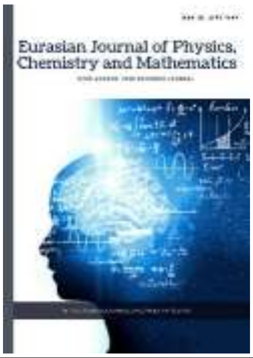

Statistical analysis of some random truncation models

Dr.Hussain Ali. Abbdllah. Aleagobe

Department of Economics
Faculty of Administration and Economics University of Thi-Qar,Iraq
hussainaliabbed@yahoo.com

ABSTRACT

Statistical analysis of some method with empirical Statistical information collected from an experimental distribution of a data set and in many fields such as statistical mechanics and economics has attracted the attention of many authors. The empirical (E), the unknown probabilities it estimates, are always negative. Unknown parameters estimated from this method have a bounded influence function. In this paper, we investigate the empirical (E) method in the truncated data and we prove that it has an asymptotic distribution, a weighted chi-square. We show that when the data is contaminated, empirical (E) method behaves better than empirical likelihood (EL) method and has a better coverage probability.

Keywords:

Abstract

1. Introduction

Survival data are incomplete in many studies and are generally presented as truncated data. This data is of great interest in various fields, including astronomy and economics. suppose $\{(X_i,Y_i)\}$ be independent and identically distribution (*i.i.d*). In the random truncation, $\{(X_i,Y_i), i = 1,..,n\}$ is observed only when $X_i \geq Y_i$. $\{(U_i,V_i)\}$ are a sub sequence of $\{(X_i,Y_i)\}$ that is also an *i.i.d.* sequence. Suppose X and Y have F and G distributions, respectively. The truncated parameter, denoted by α, is defined as follows:

(1.1) $$\alpha = P(X \geq Y) = \int G(s)dF(s)$$

If F_n and G_n are the limit estimates of F and G, respectively, we show the α estimate with α_n and define it as follows.

(1.2) $$\alpha_n = \int G_n(s)dF_n(s).$$

Key words and phrases. empirical (E), Empirical likelihood, weighted chi-square, asymptotic distribution . Email addresses: Hussain.phd@utq.edu.iq. hussainaliabbed@yahoo.com

For any cumulative distribution function F, suppose (a_F,b_F) be the confine of F defined by

$$a_F = \inf\{x : F(x) > 0\} \qquad b_F = \sup\{x : F(x) < 1\}.$$

Assuming $a_G < b_F$, then we can be sure that $\alpha > 0$. Under random truncation, assume $F*$ and $G*$ have the marginal distributions of U_i and V_i. From Woodroof (1985) we know

(2.3) $$F^*(x) = P(U_i \leq x) = \frac{1}{\alpha_0}\int_{-\infty}^{x} G(s)dF(s),$$

(2.4) $$G^*(x) = P(V_i \le x) = \frac{1}{\alpha_0}\int_{-\infty}^{x}(1-F(s))dG(s),$$

and

(2.5) $$R(x) = G^*(x) - F^*(x) = \frac{1}{\alpha_0}G(x)(1-F(x)).$$

The empirical distributions $F*(x)$,$G*(x)$ and $R(x)$ are defined by

$$F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} I(U_i \le x), \qquad G_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} I(V_i \le x)$$

and

(2.6) $$R_n(x) = G_n^*(x) - F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} I(V_i \le x \le U_i)$$

The product-limit estimators for $F(x)$ and $G(x)$ are introduced as follows:

$$F_n(x) = 1 - \prod_{s\le x}\{1 - \frac{F_n^*(s)}{R_n(s)}\}, \qquad G_n(x) = \prod_{s>x}\{1 - \frac{G_n^*(x)}{R_n(x)}\}.$$

Under truncate sampling, the distribution function estimate was obtained by Woodroofe (1985) and the properties of this estimate were also studied. Kaplan-Meyer estimate uses the MLE for nonparametric and semi-parametric models. For randomly truncated data, the MLE estimation is used to the truncation product-limit in nonparametric models. Wang (1989) introduced the MLE for estimating truncated data in semi-parametric models and the large-sample properties of the estimate are presented. the proportional hazards model for censorship and truncated data have studied by Alioum and Commenges (1996) and in this study have attained a suitable method for this model. under left truncated data, He and Yang (2003) proposed the estimation of regression parameters . Ould-Said and Mohamed Lemdani (2006) introduced a new kernel estimator of the regression function for truncated data and they constructed the pointwise and uniform strong consistency and offered a rate of convergence of the estimate. Winfried and Wang (2008) proved the central limit theorem for the truncated data. This paper is organized as follows. In a linear model with left-truncated data, the weighted composite quantile regression considered by Yao et al (2018). The adaptive penalized procedure for variable choice is offered. The asymptotic normality and oracle property of the resulting estimators are also determined.

This article is organized as follows. In Section 3, an empirical log-likelihood ratio is concluded and its asymptotic distribution is demonstrated that be a weighted chi-square. In Section 4, a simulation study is given to compare the performance of the proposed EE method to the introduced EL method for truncated data in terms of coverage probability and length of the interval. Proof of the main results have been deferred in the last Section

.

3. Main results

In this section, we introduce EE and EL methods. Both methods prove to have a weighted asymptotic chi-square distribution.

3.1. **Empirical entropy. Empirical** E method the unknown probabilities it estimates are ever nonnegative and unknown parameters, which are influence functions with abound is estimated. According to the evidence from Monte Carlo, tests based on the EE method with a limited number of samples have better characteristics than the EL method.

For a discrete probability distribution p on the countable set $x_1,x_2,...$, with $p_i = p(x_i)$, the entropy of p is shown as

$$h(p) = -\sum_{i\ge1} p_i \log p_i.$$

Then, in a continuous probability density function $p(\cdot)$ on an interval I, its entropy is defined as
$h(p) = \int p(x)\log p(x)dx.$

I

Monti and Ronchetti (1993) expressed the relationship between the EE method and the EL method. they presented that EE method provides very accurate inference in independent samples. Efron (1981) showed that the EE method is related to biased bootstrap. Test the general parametric hypothesis in time-series regressions using the EE method has performed by Bravo (2005) and used blocking techniques to receive a weak dependence of the observations. Zhao et al (2015) used the EE method in the right-censored data and showed that if the data is contaminated and right-censored, they have a better coverage probability than the EL method. From (2.3), we have

(3.1) $$\int (1 - \frac{\alpha}{G(s)}) dF^*(s) = 0$$

For *i.i.d.* random variable

(3.2) $$\psi_{ni} := 1 - \frac{\alpha}{G(s)}$$

Now consider the empirical entropy at the true value α_0,

$$\mathcal{R}(\alpha_0) := \sup\Big\{ -\sum_{i=1}^{n} p_i \log p_i : \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i \psi_{ni} = 0, p_i \geq 0, 1 \leq i \leq n \Big\}.$$

By Lagrange multiplier, we get

$$p_i = \exp(-\widehat{\lambda}_n \psi_{ni}) / \sum_{j=1}^{n} \exp(-\widehat{\lambda}_n \psi_{ni}), \quad i = 1, ..., n$$

(3.3) $$f(\lambda) = \frac{1}{n} \sum_{j=1}^{n} \psi_{ni} \exp(-\lambda \psi_{ni}) = 0$$

<u>Define</u>

(3.4) $$\rho(\widehat{\psi}_n) = \sum_{j=1}^{n} \exp(-\widehat{\lambda}_n \psi_{ni})$$

it is seen that $\widehat{\mathcal{R}}(\alpha_0) - \log \rho(\widehat{\psi}_n)$). we define the empirical entropy difference at the true value α_0 by

(3.5) $$\Delta\mathcal{R}(\alpha_0) = \widehat{\mathcal{R}}(\alpha_0) - \log n = \log(\frac{\rho(\widehat{\psi}_n)}{n}).$$

According to theorem (3.2) of Newey and Smith (2004) and entropy concentration theorem of Jaynes (1982), the following two adjusted empirical entropy differences are defined,

(3.6) $$Z1(\alpha_0) = -2n\{\exp[\Delta R(\alpha_0)] - 1\}$$

(3.7) $$Z2(\alpha_0) = -2n\Delta R(\alpha_0),$$

Correspondingly, $Z_1(\theta)$ is called the adjusted Newey-Smith empirical entropy difference at θ and $Z_2(\theta)$ the adjusted Jaynes empirical entropy difference at θ. We have the following theorem about $Z1(\alpha_0)$ and $Z2(\alpha_0)$.

Theorem 1. *Let F and G be continuous and satisfy the following moment conditions*

(3.8) $$\int_{-\infty}^{b_G} \frac{dG(s)}{(1 - F(s))} < \infty, \qquad \int_{a_F}^{\infty} \frac{dF(s)}{G(s)} < \infty.$$

Then

(3.9) $$\{\mathcal{Z}_1(\alpha_0) \quad and \quad \mathcal{Z}_2(\alpha_0)\} \xrightarrow{\mathcal{D}} \kappa\chi_1^2,$$

where χ_1^2 *is the standard* χ^2 *variable with one degree of freedom, and*

(3.10) $$\kappa = \frac{\delta_1^2}{\delta_2^2},$$

δ_1^2 *and* δ_2^2 *are obtained from the following equations:*

(3.11) $$\delta_1^2 = \int_{-\infty}^{x} \frac{dF(s)}{R(s)\overline{F}(s)} + \int_{x}^{\infty} \frac{dG(s)}{R(s)G(s)} - \frac{1}{R(x)} + 2\alpha_0 - 1, \qquad x \in (a_G, b_F),$$

(3.12) $$\delta_2^2 = \int_{a_F}^{\infty} \frac{dF(s)}{G(s)} - 1.$$

Proof. See the Proofs.

In order to obtain the confidence interval, it is necessary to use estimate (3.10). As a result, according to equation (3.13) and (3.14), we have:

(3.13) $$\hat{\delta}_1^2 = \int_{-\infty}^{x} \frac{dF_n(s}{R_n(s)\overline{F}_1} \quad \hat{\delta}_2^2 = \int_{a_F}^{\infty} \frac{dF_n(s)}{G_n(s)} - 1,$$

(3.14). χ_1^2

3.2. **Empirical likelihood.** The empirical likelihood method is a non-parametric inference method that has very interesting properties and this method was first introduced by Owen (1990).It is not easy to computation of the empirical likelihood ratios with censored/truncated data and parameter of mean Therefore, Zhou (**?**) used the EM algorithm to obtain the tests and the confidence interval. The most well-known method for interval estimation of probabilities for randomly truncated data is the asymptotic normal method, but this method has a fundamental problem and that is that the given distance may be out of range [0,1] and is not satisfactory in small samples. Li (1995) proposed another way to obtain this distance, which is to use a conditional nonparametric likelihood ratio. It also demonstrated a better small-sample performance in our simulation studies. This approach is generalized to obtain confidence intervals for the ratio of two probabilities. Shen and Hi (2006) used the empirical likelihood method to create a confidence interval for the truncation parameter in the random truncation model and showed that it has an asymptotic distribution of weighted chi-square.

Suppose **p** = ($p_1,...,p_n$) be a probability vector for which $\sum_{i=1}^{n} p_i = 1$ and $p_i \geq 0$ for each i. for $1 \leq i \leq n$ at a stable time t. So, the evaluated EL at $\mu_0(t)$, is given by

$$\psi_{ni}(t)) := \sup\left\{\prod_{i=1}^{n} np_i : \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i\psi_{ni}(t) = 0\right\}.$$

By using the Lagrange multiplier method, we get

$$p_i = \{n(1 + \beta(t)\psi_{ni}(t))\}^{-1}, \qquad i = 1,...,n,$$

where $\beta(t)$ is from solving of

(3.15) $$\frac{1}{n}\sum_{i=1}^{n} \frac{\psi_{ni}(t)}{1 + \beta(t)\psi_{ni}(t)} = 0.$$

It is explicit that $\beta(t)$ and $\psi_{ni}(t)$ are functions of t. So that, they are considered and calculated at a constant and optional time t_0, where $0 \leq t_0 < \tau$. Therefore, we utilized of $\beta(t_0)$ and $\psi_{ni}(t_0)$.

We know that $\prod_{i=1}^{n} p_i$, subjecting to the condition $\sum_{i=1}^{n} p_i = 1$, its maximum at $p_i = n^{-1}$ is n^{-n}. So, the EL ratio for ψ_{ni}, is defined by

(3.16) $$R(\psi_{ni}) := \prod_{i=1}^{n}(np_i) = \prod_{i=1}^{n}\{1 + \beta(t)\psi_{ni}(t)\}^{-1},$$

Hence the empirical log-likelihood ratio is given by

(3.17) $$R'(\psi_{ni}(t)) = -2\log R(\psi_{ni}(t)) = 2\sum_{i=1}^{n}\log\{1 + \beta(t)\psi_{ni}(t)\},$$

where $\beta(t)$ is the solution of (3.15).

Let $X_1,...,X_n$ be independent random variables with common distribution F. Let $\mu = E(X_i)$, and suppose that $0 < Var(X_i) < \infty$. Then $R'(\alpha_0)$ converges in distribution to as $n \to \infty$, that as

(3.18) $$R'(\alpha_0) \xrightarrow{\mathcal{D}} \kappa\chi_1^2,$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution. See the proof (3.18) in theorem (2.1) of Shen and He (2006).

4. Simulation studies

In the simulation, the contaminated lifetime X is a $(1 - m/n)\Gamma(2,1) + (m/n)\Gamma(2,5)$ random variable, where $\Gamma(\alpha,\beta)$ is the Gamma distribution with shape parameter α and scale parameter β and $\Gamma(2,5)$ is the contaminating distribution, *m* is the number of contaminated data and *n* is the sample size. That means there are *m* lifetimes contaminated by Γ(2,5) distributed random variables within *n i.i.d.* lifetimes. In the simulation, we use *m* = 0,1,2,...,7 for sample size *n* = 20,50,100. The process is repeated for *N* = 5000 times and the coverage ratios are calculated by using the *N* data sets. In Table 1, NS-EE is used for the adjusted Newey-Smith empirical entropy difference, AJ-EE is used for the adjusted Jaynes empirical entropy difference, EL is used for empirical likelihood , c.p is used for the coverage probability and Δ is used for the average length. The simulation results reveal that, in terms of coverage probability, for m= 0,1,2, the empirical likelihood method performs better than that of the empirical entropy method. For m = 3,4,5,6,7 the empirical entropy method performs better than that of the empirical likelihood method.

m	n	c.p.NS-EE	$\Delta.NS-EE$	c.p.AJ-EE	$\Delta.AJ-EE$	c.p.EL	$\Delta.EL$
	20	0.762	0.481	0.785	0.452	0.812	0.391
0	50	0.775	0.489	0.791	0.459	0.817	0.397
	100	0.782	0.498	0.805	0.473	0.831	0.403
	20	0.773	0.488	0.789	0.455	0.805	0.407
1	50	0.787	0.495	0.793	0.462	0.807	0.415
	100	0.794	0.507	0.809	0.476	0.819	0.427
	20	0.786	0.492	0.792	0.461	0.796	0.409
2	50	0.798	0.502	0.796	0.465	0.801	0.411
	100	0.803	0.511	0.810	0.481	0.809	0.419
	20	0.791	0.501	0.799	0.469	0.783	0.421
3	50	0.802	0.511	0.809	0.471	0.791	0.435
	100	0.811	0.513	0.815	0.485	0.801	0.443
	20	0.804	0.509	0.803	0.474	0.769	0.435
4	50	0.811	0.519	0.811	0.482	0.778	0.451
	100	0.827	0.521	0.819	0.491	0.781	0.467
	20	0.813	0.510	0.807	0.484	0.758	0.449
5	50	0.819	0.521	0.815	0.491	0.762	0.461
	100	0.832	0.527	0.821	0.502	0.771	0.472
	20	0.824	0.515	0.811	0.499	0.751	0.460
6	50	0.831	0.526	0.819	0.509	0.752	0.471
	100	0.839	0.529	0.822	0.517	0.763	0.483
	20	0.839	0.521	0.814	0.519	0.742	0.475
7	50	0.857	0.529	0.822	0.531	0.744	0.486
	100	0.861	0.531	0.829	0.542	0.751	0.501

m	n	c.p.NS-EE	$\Delta.NS-EE$	c.p.AJ-EE	$\Delta.AJ-EE$	c.p.EL	$\Delta.EL$
	20	0.801	0.511	0.813	0.491	0.829	0.479
0	50	0.809	0.519	0.821	0.501	0.838	0.488

	100	0.821	0.534	0.837	0.536	0.851	0.499
	20	0.805	0.518	0.817	0.497	0.829	0.487
1	50	0.815	0.523	0.828	0.509	0.841	0.493
	100	0.831	0.538	0.841	0.537	0.859	0.521
	20	0.812	0.522	0.821	0.503	0.845	0.491
2	50	0.823	0.529	0.832	0.511	0.851	0.497
	100	0.842	0.544	0.851	0.541	0.861	0.529
	20	0.826	0.529	0.831	0.532	0.851	0.501
3	50	0.857	0.537	0.862	0.539	0.861	0.531
	100	0.871	0.553	0.880	0.562	0.864	0.538
	20	0.839	0.531	0.833	0.532	0.852	0.511
4	50	0.868	0.542	0.864	0.539	0.863	0.532
	100	0.881	0.555	0.881	0.559	0.866	0.543
	20	0.851	0.535	0.848	0.534	0.852	0.521
5	50	0.872	0.548	0.868	0.547	0.866	0.549
	100	0.883	0.559	0.879	0.559	0.871	0.561
	20	0.882	0.541	0.871	0.538	0.841	0.531
6	50	0.895	0.552	0.883	0.549	0.858	0.558
	100	0.903	0.563	0.897	0.561	0.862	0.568
	20	0.897	0.552	0.879	0.545	0.832	0.552
7	50	0.911	0.569	0.891	0.556	0.849	0.571
	100	0.931	0.578	0.902	0.563	0.851	0.587

AJ−EE

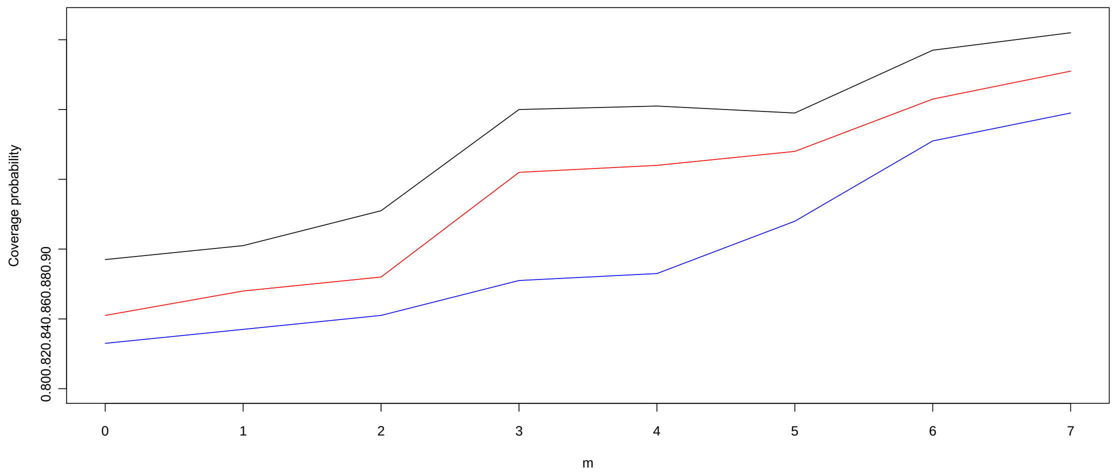

Figure 1. compare of plots for n=20, n=50 and n=100 in AJ-EE method

As n increases, the coverage probability is increased. In Figure 1, the AJ-EE method is shown in the case of n=20, n=50, and n=100, which, as we see, increases the coverage probability as n increases.

5. proof

In what follows we define

$$\psi_{ni} := 1 - \frac{\delta}{G(s)}$$

(5.1)

Lemma 1. *Under the conditions of Theorem (1), as n → ∞,*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_{ni} \xrightarrow{\mathcal{D}} N(0, \sigma_1^2)$$

(5.2) ,

where σ_1^2 *is defined by* (3.13)*.*

Proof. The proof of this Lemma is presented in Lemma (4.1) of Shen and He (2006).

□ **Lemma 2.** *Under the conditions of Theorem (1), we get*

(i) $\frac{1}{n} \sum_{i=1}^{n} \psi_{ni}^2 \xrightarrow{p} \delta_2^2$

(ii) $\widehat{\delta}_2^2 \xrightarrow{p} \delta_2^2,$

(iii) $\widehat{\delta}_1^2 \xrightarrow{p} \delta_1^2.$,

Proof. You can see the proof of this Lemma in Lemma (4.3) of Shen and He (2006). □ **Proof of**

$$0 = \frac{1}{n} \sum_{i=1}^{n} \psi_{ni} \exp(-\widehat{\lambda}_n \psi_{ni})$$

$$= \frac{1}{n} \sum_{i=1}^{n} \psi_{ni} [1 - \widehat{\lambda}_n \psi_{ni}] + \varsigma_n$$

n n

$$\lambda_n + \varsigma_{n}$$

Theorem (1) From (3.3), we get

$$= \frac{1}{n} \sum_{i=1} \psi_{ni} - \frac{1}{n} \sum_{i=1} \psi_{ni}^2 .\widehat{}$$

$|\varsigma_n| \leq \frac{1}{n} \sum_{i=1}^{n} \psi_{ni}^3 .\widehat{\lambda}_n^2.$ Then

where

$$\widehat{\lambda}_n = \sum_{i=1}^{n} \psi_{ni} / \sum_{i=1}^{n} \psi_{ni}^2 + \xi_n$$

(5.3),

where

$$\xi_n = \varsigma_n / \frac{1}{n} \sum_{i=1}^{n} \psi_{ni}^2.$$

By the result of Lemma(2) (2001), we have and $\max_{1 \leq i \leq n} |\psi_{ni}| = o_p(n^{1/2})$ of Wang and Jing

$$|\xi_n| |\psi_{ni}| \max_{1in} |\psi_{ni}| = o_p(n^{-1/2})$$.

By (3.3) and $\max_{1 \leq i \leq n} |\widehat{\lambda}_n \psi_{ni}| - o_p(1)$, we have

$$0 = \lambda_n \sum_{i=1}^{n} \psi_{ni} \exp(-\widehat{\lambda}_n \psi_{ni})$$

$$= \sum_{i=1}^{n} (\widehat{\lambda}_n \psi_{ni}) \exp(-\widehat{\lambda}_n \psi_{ni})$$

$$= \sum_{i=1}^{n} (\widehat{\lambda}_n \psi_{ni}) - \sum_{i=1}^{n} (\widehat{\lambda}_n \psi_{ni})^2 + o_p(1)$$.

That is

$$\sum_{i=1}^{n}(\widehat{\lambda}_n\psi_{ni}) = \sum_{i=1}^{n}(\widehat{\lambda}_n\psi_{ni})^2 + o_p(1)$$

(5. 4).

Hence from (3.6), (3.9), we have

$$\mathcal{Z}_1(\alpha_0) = -2n[\frac{\rho(\widehat{\lambda}_n)}{n} - 1] = 2\sum_{i=1}^{n}[1-\exp(-\widehat{\lambda}_n\psi_{ni})]$$
$$= 2\sum_{i=1}^{n}[\widehat{\lambda}_n\psi_{ni} - \frac{1}{2}(\widehat{\lambda}_n\psi_{ni})^2] + \psi = \widehat{\lambda}_n^2\sum_{i=1}^{n}\psi_{ni}^2 + \psi$$

(5.5)

where

$$|\psi| \le |\widehat{\lambda}_n^3|\sum_{i=1}^{n}\psi_{ni}^3| \le O_p(n^{-3/2})\max_{1\le i\le n}|\psi_{ni}|\sum_{i=1}^{n}\psi_{ni}^2 = o_p(1).$$

Using (5.3), we get

$$\mathcal{Z}_1(\alpha_0) = (\sqrt{n}\bar{\psi}_n)^2/(n^{-1}\sum_{i=1}^{n}\psi_{ni}^2) + o_p(1).$$

(5.6)

□

Proof of (3.9) in Theorem (1). By Taylor expansion, we have

$$\frac{\rho(\lambda_n)}{n} - 1 = n^{-1}\sum_{i=1}^{n}[\exp(-\lambda_n\psi_{ni}) - 1] = n^{-1}\sum_{i=1}^{n}[(-\lambda_n\psi_{ni}) + \frac{1}{2}(\lambda_n\psi_{ni})^2] + ,$$

where

$$n^{-1}\sum_{i=1}^{n}(\lambda_n\psi_{ni})^3 \le \lambda_n^3\max_{1in}|\psi_{ni}|.n^{-1}\sum_{i=1}^{n}\psi_{ni})^2$$

= $Op(n{-}3/2)op(n1/2)Op(n1)$

= $op(n{-}1)$.

From (38), we get

$$-2n\Delta\mathcal{D}(\alpha_0) = -2n\log\{1 + [\frac{\rho(\lambda_n)}{n} - 1]\} = -2n\{[\frac{\rho(\lambda_n)}{n} - 1] + \ell_n\}$$
$$= -2n[\frac{\rho(\lambda_n)}{n} - 1] - 2n\ell_n,$$

$$\ell_n[\frac{\rho(\lambda_n)}{n} - 1]^2.$$

$$\frac{\rho(\lambda_n)}{n} - 1 = -\frac{1}{2}n^{-1}\sum_{i=1}^{n}(\lambda_n\psi_{ni})^2 + o_p(n^{-1}).$$

Since $n^{-1}\sum_{i=1}^{n}(\lambda_n\psi_{ni})^2 = O_p(n^{-1})$, so $\frac{\rho(\lambda_n)}{n} - 1 = o_p(1)$. By Taylor expansion, we have

where

Note that, from (38), we have

$$\ell_n[-\frac{1}{n}n^{-1}\sum_{i=1}^{n}(\lambda_n\psi_{ni})^2 + o_p(n^{-1})]^2.$$

So $\ell_n = O_p(n{-}^2) + o_p(n{-}^2) = o_p(n{-}^1)$ and $n\ell_n = o_p(1)$. Hence

$$\mathcal{Z}_2(\alpha_0) = -2n\Delta\mathcal{D}(\alpha_0) = \lambda_n^2\sum_{i=1}^{n}\psi_{ni}^2 + o_p(1)$$

References:

1. Ahmadou Alioum and Daniel Commenges. A proportional hazards model for arbitrarily censored and truncated data. *Biometrics*, pages 512–524, 1996.
2. Francesco Bravo. Blockwise empirical entropy tests for time series regressions. *Journal of Time Series Analysis*, 26(2):185–210, 2005.
3. Bradley Efron. Nonparametric standard errors and confidence intervals.

canadian Journal of Statistics, 9(2):139–158, 1981.

4. Shuyuan He and Grace L Yang. Estimation of regression parameters with left truncated data. *Journal of statistical planning and inference*, 117(1):99–122, 2003.
5. Edwin T Jaynes. On the rationale of maximum-entropy methods. *Proceedings of the IEEE*, 70(9):939–952, 1982.
6. Gang Li. Nonparametric likelihood ratio estimation of probabilities for truncated data. *Journal of the American Statistical Association*, 90(431):997–1003, 1995.
7. Anna Clara Monti and Elvezio Ronchetti. On the relationship between empirical likelihood and empirical saddlepoint approximation for multivariate m-estimators. *Biometrika*, 80 (2):329–338, 1993.
8. Whitney K Newey and Richard J Smith. Higher order properties of gmm and generalized empirical likelihood estimators. *Econometrica*, 72(1):219–255, 2004.
9. Elias Ould-Sa¨ıd and Mohamed Lemdani. Asymptotic properties of a nonparametric regression function estimator with randomly truncated data. *Annals of the Institute of Statistical Mathematics*, 58(2):357–378, 2006.
10. Art Owen et al. Empirical likelihood ratio confidence regions. *The Annals of Statistics*, 18 (1):90–120, 1990.
11. Junshan Shen and Shuyuan He. Empirical likelihood for truncation parameter in random truncation model. *Journal of statistical planning and inference*, 136(1):90–107, 2006.
12. Winfried Stute and Jane-Ling Wang. The central limit theorem under random truncation. *Bernoulli: official journal of the Bernoulli Society for Mathematical Statistics and Probability*, 14(3):604, 2008.
13. Mei-Cheng Wang. A semiparametric model for randomly truncated data. *Journal of the American Statistical Association*, 84(407):742–748, 1989.
14. Michael Woodroofe et al. Estimating a distribution function with truncated data. *The Annals of Statistics*, 13(1):163–177, 1985.
15. Mei Yao, Jiang-Feng Wang, Lu Lin, and Yu-Xin Wang. Variable selection and weighted composite quantile estimation of regression parameters with left-truncated data. *Communications in Statistics-Theory and Methods*, 47(18):4469–4482, 2018.
16. Guo-qing Zhao, Wei Liang, and Shu-yuan He. Empirical entropy for right censored data. *Acta Mathematicae Applicatae Sinica, English Series*, 31(2):395–404, 2015.