A REVIEW ON PARAMETRIC AND NONPARAMETRIC EMPIRICAL LIKELIHOOD METHOD BASED ON INTERVAL-CENSORED DATA

Norazelah Zainudin[1], F. A. M. Elfaki[2,*], and M. Yeakub Ali[1]
[1]Department of Manufacturing and Materials Engineering, Faculty of Engineering, International Islamic University Malaysia, Malaysia, 50728, Kuala Lumpur, Malaysia
[2]Department of Science in Engineering, Faculty of Engineering, International Islamic University Malaysia, 50728, Kuala Lumpur, Malaysia
*Corresponding author e-mail: faizelfaki@yahoo.com; faizelfaki@iium.edu.my

ABSTRACT: This paper shall review on the estimation of distribution function for both parametric and nonparametric empirical likelihood method. The focus of this review is on empirical likelihood regression method behavior on censored data and how those methods affect the result of the simulation studies. Moreover, we looking into the case which is involve the partly interval censored data.

Keywords: Censored data; Empirical likelihood; Parametric estimation; Nonparametric estimation; Approximation of variance.

1. INTRODUCTION

Cox Proportional Hazard model has been widely used all over the world. This model is being served on the purpose mostly in survival analysis which usually employed by economist and medical officer for the sake of financial and health monitoring analysis. Numerous studies has been done regarding empirical likelihood and tons of reviews made by the various researchers and their ideas of approaching to the conclusion based on that matter might be different from each other.

The literature on Empirical Likelihood (EL) are increasing rapidly nowadays and it would surely be a perplex job to review each and every existing written paper in details. Therefore, to compress the scope of this review paper into a more understandable and apprehensible reference for further research, this paper will concentrate on the partly interval-censored data since existing studies pay less attention on this area of interest. However, the things that will make this paper more interesting are that it's reviewed from different impendency: parametric and nonparametric.

Based on empirical likelihood method, there are many advantages that it may afford compared to other method. However, in this paper we are focusing on EL method solely in order to see the dissimilarity of yielding results from the previous studies regarding the parametric and nonparametric approach. One of the advantages the EL may provide to us is that the ability of the method to produce confidence region for the particular targeted parameter we are using. They are several authors have been used EL in their study such as; Zhang and Zhao (2013)[1] they have agreed that an EL is one of the nonparametric methods being used to construct the confidence region for the regression studies and being used to make the inferences of major estimating equation with different views since then. While this point is being agreed by Rao (2009)[2] that EL method may provide nonparametric confidence intervals on parameters of interest which likely being produced by traditional parametric likelihood ratio intervals. In addition to that, the confidence interval used by EL could be determined by the data whereby the range of the intervals included is upheld Rao (2013)[3]. Qin and Zhang (2007, 2008),[4,5] they have implemented EL with confidence intervals to interpret the uncertain data of structural differences within the populations which is specially made for mean and distribution function differences. The result found by this research proved experimentally that the confidence region produce by EL works substantially in making conclusion for diverse dissimilarity of two tested groups. Meanwhile, Wong et al. (2009)[6] stated that diagnostic technique based on EL approach was being developed in their paper using partial linear model. Zhao *et al.* (2013)[7] is being used as a function to estimate mean functional missing response that were missing without notice (non-ignorable mechanism) and proved to be a great nonparametric method which has some qualities over Normal Approximation (NA) based methods.

In a simple comparison made between EL and NA, early conclusion that could be made is EL method provide more advantages than NA. As it was being discussed by Ali Jinnah (2007)[8], EL overcome NA method for its capability of providing better coverage probability for small sample sizes while it was agreed by Zhang and Zhao (2013)[9] as they were referring it as the under coverage problem. If the regression parameter in NA method should first be estimated to construct the confidence region, this step could be eliminated in EL and able to yield the confidence region that is conformed to dataset without the need of being symmetric. According to Rao (2009[2], the requirement to measure the standard errors of estimator in Independent and Identically Distributed (I.I.D) cases is abolished in EL and somewhat useful to the analysis if good balanced tail error rates could be obtain compared to NA method.

2. PARAMETRIC

Consider a regression model written by Chen et al. (2007) as follow:

$$Y_i = m(X_i) + \sigma(X_i)e_i, \quad i = 1, 2, \ldots, n, \tag{2.1}$$

where $m(.)$ and $\sigma(.)$ are both unknown function defined over R^d, while X_i and e_i are the weakly dependent variable which e_i is an error with zero mean and variance.

For this parametric regression, we are dealing with kernel estimator of conditional mean function $m(x)$., with K be the *r*th order *d*-dimensional kernel and *h* be the smoothing bandwidth, therefore $K_h(u) = h^{-d}K(\frac{u}{h})$. The Nadaraya-Watson estimator for $m(x)$ could be written as:

$$\hat{m}(x) = \frac{\sum_{i=1}^{n} K_h(x-X_i)Y_i}{\sum_{i=1}^{n} K_h(x-X_i)} \quad (2.2)$$

Now we let θ be the consistent estimator of equation (2,2). Thus, we get:

$$\tilde{m}_{\hat{\theta}}(x) = \frac{\sum_{i=1}^{n} K_h(x-X_i)m_{\hat{\theta}}(X_i)}{\sum_{i=1}^{n} K_h(x-X_i)} \quad (2.3)$$

as the kernel smooth of the parametric estimate having the same kernel and bandwidth as in equation (2.2). Therefore, we can define EL for $m(x)$ smoothed at parametric model $\tilde{m}_{\hat{\theta}}(x)$ written as:

$$L\{\tilde{m}_{\hat{\theta}}(x)\} = \max \prod_{i=1}^{n} p_i(x) \quad (2.4)$$

with respect to $\sum_{i=1}^{n} p_i(x) = 1$ and as we define $Q_i(x) = K_h(x-X_i)\{Y_i - \tilde{m}_{\hat{\theta}}(x)\}$, equation (2.4) also subjected to $\sum_{i=1}^{n} p_i(x)Q_i(x) = 0$. A standard deviation of EL show us that the optimal weight of p_i is:

$$p_i(x) = \frac{1}{n}\left\{1 + \lambda(x)K_h(x-X_i)\{Y_i - \tilde{m}_{\hat{\theta}}(x)\}\right\}^{-1} \quad (2.5)$$

in which $\lambda(x)$ is the solution of:

$$\sum_{i=1}^{n} \frac{K_h(x-X_i)\{Y_i - \tilde{m}_{\hat{\theta}}(x)\}}{1+\lambda(x)K_h(x-X_i)\{Y_i - \tilde{m}_{\hat{\theta}}(x)\}} = 0 \quad (2.6)$$

Therefore, if we substitute the optimal weight into the empirical likelihood, hence we will obtain:

$$L_n(\theta) = \prod_{i=1}^{n} \frac{1}{n}\frac{1}{1+\lambda(x)K_h(x-X_i)\{Y_i - \tilde{m}_{\hat{\theta}}(x)\}}, \quad (2.7)$$

As EL is maximized at $p_i(x) = \frac{1}{n}$, the log-likelihood of equation (2.7) could be obtain by assigning log operation to these equation to get:

$$\log\{L_n(\theta)\} = -\sum_{i=1}^{n} \log[1+\lambda(x)K_n(x-X_i)\{Y_i - \tilde{m}_{\theta}(x)\}\} - n\log(n) \quad (2.8)$$

Following the standard parametric likelihood convention, we finally can define from equation (2.8), the EL ratio for parametric regression. It is to mention that other dual problem that involves in minimizing the first part of equation (2.8) should be ignored since we note that the profile likelihood $\prod_{i=1}^{n} p_i$ achieve it maximum n^{-n} when all the weight of p_i is equal to n^{-1} for $i = 1, \dots, n$. In addition, based on Chen et al., (2007)[10], showed that, the basic idea of EL is to maximize a product of probability weight that is being used for observations under certain limit. Therefore, the EL ratio is:

$$r_n(\theta) = -2\log\left[\frac{L_n(\theta)}{L_n(\hat{\theta})}\right] = 2\sum_{i=1}^{n}\log\{1+\lambda(x)K_n(x-X_i)\{Y_i - \tilde{m}_{\theta}(x)\} \quad (2.9)$$

Chen et al (2009), they stated that Wilk's theorem is the key point for parametric regression successfulness. However, this is depending on how we play with this theorem to achieve our main result since Wilk's theorem could be parametric and non-parametric. In the article written by Huang (2012)[11], earlier he induced the parametric solution in order to construct the confidence region for partially linear error-in-function model, and later on, he derived the non-parametric version Wilk's theorem in his research. Therefore, the study might not be claim as a parametric regression since in the simulation, he used those derived theorem to obtain the result.

Nevertheless, as Wilk's theorem considered as important in determining the EL ratio, this paper hereby takes the opportunity to appreciate the works of non-parametric Wilk's theorem under certain regularity condition. If we used the notation of $r_n(\theta)$ as the ratio for EL, therefore with regard to Wilk's theorem and in context of parametric regression, we get:

$$r_n(\theta_0) \xrightarrow{d} \chi_p^2 \ as \ n \to \infty \quad (2.10)$$

However, to appreciate more on Wilk's theorem to prove it validity in parametric regression, few steps of nonparametric likelihood need to be considered. In general, the expansion of λ in equation (2.6) at β_0, true value of β parameter, and specifying its order magnitude are the initial steps taken in considering EL approach. From those steps, the current parametric regression could be written as:

$$\lambda = O_p(n^{-1/2}) \quad (2.11)$$

With regards to (2.11) and (2.6), and the simplification of $Z_{ni} = (x)K_h(x-X_i)\{Y_i - \tilde{m}_{\hat{\theta}}(x)\}$, the equation of (2.6) can be inverted to:

$$n^{-1}\sum_{i=1}^{n} Z_{ni}(1-\lambda^T Z_{ni}) + n^{-1}\sum_{i=1}^{n} Z_{ni}\frac{\lambda^T Z_{ni}Z_{ni}^T\lambda}{1+\lambda^T Z_{ni}} = 0 \quad (2.12)$$

From all of the above notation on non-parametric version Wilk's theorem, an EL confidence region could be construct for β_0. With the nominal level of confidence $1-\alpha$, the confidence region could be written as:

$$I_{1-\alpha} = \{\beta : r_n(\beta) \le \chi_{p,1-\alpha}^2\} \quad (2.13)$$

with $\chi_{p,1-\alpha}^2$ is the ($1-\alpha$) quartile of χ_p^2 distribution, following the condition of $P\{\beta_0 \in I_{1-\alpha}\} \to 1-\alpha$ as

$n \to \infty$. Pursuing the Bartlett correction in order to obtain the confidence region of parametric regression,

also considering the $I_{1-\alpha}$ confidence region having the coverage error of n^{-1}, therefore the adjusted confidence region following Barlett correction technique with some improvement, is written as:

$$P\{r_n{}^*(\beta_0) \le \chi^2_{p,1-\alpha}\} = 1 - \alpha + O(n^{-2}) \quad (2.14)$$

3. NONPARAMETRIC REGRESSION

The nonparametric offered different path of solving empirical likelihood problems. Consider the linear transformation model proposed by Zhang et al. (2013)[10] is as follow:

$$u(T) = Z^T\beta + \epsilon \quad (3.1)$$

where $u(T)$ denoted as a strictly unknown increasing function while β is the $p\ x\ 1$ vector of regression parameters and ϵ is the pre-specified distribution function of *F*. Commonly the proportional hazard model for this method could be written as follow:

$$F(t) = 1 - exp\{-exp(t)\} \quad (3.2)$$

However in other model, equivalent to equation (3.2) as when F is increasing, the notation of the model could be written as:

$$g\{S_z(t)\} = u(t) - z^\tau\beta \quad (3.3)$$

where $g^{-1}(s) = 1 - F(s)$ and S_z act as the survival function of T given that $Z = z$. Similar to parametric regression, non-parametric comes with the assumption of $E(\varepsilon_i) = 0$, and $Var(\varepsilon_i) = \sigma^2.m_i$, with all $m_i > 0$. As for m_i is assumed to have a function in form of:

$$m_i = \mathrm{m}(z_i, \delta),\ i = 1, 2, \ldots., n \quad (3.4)$$

According to Zhao et al. (2013), this m_i act as the consistent estimator of $m_i(X_i) = E(Y_i|X_i, \delta_i = 1)$. While taking the advantage on the proposed nonparametric regression estimator by Kim and Yu (2011), therefore, $\hat{m}_0(x)$ could be written as:

$$\hat{m}_0(x) \coloneqq \hat{m}_0(x;\gamma) = \sum_{i=1}^n \omega_{i0}(x;\gamma)Y_i, \quad (3.5)$$

Identically, Wong et al. (2009) and Wang et al. (2003)[12] have applied the same technique involving these regression estimators which require the weigh of the estimator that could be written as:

$$W_{nj}(t) = \frac{K\left(\frac{t-T_j}{h_n}\right)}{\sum_{j=1}^n K\left(\frac{t-T_j}{h_n}\right)} \quad (3.6)$$

Here, it is to mention that $K(.)$ is defined as the Kernel function that has h_n with a bandwidth approaching zero and $nh_n \to \infty$ as $n \to \infty$. Now, to satisfy the empirical log-likelihood estimator for nonparametric regression, let us define $\tilde{Z}_i = \tilde{X}_i(\tilde{Y}_i - \tilde{X}_i^\tau\beta)$ and use Lagrange multiplier method to find the optimal value of p_i. This can be shown as:

$$p_i = \frac{1}{n}\frac{1}{1+\lambda^\tau\tilde{X}_i(\tilde{Y}_i - \tilde{X}_i^\tau\beta)}, \quad (3.7)$$

where the λ here is the solution of the equation

$$\frac{1}{n}\sum_{i=1}^n \frac{\tilde{X}_i(\tilde{Y}_i - \tilde{X}_i^\tau\beta)}{1+\lambda^\tau\tilde{X}_i(\tilde{Y}_i - \tilde{X}_i^\tau\beta)} = 0, \quad (3.8)$$

Therefore, the empirical likelihood evaluated at β after the substitution of optimal weight into the equation became:

$$L_n(\beta) = \prod_{i=1}^n \frac{1}{n}\frac{1}{1+\lambda^\tau\tilde{X}_i(\tilde{Y}_i - \tilde{X}_i^\tau\beta)}, \quad (3.9)$$

Hence, the empirical log-likelihood ratio can be written as:

$$l_n(\beta) = -\sum_{i=1}^n \log[1 + \lambda^n \tilde{X}(\tilde{Y}_i - \tilde{X}_i^T)] - n\log(n) \quad (3.10)$$

As the overall empirical likelihood is maximized at $p_i = 1$, thus the definition of log-likelihood ratio of β is written as:

$$r_n = 2\sum_{i=1}^n \log[1 + \lambda^\tau\tilde{X}_i(\tilde{Y}_i - \tilde{X}_i^\tau\beta)] \quad (3.11)$$

The distribution of $\ell_n(\beta)$ is somehow determined by the asymptotic distribution of the equation $\frac{1}{\sqrt{n}}\sum_{i=1}^n \tilde{Z}_i$ and the constituency of $\frac{1}{n}\sum_{i=1}^n \tilde{Z}_i\tilde{Z}_i^\tau$. Since we define that

$$Z_i = (X_i E[X_i/T_i])(Y_i - E[Y_i/T_i] - (X_i - E[Y_i/T_i]^\tau \beta_i$$

, therefore if we can prove that $\frac{1}{\sqrt{n}}\sum_{i=1}^n \tilde{Z}_i = \frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i + o_p(1)$ and $\frac{1}{n}\sum_{i=1}^n \tilde{Z}_i\tilde{Z}_i^\tau = \frac{1}{n}\sum_{i=1}^n Z_iZ_i^\tau + o_p(1)$, hence $\ell_n(\beta)$ will have the same distribution as mention earlier.

There are some assumptions or conditions that should be made in order for the equation to be distributed asymptotically as required. However, different studies might provide dissimilar assumption based on the parameters, approach and limits they are using. Under some circumstances, the non-parametric estimation does not have any effect on the asymptotic result of the empirical log-likelihood ratio. Therefore, when those conditions are fulfill, we have the equation of:

$$l_n(\beta) = \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{Z}_i\right)^{\tau}\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{Z}_i\,\tilde{Z}_i^{\tau}\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{Z}_i\right) + O_p(1) \quad (3.12)$$

Again, under the same condition applied, when the β parameter have the true value of β_0, hence the $\ell_n(\beta_0)$ has the distribution of χ_p^2 which can be written as:

$$P\left(\tilde{\ell}_n(\beta) < c_\alpha\right) = 1 - \alpha + o(1) \quad (3.13)$$

with $P\left(\chi_p^2 \le c_\alpha\right) = 1 - \alpha$.

Equation (3.13) it's not only has the ability to test the hypotheses of $H_0: \beta = \beta_0$, but also help to construct the confidence region β for the non-parametric method. Therefore, we let $I_\alpha(\beta) = \{\beta: \tilde{l}_n(\beta) \le c_\alpha\}$ in which $I_\alpha(\beta)$ gives the approximate confidence region for β with asymptotically correct coverage probability $1 - \alpha$. Hence, the confidence region is stated as:

$$\mathcal{R} = \{\beta: \tilde{l}_n(\beta) \le \chi_p^2(\alpha)\} \quad (3.14)$$

4. RESULT COMPARISONS

4.1. PARAMETRIC RESULT

Parametric and non-parametric results are to be compare in this section. In general, there are a lot more non-parametric-regression-base research had been done so far compared to parametric studies. However, there are some parametric studies had been done, therefore included in this review to observe the differences between those two regression type.

In the study conducted by Chen et al. (2007), the result obtained from the simulation studies are done on sample sizes of $n = 300$ and $n = 500$ whereby the bandwidth, h value was remain constant at first simulation study and keep changing for the next simulation. The study was done to see the performance of their EL method and compare it with others adaptive test method. According to this result**,** it shows that the proposed EL method have a better power over the method used by Horowitz and Spokoiny (2001) [15], HS test). However, the simulation later being further to the adaptive EL test proposed in Chen. et al (2003[17], CHL test) to see the performance of it based on single bandwidth simulation. The result shown acknowledged the powerfulness of CHL test over the their methods and HS test.

Similar simulation was done but with h value being change where the result reported that both the their EL test and HS test express a good approximation to the nominal level of $\alpha = 0.05$. Which is indicates, it might due to internal studentization of EL method of proposed method which enhanced its power over HS test.

Likewise, regression type was later being employed by Elfaki. et al (2012), for the study of partly interval-censored data with the application to AIDS studies. However, the bandwidth for the simulation study was not stated and the result yield was has similarities with simulation done by Kim (2003).

The study later on being subjected to real world that are onto HIV/AIDS data for the Elfaki et al (2012)[14] and onto Diabetis Mellitus data for Kim (2003)[15]. Surprisingly, the result obtained for the application studies were similar to each other in which the diagnosis had been fixed in terms of gender and age whereby male and very young patient has lower hazard rate. The latter research therefore strongly support the result obtained by the former study that is the more exact data available, the better estimates are and since smaller bias is obtained with larger sample sizes.

4.2. NONPARAMETRIC RESULT

There are quite a number of studies that have been sum up in this review for the purpose of finding the similarities in the method and equations used but on the other hand to analyze the differences amongst those studies in terms of their result, either it yield most likely conclusion or resolve for other. In addition, to meet the comparison procedure, there are many elements that will be considered such as the bandwidth, h_n and the sample size involve n other than the type of censoring which consider as important for this review paper.

In the study conducted by Wang et al. (2003) that implied non-parametric approach which consider the kernel function $K(t)$, involving the biweight kernel function below:

$$K(x) = \frac{15}{16}(1 - x^2)^2, \quad |x| \le 1 \quad (4.1)$$

share the same biweight kernel function with the study held by Wong et al. (2009)[16]. However, as mention before, bandwidth of each study will be count for comparison purposes. For Wong et al (2009[16], under the selection of bandwidth h_n to be $2t_{sd}n^{-1/5}$, in which t_{sd} is the standard deviation of t_i.

For respected simulation study, the processes for both models, it can be seen that empirical likelihood ratio test confirmed that it has some mild size disturbance mainly for the small sample size which likely caused by the limiting distribution of χ^2. This is referred back to the confidence region established by the researcher earlier which later gave slightly distortion towards the result obtained.

The result of Empirical Likelihood Ratio test (ELR) later being compared to the other method of Dette and Munk's (D-M) and showed that it is more powerful than D-M. The author stated, in their opinion, since D-M method cast off necessitate of estimating non-parametric function cause it to loss some power in the test. However, this is not the main concern in this paper to point it out but it always good start to associate our studied cases with others.

In comparison with Wong (2009) study, Wang et al. (2003) which has similar biweight kernel function aforementioned however produced slightly different result in their simulation study. In this study, there are three different bandwidth being used that are $(nlogn)^{-1/2}, (nlogn)^{-1/3}$ and

$5(nlogn)^{-1/3}$. While the sample size are smaller compared to Wong et al. (2009)[17] simulations that are 10, 20 and 50 and the nominal levels α are 0.10 and 0.05. From their results it can concluded that for both methods, studentize-t and empirical likelihood, the coverage probabilities is increasing as the sample size increase. On the bright side, empirical likelihood method seems to surpass studentize method, particularly for the small sample size. However, the capability to outperform decrease as the sample size gets larger. Out of three bandwidth used in the simulation, the third bandwidth seems to have better result since the coverage probabilities for both method have small gap in number.

Here, it can be conclude that for both non-parametric studies explained, with the same biweight kernel function value, the result yield the similar problem that is the distortion in the reading of smaller sample size. Although the bandwidth used for both studies seems to be different, it clarifies that the same difficulties will occur when dealing with kernel function. Though the importance of choosing proper bandwidth for the simulation purpose is made crucial, but the problem will not be address in this review.

Other simulation study on non-parametric regression is one held by Zhang et al. (2013)[18] in which they used other method than kernel function that is pseudo empirical likelihood approach proposed by previous researcher. One thing that is similar to aforementioned studied is that its lingered to Lagrange multiplier to obtain the confidence region. In this study, the sample sizes is limit to 30, 50 and 100 while the nominal level $\alpha = 0.5$. This is to mention that the sample sizes are not too big compared to Wong et al. (2009) study and the simulation are based on simple work since the nominal level only set once. However, from their result, clearly we have that EL and JEL that have closer value of nominal level compared to NA method. This is, from the author opinion, due to the widening in confidence intervals of the simulation studies which later lead NA to underestimate the standard error of regression parameter estimator. Further, the author stated that due to the approximation of variance adopted by EL and JEL, the estimation bias did not completely abolish in this matter. This is why EL and JEL seems to be overpowering the NA method. Therefore, once again, it showed that EL method that is used in terms of non-parametric regression help in solving the problem that NA brought up.

5. CONCLUSION

The determination of parametric and non-parametric regression type in our research is made important since both of the regression type adhere different approach and solved by variety of method exist. The result obtained by both type might be similar or may have come out to the same conclusion at the end, but the significance of choosing the right regression type may help in obtaining more accurate and precise data reading. Choosing the right bandwidth is also an important issue when dealing with simulation study, however some of the researcher does not stress on this matter but rather focusing on the main result. The confidence region constructed made crucial to the study since it is determined by the nominal level.

Some of the research did pursue the parametric or non-parametric regression type from start to the end of study. But some of it might start as parametric but non-parametric methods are being introduced half-way through the research. This is might due to some appreciation they did to use the theorem they want to use and made the research be more reliable. Therefore, in conclusion, this review paper has made some effort to expand the knowledge of parametric and non-parametric regression on empirical likelihood based on interval-censored data with the ability to show the comparison of their result in order to differentiate of both types. However, in our future research we will investigate how the non-parametric regression plays it role in determining the differences adhere by varying the method such as Normal Approximation and Empirical Likelihood method and propose a new method which may improve both of those methods if the case of losing in power is to occur.

REFERENCES

Chen S. X., Gao J., (2007). *An adaptive empirical likelihood test for parametric time series regression models.* Journal of Econometrics. 141: 950-972

Chen S. X., Keilegom I. V., (2009). *A review on empirical likelihood methods for regression.* Test (2009) 18: 415–447

Chen, S.X., Hardle W., Li, M., 2003. *An empirical likelihood goodness-of-fit test for time series*. Journal of the Royal Statistical Society Series B 65, 663–678

Elfaki F. A. M., Azram M., Usman M., (2012). *Parametric Cox's Model for Partly Interval Censored Data with Application to AIDS Studies.*Int. Journal of Applied Physics and Mathematics. 5: 352-354

Horowitz, J.L., Spokoiny, V.G., 2001. *An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative*. Econometrica 69, 599–632

Huang J. (1999). *Asymptotic Properties of Nonparametric Estimation Based on Partly Interval-Censored Data.* StatisticaSinica. 9: 501-519

Huang Z., (2012). *Empirical likelihood for the parametric part in partially linear errors-in-function models.* Statistics and Probability Letters. 82: 63-66

Jinnah A., (2007). *Inference for Cox's Regression Model via a New Version of Empirical Likelihood.* Mathematics Theses. 1-47

Kim, J. K., Yu, C.L., 2011. *A semiparametric estimation of mean functionals with nonignorable missing data*. Journal of the American Statistical Association. 106, 157–165.

Kim J. S., (2003). *Maximum likelihood estimation for the proportional hazards model with partly interval-censored data.* J. R. Statist. Soc. B. 65:489–502

Qin Y., Zhang S., (2008). *Empirical likelihood confidence intervals for differences between two datasets with missing data.* Pattern Recognition Letters. 29:803–812

Wang Q. H., Jing B. Y., (2003). *Empirical likelihood for partial linear models.* Ann. Inst. Statist. Math. 55: 585-595

Wei C., Luo Y., Wu X., (2012). *Empirical likelihood for partially linear additive errors-in-variables models.* Stat Papers. 53: 485–496

Wong H., Liu F., Chen M., Ip W. C., (2009). *Empirical likelihood based diagnostics for heteroscedasticity in partial linear models.* Computational Statistics and Data Analysis. 53: 3466-3477

Wu C., Rao J.N.K., (2010). *Bootstrap procedures for the pseudo empirical likelihood method in sample surveys.* Stat Prob Letters. 80: 1472-1478

1. Zhang Z., Zhao Y., (2013). *Empirical likelihood for linear transformation models with interval-censored failure time data.* Journal of Multivariate Analysis. 116: 398-409

Zhao H., Zhao P. Y., Tang N. S., (2013). *Empirical likelihood inference for mean functionals with nonignorably missing response data.* Computational Statistics and Data Analysis. 66: 101-116