Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT

Statistics & Probability Letters 65 (2003) 29–37

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

Empirical likelihood-based confidence intervals for data with possible zero observations

Song Xi Chen[a,*], Jing Qin[b]

[a] *Department of Statistics and Applied Probability, National University of Singapore, 117543 Singapore*
[b] *Department of Epidemiology and Biostatistics, Memorial Sloan–Kettering Cancer Center, 1275 York Avenue, New York, NY 10021, USA*

Received October 2002; received in revised form April 2003

Abstract

In statistical applications, we often encounter a situation where a substantial number of observations takes zero value and at the same time the non-zero observations are highly skewed. We propose empirical likelihood-based non-parametric confidence intervals for the mean parameter which have two unique features. One is that the information contained in the zero observations is fully utilized. The other is that the proposed confidence intervals are more reflective to the skewness in the non-zero observations than those based on the asymptotic normality.

Keywords: Confidence intervals; Empirical likelihood; Skewed distribution; Zero values

1. Introduction

In statistical applications, we often encounter a situation where a substantial number of observations collected takes zero value. In a recent paper, Zhou and Tu (2000) reported a diagnostic charge study, where the interest was to find the mean charge on patients for certain diagnostic test. As certain proportion of patients had not undertaken the test, the observed values of diagnostic charge for those patients were zero. At the same time, the non-zero observations were highly skewed to the right. The above features are not isolated cases. In survey sampling, a variable may have a lower or upper limit and a substantial number of observed values of the variable takes the limit as its value. For

* Corresponding author. Fax: +65-6874-6624.
E-mail address: stacsx@nus.edu.sg (S.X. Chen).

0167-7152/$ - see front matter
doi:10.1016/S0167-7152(03)00216-5

instance, the expenditure on certain luxury items in a household survey as in Tobin (1958) may take zero values.

The statistical inference for data containing many zero values has been investigated by some authors in the literature. Cox and Snell (1979) studied the inference on the total population error when most observations were zero. A parametric approach was used by assuming errors were positive. Tamura (1988) pointed out that the error distribution was unlikely to be modeled by any standard parametric distribution. By using Ferguson's Dirichlet process, he proposed a non-parametric approach. In the above mentioned paper, Zhou and Tu (2000) studied different confidence intervals for diagnostic test charge data by modeling non-zero value distribution as a lognormal.

Without assuming a parametric model, in this paper we employ Owen's (1988) empirical likelihood to construct confidence intervals for the mean parameter of the population. There are two advantages of this empirical likelihood formulation. One is that the information contained in the zero observations is fully utilized. The other is that the proposed confidence intervals are more reflective to the likely situation that the non-zero value distribution is skewed. To further improve the coverage of the empirical likelihood confidence interval we proposed an empirical Bartlett correction to the empirical likelihood confidence intervals based on the bootstrap. Readers should refer to Owen (2001) for a comprehensive review of the empirical likelihood method. We have noticed a recent University of Waterloo technical report by Chen et al. (2002) on using empirical likelihood method for data with zero values. However, their empirical likelihood approach does not separate zero and no zero observations.

The paper is structured as follows. In Section 2, we construct a joint empirical likelihood for the mean parameter and the probability of taking zero value to fully use all the data information. The empirical likelihood for the mean parameter is derived by profiling out the zero-probability parameter. Confidence intervals are formulated based on a limiting χ_1^2 distribution of the empirical likelihood ratio. We then propose empirical Bartlett corrected confidence intervals. The empirical performance of the empirical likelihood confidence intervals is evaluated in a simulation study in Section 3. It also includes an analysis of the diagnostic charge data given in Zhou and Tu (2000). All the technical details are given in the appendix.

2. Main results

Let $w_1, \ldots, w_n$ be a random sample from a certain population that contains both zero and positive observations. Let $\delta = P(w = 0) > 0$ be the probability of taking zero values and $\mu = EW = (1 - \delta)$ $E(w|w > 0)$ be the population mean. In the literature, the popular assumption for the non-zero observation is lognormal. We are interested in finding a confidence interval for μ. We consider empirical likelihood-based confidence intervals without assuming a parametric distribution for the non-zero observations. Let $\mu^* = E(w|w > 0)$ be the mean of the non-zero observations, then $\mu^* = \mu/(1 - \delta)$. Let n_0 and n_1 be the number of zero and non-zero observations in the sample. For convenience, we denote the non-zero observations as $x_1, \ldots, x_{n_1}$.

A combined binomial likelihood for δ and the empirical likelihood for μ is defined as

$$L(\delta, \mu) = \delta^{n_0}(1 - \delta)^{n_1} \prod_{i=1}^{n_1} p_i$$

subject to the constraints

$$\sum_{i=1}^{n_1} p_i = 1, \quad \sum_{i=1}^{n_1} p_i\{x_i - \mu/(1-\delta)\} = 0, \quad p_i \geqslant 0.$$

It may be easily shown after profiling p_i's by introducing a Lagrange multiplier λ that the log-likelihood

$$\ell(\delta, \mu) = n_0 \log \delta + n_1 \log(1-\delta) - \sum_{i=1}^{n_1} \log\{1 + \lambda[x_i - \mu/(1-\delta)]\},$$

where λ is determined by

$$\sum_{i=1}^{n_1} \frac{x_i - \mu/(1-\delta)}{1 + \lambda[x_i - \mu/(1-\delta)]} = 0. \tag{1}$$

The likelihood ratio statistic is

$$R(\mu) = 2\left\{\max_{\delta,\mu} \ell(\delta, \mu) - \max_{\delta} \ell(\delta, \mu)\right\}.$$

Theorem 1. *Let $w_1, \ldots, w_n$ be a random sample from a population that contains both zero and non-zero positive observations. If the distribution of w_i is non-degenerate and $E(|w_i|^3 | w_i > 0) < \infty$, and $n_0/n \to \delta \in (0,1)$ as $n \to \infty$, then $R(\mu)$ converges to the χ_1^2-distribution when μ is the true mean parameter.*

The limiting χ^2 distribution of $R(\mu)$ implies that an empirical likelihood confidence interval for μ with $1-\alpha$ level of confidence is $I_{1-\alpha} = \{\mu | R(\mu) \geqslant c_\alpha\}$, where c_α is the α-upper quantile of χ_1^2 distribution.

Considering the skewed nature of the underlying distribution, it is expected that there will be some coverage discrepancy in the above confidence interval. Bartlett correction is a novel property of conventional parametric likelihood and has been used for improving the coverage of likelihood ratio based confidence intervals. In the context of empirical likelihood, the validity of Bartlett correction has been formally established in many situations, see for instance DiCiccio et al. (1991) and Chen (1993). The virtue of Bartlett correction is to adjust the mean of the likelihood ratio such that its difference from the mean of the limiting χ^2 becomes a smaller order. This simple mean adjustment then, remarkably, improves the approximation of the likelihood ratio to the χ^2 by one order of magnitude. A formal establishment of Bartlett correction requires quite lengthy derivations of cumulants. To avoid this, we propose in the following empirical mean adjustment to the empirical likelihood ratio via the bootstrap.

Step 1: Generate a bootstrap resample of size n by sampling with replacement from the original sample $\{w_i\}_{i=1}^n$ which includes both zero and non-zero observations and compute the empirical likelihood ratio based on the resample and denote it as $\ell^*(\hat{\mu})$, where $\hat{\mu}$ is the maximum empirical likelihood estimate;

Step 2: For a large integer B, repeat Step 1 B times and obtain $\ell^{*1}(\hat{\mu}), \ldots, \ell^{*B}(\hat{\mu})$.

The empirical Bartlett factor $\eta = B^{-1}\sum_{b=1}^{B} \ell^{b*}(\hat{\mu})$, which is also the bootstrap estimate of $E\{\ell(\mu)\}$ for this case. The Bartlett corrected empirical likelihood confidence interval is $I_{1-\alpha,\mathrm{ebc}} = \{\mu|\ell(\mu) \leqslant c_\alpha * \eta\}$. As the empirical likelihood ratio tends to take values larger than 1, the Bartlett correction shifts the body of the distribution of the likelihood ratio to the left and makes it closer to the χ_1^2 distribution.

3. Simulation results and a real example

In this section we report simulation results and analyse a real dataset from diagnosis test charges. For comparison we also report the confidence intervals based on an asymptotic normal approach which is a non-parametric competitor of the proposed empirical likelihood confidence intervals.

Since the non-parametric maximum likelihood of δ and μ^* are, respectively, $\hat{\delta} = n_0/n$ and $n_1^{-1}\sum_{i=1}^{n_1} w_i$, the non-parametric maximum likelihood estimation of μ is $\hat{\mu} = n^{-1}\sum_{i=1}^{n_1} w_i$. Its variance is $(\mu^*)^2\delta(1-\delta)/n + (1-\delta)\sigma^2/n$. Replacing μ^* by $\bar{w} = n_1^{-1}\sum_{i=1}^{n_1} w_i$ and σ^2 by $\sum_{i=1}^{n_1}(w_i - \bar{w})^2/(n_1-1)$, we have the sample variance $\hat{\sigma}^2$. Therefore, a normal approximation based confidence interval with $1-\alpha$ level of confidence is

$$\left(n^{-1}\sum_{i=1}^{n_1} w_i - z_\alpha\hat{\sigma}, n^{-1}\sum_{i=1}^{n_1} w_i + z_\alpha\hat{\sigma} \right), \tag{2}$$

where z_α is the upper α-quantile of $\mathrm{N}(0,1)$.

We generated two distributions for non-zero values, namely the standard lognormal and the standard exponential distributions, and chose three different proportions of zero values $\delta = 0.2, 0.3$ and 0.5, respectively. Three confidence intervals are considered: the normal approximation intervals, the empirical likelihood intervals $I_{1-\alpha}$ and the empirical Bartlett corrected intervals $I_{1-\alpha,\mathrm{ebc}}$. Table 1 contains the coverage of these confidence intervals with nominal 95% and 99% confidence levels based on 10,000 repetitions for the log-normal case, whereas those for the exponential case are presented in Table 2. The sample sizes considered were $n = 40$ and $n = 100$, and the number of bootstrap resamples was $B = 1000$ in obtaining $I_{1-\alpha,\mathrm{ebc}}$. To assess whether or not coverage errors are symmetric between the two tails, we report also the percentage P_L of intervals in which the lower limit is greater than the true value of μ and the percentage P_R of intervals in which the higher limit is smaller than the true value of μ.

We observe that for both types of non-zero distributions the empirical likelihood confidence intervals $I_{1-\alpha}$ had better coverage than the normal approximation based confidence intervals in all the cases considered. The empirical likelihood confidence intervals were more equal-tailed than the normal approximation as shown by more balanced values of $(P_\mathrm{L}, P_\mathrm{R})$ for all the cases considered. The normal approximation-based confidence intervals produced larger differences between P_L and P_R. However, there were still some under-coverage with $I_{1-\alpha}$. The under-coverage was restored to certain degree by the proposed empirical Bartlett correction. The restoration in coverage was substantial when the sample size was small ($n = 40$) and when the percentage of zero-observation was larger ($\delta = 0.5$). We also observe that the empirical Bartlett correction reduced both P_L and P_R of $I_{1-\alpha}$, but slightly more in P_R. We observe the results in Table 2 were generally better than those in Table 1, which was due to the log-normal distribution is more skewed and has a heavier tail.

Table 1
Empirical coverage in percentage and indicators of symmetry (P_L, P_R) for the normal approximation interval (2) (NORM), the empirical likelihood ratio interval $I_{1-\alpha}$ (EL), the empirical Bartlett corrected interval (EBCEL)

δ	n	Feature	NORM	EL	EBCEL
(a) *Nominal coverage* 95%					
0.2	40	Obs. Cov.	88.50	90.00	92.85
		(P_L, P_R)	(0.20,11.3)	(2.35,7.65)	(0.85,6.3)
0.2	100	Obs. Cov.	91.00	92.00	93.05
		(P_L, P_R)	(0.70,8.3)	(2.10,5.9)	(1.45,5.5)
0.3	40	Obs. Cov.	88.75	89.40	92.15
		(P_L, P_R)	(0.65,10.6)	(3.35,7.25)	(1.70,6.15)
0.3	100	Obs. Cov.	91.40	92.80	93.70
		(P_L, P_R)	(0.4,8.2)	(2.0,5.2)	(1.40,4.9)
0.5	40	Obs. Cov.	88.05	89.50	92.15
		(P_L, P_R)	(0.5,11.45)	(2.65,7.85)	(1.65,6.2)
0.5	100	Obs. Cov.	90.80	91.45	93.30
		(P_L, P_R)	(0.6,8.6)	(2.45,6.1)	(1.30,5.4)
(b) *Nominal coverage* 99%					
0.2	40	Obs. Cov.	94.05	96.75	97.65
		(P_L, P_R)	(0.0,5.95)	(0.4,2.85)	(0.05,2.3)
0.2	100	Obs. Cov.	96.1	97.40	97.70
		(P_L, P_R)	(0.05,3.85)	(0.5,2.1)	(0.4,1.9)
0.3	40	Obs. Cov.	94.35	96.00	97.30
		(P_L, P_R)	(0.0,5.65)	(0.7,3.3)	(0.1,2.6)
0.3	100	Obs. Cov.	96.15	98.40	98.60
		(P_L, P_R)	(0.05,3.8)	(0.25,1.35)	(0.1,1.3)
0.5	40	Obs. Cov.	93.50	96.65	97.65
		(P_L, P_R)	(0.0,6.5)	(0.5,2.85)	(0.2,2.15)
0.5	100	Obs. Cov.	96.00	97.60	98.35
		(P_L, P_R)	(0.0,4.0)	(0.45,1.95)	(0.15,1.5)

Next we consider a real data example given by Zhou and Tu (2000). The data set has 40 patients, but 10 of them have no diagnostic tests during the study period. The 90%, 95% and 99% confidence intervals are $(887.68, 2689.67)$, $(715.07, 2862.27)$ and $(514.38, 3062.97)$ for the normal approximation confidence intervals, and $(1132.26, 2899.15)$, $(1040.78, 3177.22)$ and $(885.65, 3778.85)$ for the empirical likelihood ratio confidence intervals. The empirical Bartlett corrected intervals are $(974.74, 3410.8)$, $(873.11, 3836.27)$ and $(708.54, 4753.64)$, respectively. The two types of confidence intervals are quite different. The empirical likelihood intervals situated to the right of the normal intervals by a large amount and are longer. These are consistent with the just reported simulation results.

Acknowledgements

The authors thank a referee and the Associate Editor for constructive comments and suggestions.

Table 2
Empirical coverage in percentage and indicators of symmetry (P_L, P_R) for the normal approximation confidence interval (2) (NORM), the empirical likelihood ratio confidence interval (EL), the empirical Bartlett empirical likelihood confidence interval (EBCEL) and the parametric likelihood confidence interval (PL)

δ	n	Feature	NORM	EL	EBCEL
(a) *Nominal coverage* 95%					
0.2	40	Obs. Cov.	92.65	93.18	94.32
		(P_L, P_R)	(1.19,6.16)	(2.70,4.12)	(2.19,3.49)
0.2	100	Obs. Cov.	93.95	94.52	95.10
		(P_L, P_R)	(1.11,4.94)	(2.44,3.04)	(2.09,2.81)
0.3	40	Obs. Cov.	92.21	92.99	94.24
		(P_L, P_R)	(0.93,6.86)	(2.48,4.53)	(1.97,3.79)
0.3	100	Obs. Cov.	93.95	94.42	94.81
		(P_L, P_R)	(1.36,4.69)	(2.53,3.05)	(2.28,2.91)
0.5	40	Obs. Cov.	90.67	92.48	94.03
		(P_L, P_R)	(0.68,8.65)	(2.18,5.34)	(1.81,4.16)
0.5	100	Obs. Cov.	93.33	94.17	94.74
		(P_L, P_R)	(0.99,5.68)	(2.48,3.35)	(2.18,3.08)
(b) *Nominal coverage* 99%					
0.2	40	Obs. Cov.	97.13	98.06	98.54
		(P_L, P_R)	(0.1,2.77)	(0.55,1.39)	(0.38,1.08)
0.2	100	Obs. Cov.	98.24	98.74	98.92
		(P_L, P_R)	(0.07,1.69)	(0.33,0.93)	(0.25,0.83)
0.3	40	Obs. Cov.	96.48	97.92	98.38
		(P_L, P_R)	(0.06,3.46)	(0.50,1.58)	(0.30,1.32)
0.3	100	Obs. Cov.	98.20	98.73	98.88
		(P_L, P_R)	(0.10,1.70)	(0.45,0.82)	(0.38,0.74)
0.5	40	Obs. Cov.	95.72	97.83	98.55
		(P_L, P_R)	(0.05,4.23)	(0.36,1.81)	(0.29,1.16)
0.5	100	Obs. Cov.	97.58	98.73	98.92
		(P_L, P_R)	(0.07,2.35)	(0.44,0.84)	(0.34,0.74)

Data are generated from the mixture of zero and standard exponential distribution.

Appendix

Proof of Theorem 1. We first derive $\max_{\delta,\mu} \ell(\delta, \mu)$. Differentiating ℓ with respect to δ and μ,

$$\frac{\partial \ell}{\partial \mu} = -\sum_{i=1}^{n_1} \frac{\partial \ell/\partial \mu\{x_i - \mu/(1-\delta)\} - \lambda/(1-\delta)}{1 + \lambda\{x_i - \mu/(1-\delta)\}} = \lambda n_1/(1-\delta), \tag{A.1}$$

$$\frac{\partial \ell}{\partial \delta} = n_0/\delta - n_1/(1-\delta) - \sum_{i=1}^{n_1} \frac{\partial \ell/\partial \delta\{x_i - \mu/(1-\delta)\} - \lambda\mu/(1-\delta)^2}{1 + \lambda\{x_i - \mu/(1-\delta)\}}$$

$$= \frac{n_0 - \delta n}{\delta(1-\delta)} + \frac{\lambda n_1 \mu}{(1-\delta)^2}. \tag{A.2}$$

Setting (A.1) and (A.2) equal to zero, we have $\lambda=0$ and the maximum likelihood estimators $\hat{\delta}=n_0/n$ and $\hat{\mu}=n^{-1}\sum_{i=1}^{n_1}x_i$. This means that

$$\max_{\delta,\mu} l(\delta,\mu)=n_0\log(n_0/n)+n_1\log(n_1/n). \tag{A.3}$$

We now turn our attention to $\max_\delta \ell(\delta,\mu)$ when μ is the real parameter value. The maximum likelihood estimator of δ given μ, denoted as $\hat{\delta}_\mu$, must satisfy (A.2) which is equivalent to

$$n\hat{\delta}_\mu^2+\{\lambda\mu n_1-(n+n_0)\}\hat{\delta}_\mu+n_0=0. \tag{A.4}$$

Let $\hat{\delta}_\mu=\hat{\delta}_0+\hat{\delta}_1$, where $\delta_j=\mathrm{O_p}(n^{-j/2})$, and $\hat{\lambda}_\mu$ be the solution of (1) at $(\hat{\delta}_\mu,\mu)$. Standard derivation in empirical likelihood for instance that given in Owen (1988) may show that $\hat{\lambda}_\mu=\mathrm{O_p}(n^{-1/2})$. We are going to develop expressions for $\hat{\delta}_j$ and $\hat{\lambda}_\mu$ in the following:

From (A.4) and ignoring terms of $\mathrm{O_p}(n^{-1})$,

$$n(\hat{\delta}_0+\hat{\delta}_1)^2+\{\hat{\lambda}_\mu\mu n_1-(n+n_0)\}(\hat{\delta}_0+\hat{\delta}_1)+n_0=0 \tag{A.5}$$

which means $n\hat{\delta}_0^2-(n+n_0)\hat{\delta}_0+n_0=0$ and hence $\hat{\delta}_0=n_0/n$ coinciding with $\hat{\delta}$, the global maximum likelihood estimator of δ. From (A.4) and (A.5), and ignoring terms of $\mathrm{O_p}(n^{-3/2})$,

$$2n\hat{\delta}_0\hat{\delta}_1+\hat{\lambda}_\mu\mu n_1\hat{\delta}_0-(n+n_0)\hat{\delta}_0=0 \tag{A.6}$$

which means

$$\hat{\delta}_1=\hat{\lambda}_\mu\mu\hat{\delta}_0. \tag{A.7}$$

Apply the Taylor expansion on (1),

$$\sum_{i=1}^{n_1}\{x_i-\mu/(1-\hat{\delta}_\mu)\}[1-\hat{\lambda}_\mu\{x_i-\mu/(1-\hat{\delta}_\mu)\}+\hat{\lambda}_\mu^2\{x_i-\mu/(1-\hat{\delta}_\mu)\}^2+\cdots]=0. \tag{A.8}$$

Since

$$\mu/(1-\hat{\delta}_\mu)=\mu/(1-\hat{\delta}_0)[1+\hat{\delta}_1/(1-\hat{\delta}_0)+\hat{\delta}_1^2/(1-\hat{\delta}_0^2)+\cdots],$$

and after converting (A.8) we have

$$\begin{aligned}\hat{\lambda}_\mu&=\frac{\{\bar{x}-\mu/(1-\hat{\delta}_0)\}-\mu\hat{\delta}_1/(1-\hat{\delta}_0)^2+\mathrm{O_p}(n^{-1})}{\hat{\sigma}_x^2-2\mu\hat{\delta}_1\{\bar{x}-\mu/(1-\hat{\delta}_0)\}/(1-\hat{\delta}_0)^2+\mathrm{O_p}(n^{-1})}+\mathrm{O_p}(n^{-1})\\&=\hat{\sigma}_x^{-2}\{\bar{x}-\mu/(1-\hat{\delta}_0)-\mu\hat{\delta}_1/(1-\hat{\delta}_0)^2\}+\mathrm{O_p}(n^{-1}),\end{aligned} \tag{A.9}$$

where $\bar{x}=n^{-1}\sum_{i=1}^{n_1}x_i$ and $\hat{\sigma}_x^2=n_1^{-1}\sum_{i=1}^{n_1}\{\bar{x}-\mu/(1-\hat{\delta}_0)\}$ are, respectively, the sample mean and variance of the non-zero data.

Solving (A.9) jointly with (A.7) for $\hat{\lambda}_\mu$ and $\hat{\delta}_1$, we have

$$\hat{\lambda}_\mu = \frac{\bar{x} - \mu/(1-\hat{\delta}_0)}{\hat{\sigma}_x^2 + \mu^2\hat{\delta}_0/(1-\hat{\delta}_0^2)} + \mathrm{O_p}(n^{-1}) \tag{A.10}$$

$$\hat{\delta}_1 = \frac{\{(1-\hat{\delta}_0)\bar{x} - \mu\}\mu\hat{\delta}_0}{(1-\hat{\delta}_0)\hat{\sigma}_x^2 + \mu^2\hat{\delta}_0/(1-\hat{\delta}_0)} + \mathrm{O_p}(n^{-1}). \tag{A.11}$$

Now we are ready to give an expansion for the log-likelihood ratio R. Note that

$$\begin{aligned}
\tfrac{1}{2}R(\mu) &= \max_{\delta,\mu} \ell(\delta,\mu) - \max_{\delta} \ell(\delta,\mu) \\
&= n_0 \log\left(\frac{n_0/n}{n_0/n + \hat{\delta}_1}\right) + n_1 \log\left(\frac{n_1/n}{n_1/n - \hat{\delta}_1}\right) \\
&\quad + \hat{\lambda}_\mu n_1\{\bar{x} - \mu/(1-\hat{\delta}_\mu)\} - \tfrac{1}{2}\hat{\lambda}_\mu^2 n_1 \hat{\sigma}_x^2 + \mathrm{o_p}(1) \\
&= -n_0 \log(1 + n\hat{\delta}_1/n_0) - n_1 \log(1 - n\hat{\delta}_1/n_1) \\
&\quad + \tfrac{1}{2}\hat{\lambda}_\mu^2\hat{\sigma}_x^2 n_1 + \mathrm{o_p}(1) \\
&= -n_0\left(n\hat{\delta}_1/n_0 - \tfrac{1}{2}n^2\hat{\delta}_1^2/n_0^2\right) - n_1\left(-n\hat{\delta}_1/n_0 - \tfrac{1}{2}n^2\hat{\delta}_1^2/n_0^2\right) + \tfrac{1}{2}\hat{\lambda}_\mu^2\hat{\sigma}_x^2 n_1 + \mathrm{o_p}(1).
\end{aligned}$$

Substituting (A.10) and (A.11)

$$R(\mu) = \hat{\lambda}_\mu^2(\hat{\sigma}_x^2 n_1 + n_0\mu^2/n_1) + \mathrm{o_p}(1) = \frac{(n_1\bar{x}/n - \mu)^2}{\{n_1\hat{\sigma}_x^2/n^2 + n_0\mu^2/(nn_1)\}} + \mathrm{o_p}(1). \tag{A.12}$$

It may be shown that

$$\begin{aligned}
\mathrm{Var}\left(\frac{n_1}{n}\bar{x}\right) &= E\left\{\frac{n_1^2}{n^2}\mathrm{Var}(\bar{x}|n_1)\right\} + \mathrm{Var}\left\{\frac{n_1}{n}E(\bar{x}|n_1)\right\} \\
&= E\left(\frac{n_1}{n^2}\sigma_x^2\right) + \frac{\mu^2}{n^2(1-\delta)^2}\mathrm{Var}(n_1) = \frac{(1-\delta)\sigma_x^2}{n} + \frac{\delta\mu^2}{n(1-\delta)}.
\end{aligned}$$

This means that as $n \to \infty$

$$n_1\hat{\sigma}_x^2/n^2 + n_0\mu^2/(nn_1) \xrightarrow{\mathrm{p}} \mathrm{Var}\left(\frac{n_1}{n}\bar{x}\right). \tag{A.13}$$

Also it may be shown via the central limit theorem that $n_1\bar{x}/n - \mu \xrightarrow{\mathrm{d}} \mathrm{N}\{0, \mathrm{Var}((n_1/n)\bar{x})\}$. This together with (A.13) and (A.12) means that $R(\mu) \xrightarrow{\mathrm{d}} \chi_1^2$ as $n \to \infty$. This completes the proof of the theorem. □

References

Chen, J., Chen, S.-Y., Rao, J.N.K., 2002. Empirical likelihood confidence intervals for a population containing many zero values. Technical report 2002–04, Department of Statistics and Actuarial Science, University of Waterloo, Canada.

Chen, S.X., 1993. On the coverage accuracy of empirical likelihood confidence regions for linear regression model. Ann. Inst. Statist. Math. 45, 621–637.

Cox, D.R., Snell, E.J., 1979. On sampling and the estimation of rare errors. Biometrika 66, 125–132.

DiCiccio, T.J., Hall, P., Romano, J.P., 1991. Empirical likelihood is Bartlett-correctable. Ann. Statist. 19, 1053–1061.

Owen, A., 1988. Empirical likelihood ratio confidence intervals for a single functional. Biometrika 75, 237–249.

Owen, A., 2001. Empirical Likelihood. Chapman & Hall, London.

Tamura, H., 1988. Estimation of rare errors using expert judgment. Biometrika 75, 1–9.

Tobin, J., 1958. Estimation of relationships for limited dependent variables. Econometrica 26, 24–36.

Zhou, X.H., Tu, W., 2000. Confidence intervals for the mean of diagnostic test charge data containing zeros. Biometrics 56, 1118–1125.