysis under random censorship. *Journal of Multivariate Analysis*, **83**, 469-486.

Wei, L. J. (1992). The accelerated failure time model: a useful alternative to the Cox regression model in survival analysis. *Statist. Med.* **11**, 1871-79.

Zhou, M. (2002). Empirical likelihood ratio with arbitrarily censored/truncated data by a modified EM algorithm. Tech Report. University of Kentucky. To appear in *Journal of Computational and Graphical Statistics*

Zhou, M. (2003). Cox proportional hazards model with additional information on the baseline. Tech Report. University of Kentucky.

Zhou, M. and Li, G. (2004). Empirical likelihood analysis of the Buckley-James estimator. Tech Report. UCLA.