Acta Mathematicae Applicatae Sinica, English Series
Vol. 21, No. 2 (2005) 295–302

Acta Mathematicae Applicatae Sinica, English Series

L_1-Norm Estimation and Random Weighting Method in a Semiparametric Model

Liu-gen Xue[1], **Li-xing Zhu**[2]

[1]College of Applied Sciences, Beijing University of Technology, Beijing 100022, China (E-mail: lgxue@bjpu,edu.cn)
[2]Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences, Beijing 100080, China & University of Hong Kong, Hong Kong, China (E-mail: lzhu@hku.hk)

Abstract In this paper, the L_1-norm estimators and the random weighted statistic for a semiparametric regression model are constructed, the strong convergence rates of estimators are obtain under certain conditions, the strong efficiency of the random weighting method is shown. A simulation study is conducted to compare the L_1-norm estimator with the least square estimator in term of approximate accuracy, and simulation results are given for comparison between the random weighting method and normal approximation method.

Keywords L_1-norm estimation, random weighting method, semiparametric regression model
2000 MR Subject Classification 62G05, 62F12

1 Introduction

In a semiparametric regression model, one observes $\{(T_i, X_i, Y_i),\ 1 \le i \le n\}$ of which the Y_i's are response variable depending on covariates (T_i, X_i) through the relationship

$$Y_i = X_i^\tau \beta + g(T_i) + e_i, \qquad i =, 1, \cdots, n, \tag{1.1}$$

where $\{(T_i, X_i, Y_i), 1 \le i \le n\}$ are independent and identically distributed (i.i.d.) as (T, X, Y), the covariate (T, X) is $R^d \times [0, 1]$ valued, β is a d-vector of unknown parametric, and g is an unknown smooth function on [0,1], $\{e_i, 1 \le i \le n\}$ are i.i.d. random error which are independent of $\{(X_i, T_i), 1 \le i \le n\}$.

Model (1.1) was proposed and studied by Engle, Granger and Rice[6] and has been extensively investigated in recent years. For example, see [2,8–10]. By use of piecewise polynomial to approximate g, Chen[2] acquired the estimators of β and g based on least squares (LS) principle. Under some mild conditions, he concluded that the underlying estimators of the parametric components achieved the convergence rate $n^{-1/2}$. Since the LS-based estimators are not robust, Shi and Li[9] gave the more robust estimators, the L_1-norm estimators of β and g. Specifically, they used a piecewise polynomial to approximate g and obtained the estimators of β and g, by way of solving absolution value equations. However, a key condition in Shi and Li[9] was that T and X were independent. Later, Shi and Li[10] improved their results to a wider class

Manuscript received July 18, 2004.
Supported by the Natural Science Foundation of Beijing City of China (1042002), the Science and Technology Development Foundation of Education Committee of Beijing City, the Special Expenditure of Excellent Person Education of Beijing (20041D0501515) and the grants (HKU7181/02H, HKU7060/04P) from Research Grants Council of Hong Kong, Hong Kong.

of ρ-function, and obtained the M-estimators of β and g, say respectively $\widehat{\beta}$ and $\widehat{g}$. They gave the convergence rates in probability for $\widehat{\beta}$ and $\widehat{g}$, and proved the asymptotic normality.

The confidence interval of β naturally lay in the limiting distribution and accuracy depends on how fast the theoretical distribution converges to its limiting distribution. However, we can not directly use the asymptotic distribution of $\widehat{\beta}$ because there are unknown parameters in the asymptotic distribution. Hence, the confidence interval of β is not obtained by the asymptotic normal distribution. The asymptotic variance of $\widehat{\beta}$ need to be estimated.

To simulate the distribution of the estimator $\widehat{\beta}$, a bootstrap method can be used (see [4]). The method need to resample the data to obtain a Monte Carlo distribution for the distribution of estimator. The bootstrap is a very nice method, and had be used comprehensively. The purpose of this paper is to use an alternative method, the random weighting method, which was originally suggested by Rubin[7]. It is a effective method to approximate the unknown distribution of pivotal quantities, which is often treated by the normal approximation or bootstrap method. This method is motivated by the bootstrap method and can be seen as a kind of smoothing of bootstrap. The main idea of this approach is: Instead of resampling from the original data set, we generate a group of random weights directly from the computer and use them to weight the original samples. The conditional distributions of suitable sums of the weighted samples are employed to approximate the distribution of original pivotal statistics. The weights can be chosen from a sequence of suitable i.i.d. random variables.

In this paper, the random weighted statistic of $\widehat{\beta}$ is constructed. The distribution of $\widehat{\beta}$ is simulated by the distribution of random weighted statistic. Therefore, the confidence interval of β is obtained by using the quantile of the distribution of random weighted statistic. Our paper shows that the random weighting method produces the confidence intervals with greater coverage accuracy than those obtained by normal approximation method. This paper also gives the strong convergence rates of $\widehat{\beta}$ and $\widehat{g}$, and the convergence rate of $\widehat{g}$ achieves the optimal global rate of convergence. Hence, our results improve the results of weak convergence rates in Shi and Li[10]. A simulation study shows that the L_1-norm estimators have many advantage, and the simulation results are given for comparison between the random weighting method and normal approximation.

The organization of this paper is as follows. In Section 2, formal definitions of the L_1-norm estimator for β and the piecewise polynomial L_1-norm estimator for $g(t)$ are given, and the main results are presented. In section 3, we conduct a simulation study to compare the advantage of the L_1-norm estimator with the least square estimator, and we also compare the random weighting method and the normal approximation method.

2 Main Results

This section has two subsections. Section 2.1 gives the asymptotic behaviors of the estimator $\widehat{\beta}$ and its random weighted statistic. In Section 2.2, the confidence intervals of β are constructed by using our results.

2.1 Asymptotic Behaviors of the Random Weighted Statistic

We first give some notation. Let m be a nonnegative integer. For an integer M_n, denoting $\delta_n = 1/(2M_n)$, $p = (m+1)M_n$, $p_1 = p + d$ and $t_v = (2v-1)\delta_n$, $1 \le v \le M_n$. We define

$$I_{mv} = \begin{cases} [2(v-1)\delta_n, 2v\delta_n), & 1 \le v < M_n, \\ [1-2\delta_n, 1], & v = M_n, \end{cases}$$
$$\varphi_v(t) = I_{nv}(t)\big(1, t-t_v, \cdots, (t-t_v)^m\big)^\tau,$$

$$\varphi(t) = \big(\varphi_1(t)^\tau, \cdots, \varphi_{M_n}(t)^\tau\big)^\tau,$$

where $I_{nv}(\cdot)$ stands for the indicator function of I_{nv}.

For $r' \in (0,1]$, and a nonnegative integer m', let

$$\mathcal{G}_{m',r'} = \big\{g: \ g \ \text{ has } m' \text{ derivative, and } |\ g^{(m')}(t) - g^{(m')}(t') \le C_0|t-t'|^{r'}, \ \forall\, t,t' \in [0,1]\big\},$$

where C_0 is a positive constant. Let $|\cdot|$ denote either the Euclidean norm of a vector or the absolute value of a real number according to the context. Let $\|g\|^2_{L^2}$ denote the L^2 norm, defined by $\|g\|^2_{L^2} = \int_0^1 g^2(t)f(t)dt$.

Now, we define the estimators of β and g. Note that $\varphi(t)$ is a $p\times 1$ vector with components $(t-t_v)^j$, $v = 1,\cdots,M_n$, $j = 0,1,\cdots,m$, then $\varphi(t)^\tau\alpha(\alpha \in R^p)$ is a piecewise polynomial of degree m on [0,1]. Let $\widehat{\beta}$ and $\widehat{g}(t) = \varphi(t)^\tau\widehat{\alpha}$ be the L_1-norm estimators for β and $g(t)$ respectively, where $\widehat{\beta}$ and $\widehat{\alpha}$ are chosen to satisfy

$$\sum_{i=1}^n \big|Y_i - X_i^\tau\widehat{\beta} - \varphi(T_i)^\tau\widehat{\alpha}\big| = \min_{\alpha,\beta}\sum_{i=1}^n \big|Y_i - X_i^\tau\beta - \varphi(T_i)^\tau\alpha\big|.$$

To simulate the distribution of $\widehat{\beta}$, the random weighting L_1-norm estimator $\widehat{\beta}^*$ of $\widehat{\beta}$ is introduced, in which $\widehat{\beta}^*$ satisfies

$$\sum_{i=1}^n W_i\big|Y_i - X_i^\tau\widehat{\beta}^* - \varphi(T_i)^\tau\widehat{\alpha}^*\big| = \min_{\alpha,\beta}\sum_{i=1}^n W_i\big|Y_i - X_i^\tau\beta - \varphi(T_i)^\tau\alpha\big|,$$

where $\{W_i, 1\le i\le n\}$ are i.i.d. random variable, satisfying

W. $P\{W_1>0\} = 1$, $E(W_1) = 1$, $E(W_1^2) = \gamma \ge 1$, $E(|W_1|^{2(m+r)+1}) < \infty$, $r \in (0,1]$, and $\{W_i, 1\le i\le n\}$ which are independent of $\{(T_i, X_i, Y_i), 1\le i\le n\}$.

There exist the weights satisfying condition W, for example, the exponential mean 1 weights, that is that the distribution of W_i is the exponential distribution with mean 1 and variance 1.

In order to study the asymptotic behaviors of the estimators, we first give a group of conditions.

C1. The distribution of T is absolutely continuous having a density function $p(t)$, and there exist two constants c_1 and c_2 such that $0 < c_1 \le p(t) \le c_2 < \infty$, for all $t \in [0,1]$.

C2. $g \in \mathcal{G}_{m,r}$ with $r \in (0,1]$ and $m + r > 1/2$.

C3. There exists a positive constant $\delta_0 \in (0, m+r)$ such that

$$\max_{1\le i\le n}|X_i| = O\big(n^{\tilde{\delta}}\big), \quad \text{a.s.}, \quad \tilde{\delta} = (m+r-\delta_0)/[2(m+r)+1].$$

C4. $P\{e_1 \le 0\} = 0.5$, the distribution of e_1 has a density f in a neighborhood of zero with $f(0) > 0$, and there exist two constants c_3 and c_4 such that

$$|f(t) - f(0)| \le c_4|t|, \quad \text{for all } |t| \le c_3.$$

C5 $E(X) = 0$; for $\xi(t) := \big(\xi_1(t),\cdots,\xi_d(t)\big)^\tau = E(X|T=t)$, $\xi_i(t) \in \mathcal{G}_{m_1,r_1}$, $\Sigma = \text{cov}(X - \xi(T))$ is positive definite and there exists a positive definite matrix Σ_0 such that $\Sigma_0 - \text{cov}(X\xi(t)) > 0$ for all $t \in (0,1]$, where m_1 is a nonnegative integer and $r_1 \in (0,1]$, satisfying $m_1 + r_1 > 1/2$.

Remark 2.1. Condition C1 ensures that $\widehat{g}$ has high rate of convergence. Condition C2 demands that g has m derivative $g^{(m)}$, and $g^{(m)}$ satisfies Lipschitz condition of order r. This ensures the order of the convergence rates of $\widehat{g}$ and $\widehat{\beta}$. Condition C3 is a constraint for samples X_i. If X_i has $2\widetilde{\delta}$th moment, then, condition C3 is satisfied. Condition C4 demands the distribution of ε_i is symmetric and positive in a neighborhood of zero, and has a bounded first derivative. Condition C5 ensures that there exits the limiting variance for the estimator $\widehat{\beta}$. Thus, C1–C5 are some common condition.

We now introduce some notations. Let $A_1, \cdots, A_k$ respectively be $l_1 \times l_1, \cdots, l_k \times l_k$ matrices and let $\mathrm{DIAG}(A_1, \cdots, A_k)$ denote a $k \times k$ block diagonal matrix with A_i as the (i,i) block. Denoting

$$
\begin{aligned}
&D_0 = \mathrm{diag}(1, \delta_n, \cdots, \delta_n^m)_{(m+1)\times(m+1)},\\
&D = \mathrm{DIAG}(D_0, \cdots, D_0)_{p\times p},\\
&\pi_v(t) = D_0^{-1}\varphi_v(t) = I_{nv}(t)\big(1, (t-t_v/\delta_n), \cdots, (t-t_v/\delta_n)^m\big)^\tau,\\
&\pi(t) = D^{-1}\varphi(t) = \big(\pi_1^\tau(t), \cdots, \pi_{M_n}^\tau(t)\big)^\tau,\\
&\Pi_n^\tau = \big(\pi(T_1), \cdots, \pi(T_n)\big)_{p\times n},\\
&G_n = \Pi_n(\Pi_n^\tau\Pi_n)^+\Pi_n^\tau,\\
&\widehat{\Sigma}_n = (X_1, \cdots, X_n)(I - G_n)(X_1, \cdots, X_n)^\tau,
\end{aligned}
$$

where A^+ is the Moore inverse of a matrix A. Let $H_{1n}^2 = \widehat{\Sigma}_n$, $H_{2n}^2 = 2M_n\Pi_n^\tau\Pi_n$, $H_n = \mathrm{DIAG}(H_{1n}, H_{2n})$. Then, for $i = 1, \cdots, n$, we define

$$
Z_i = (Z_{1i}^\tau, Z_{2i}^\tau)^\tau, \qquad Z_{2i} = H_{2n}^+\pi(T_i)\delta_n^{-1/2}, \tag{2.1}
$$

$$
Z_{1i} = H_{1n}^+\Big(X_i - \sum_{k=1}^n \pi(T_k)^\tau(\Pi_n^\tau\Pi_n)^+\Pi(T_i)X_k\Big). \tag{2.2}
$$

Given positive numbers $\{a_n\}$ and $\{b_n\}$, let $a_n \sim b_n$ denote $0 < \inf_n(a_n/b_n) \le \sup_n(a_n/b_n) < \infty$. $\mathrm{sgn}(x)$ stands for the sign function of x, that is $\mathrm{sgn}(x)$=1 if $x > 0$, 0 if $x = 0$, -1 if $x < 0$.

Theorem 2.1. *Let $M_n \sim n^{1/[2(m+r)+1]}$ with $m + r > 1/2$. If conditions* C1–C5 *hold, then, we have*

$$
\begin{aligned}
&|\widehat{\beta} - \beta| = O\big(n^{-(m+r)/[2(m+r)+1]}\big), \qquad \text{a.s.},\\
&\|\widehat{g}(t) - g(t)\|_{L^2} = O\big(n^{-(m+r)/[2(m+r)+1]}\big), \qquad \text{a.s.}
\end{aligned}
$$

Theorem 2.2. *Addition to the conditions of Theorem 2.1, further, if $m+r > 1$ and condition* W *hold, then, for almost all sample sequences, we have*

$$
\widehat{\Sigma}_n^{1/2}(\widehat{\beta} - \beta) = \big(2f(0)\big)^{-1}\sum_{i=1}^n Z_{1i}\, sgn(e_i) + o_P(1), \tag{2.3}
$$

and when $(T_1, X_1, Y_1), \cdots, (T_n, X_n, Y_n)$ are given, for almost all sample sequences,

$$
\widehat{\Sigma}_n^{1/2}(\widehat{\beta}^* - \beta) = \big(2f(0)\big)^{-1}\sum_{i=1}^n W_i Z_{1i}\, sgn(e_i) + o_P(1), \qquad \text{a.s.} \tag{2.4}
$$

Theorem 2.3. *Under the conditions of Theorem 2.2 and $\gamma = 2$ in condition W, then, along almost sample sequences,*

$$\sqrt{n}(\widehat{\beta}^* - \widehat{\beta}) \xrightarrow{\mathcal{L}^*} N(0, (2f(0))^{-2}\Sigma^{-1}), \qquad \text{a.s.}, \tag{2.5}$$

$$\sup_{u \in R^d} \left|P^*\{\sqrt{n}(\widehat{\beta}^* - \widehat{\beta}) \le u\} - P\{\sqrt{n}(\widehat{\beta} - \beta) \le u\}\right| \longrightarrow 0, \qquad \text{a.s.} \tag{2.6}$$

where Σ is defined in condition C5, $\mathcal{L}^*$ and P^* represent respectively the convergence in distribution and the conditional probability operation when $(T_1, X_1, Y_1), \cdots, (T_n, X_n, Y_n)$ are given.

By (2.3) and Lindeberg Central Limit Theorem, for almost all sample sequences, we obtain

$$\sqrt{n}(\widehat{\beta} - \beta) \xrightarrow{\mathcal{L}} N\big(0, (2f(0))^{-2}\Sigma^{-1}\big).$$

Remark 2.2. The proofs of Theorem 2.1–2.3 can be found in another technical report. (See [11]).

Remark 2.3. Shi and Li[10] prove that the weak convergence rates of $\widehat{g}$ and $\widehat{\beta}$ can attain $n^{-(m+r)/[2(m+r)+1]}$. Our Theorem 2.1 shows that the strong convergence rates of $\widehat{g}$ and $\widehat{\beta}$ also can attain $n^{-(m+r)/[2(m+r)+1]}$. Hence, the result of Theorem 2.1 is stronger than that in [10]. Specially, when $m = 1$ and $r = 1$, the convergence rate of $\widehat{\beta}$ can achieves the optimal global rate $n^{-2/5}$. Theorem 2.2 are the asymptotic expression for the estimators $\widehat{\beta}$ and $\widehat{\beta}^*$, respectively, which is the sum of i.i.d. random variables. Using (2.3) and (2.4), we can prove the asymptotic normality of the estimators $\widehat{\beta}$ and $\widehat{\beta}^*$ respectively. (2.6) of Theorem 2.3 shows that the distribution of error of estimator $\widehat{\beta}$ can be approximated by a distribution of random weighted statistics. (2.6) can be used to construct the confidence intervals of β.

2.2 The Confidence Intervals of β

To construct a confidence interval of β, we would ideally to know the exact distribution of Q_n, where $Q_n = \sqrt{n}(\widehat{\beta} - \beta)$, from which we would compute the quantile u_α defined by $P\{Q_n \le u_\alpha\} = 1 - \alpha$. The ideal $(1 - \alpha)\%$ one-sided confidence interval for β would be $I_1 = [\widehat{\beta} - n^{-1/2}u_\alpha, \infty)$. Its converge probability is precisely $1 - \alpha$. However, the distribution of Q_n is unknown, and the asymptotic normal distribution of Q_n has the unknown parameters $f(0)$ and Σ. Hence, it is not used to make statistical inference. In such cases, we might replace u_α by its random weighted estimate $\widehat{u}_\alpha$, defined by

$$P^*\{\sqrt{n}(\widehat{\beta}^* - \widehat{\beta}) \le \widehat{u}_\alpha\} = 1 - \alpha.$$

Hence, the corresponding $(1 - \alpha)\%$ one-sided confidence interval for β is

$$\widehat{I}_1 = [\widehat{\beta} - n^{-1/2}\widehat{u}_\alpha,\ \infty).$$

Its two-sided counterpart is

$$\widehat{I}_2 = [\widehat{\beta} - n^{-1/2}\widehat{u}_{\alpha/2},\ \widehat{\beta} + n^{-1/2}\widehat{u}_{\alpha/2}]. \tag{2.7}$$

The interval $\widehat{I}_2$ has the nominal coverage $1 - \alpha$.

3 Simulation Study

In this section, we would do some simulation works to show the practicability of the Theorem 2.1 and Theorem 2.3. We consider three main points. The first, we compare the range of absolute errors and root mean squared error for the L_1-norm estimators and the least square estimators; The second is comparison of the random weighted distribution with the true and asymptotic ones. The third is investigation of how much practical difference there it is between random weighting and asymptotic normally confidence intervals, and compute the coverage probabilities of the random weighted confidence intervals and the asymptotic normal confidence intervals.

3.1 Range of the Errors for Estimators

Consider a semiparametric regression model,

$$Y_i = X_i\beta + 6T_i^2 + e_i, \quad i = 1, \cdots, n, \tag{3.1}$$

where $\beta = 0.75$, $X_i \sim N(0,1)$, $T_i \sim U(0,1)$ and $e_i \sim N(0, 0.3^2)$, $i = 1, \cdots, n$. We take $M_n = [0.8n^{1/5}]$, where $[x]$ represents the integer part of x. It is the optimum chose of the estimators based on the 500 simulations(Mean square error is the best small). The estimators $\{\widehat{g}(\cdot)\}$ are assessed via the root mean squared error(RMSE),

$$RMSE = \left\{ n_{grid}^{-1} \sum_{k=1}^{n_{grid}} \left[\widehat{g}(t_k) - g(t_k)\right]^2 \right\}^{1/2},$$

where $\{t_k, 1 \le k \le n_{grid}\}$ are regular grid points. The run results are drawn in Figure 1 and Figure 2. In the two figures, we use L1NE and LSE to represent the L_1-norm estimate and the least squares estimate respectively.

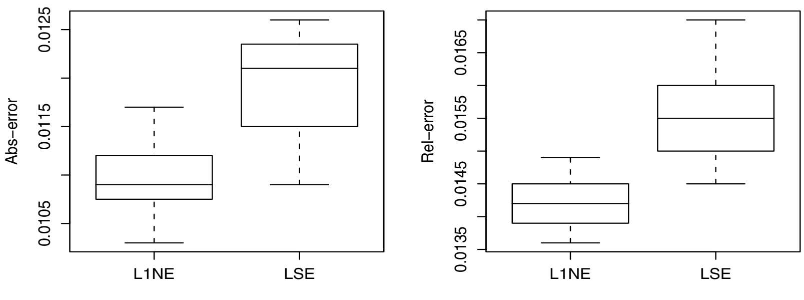

Figure 1. Range of the absolute errors and the relative errors for the estimators of β.

Figure 1 compares the range of absolute errors and relative errors for 500 error values in estimation $\widehat{\beta}$. The LSE range is seen to be much broader than L_1-norm. This illustrates the fact that the L_1-norm estimate is more robust than the least squares estimate.

Figure 2 shows the range of absolute errors and root mean squared errors for 500 error values in estimation $\widehat{g}$. From figure 2, it is clear to see that the L_1-norm estimate is more robust than the least squares estimate.

3.2 Comparison of the Error Distribution of Estimators

We consider the following three distribution functions:

$$G(u) = P\{\sqrt{n}(\widehat{\beta} - \beta) \le u\},$$
$$G^*(u) = P^*\{\sqrt{n}(\widehat{\beta}^* - \widehat{\beta}) \le u\},$$
$$B(u) = P\{\xi \le u\}, \quad \xi \sim N(0, (2f(0))^{-2}\Sigma^{-1}).$$

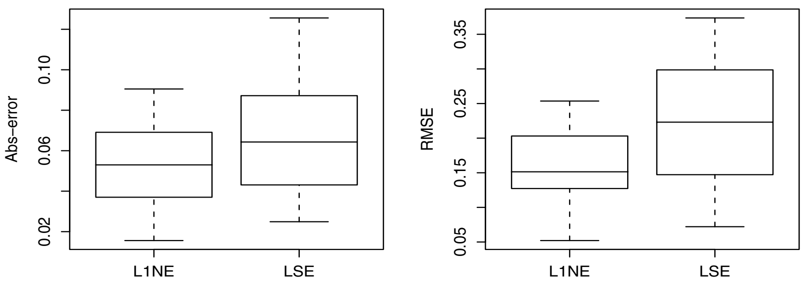

Figure 2. Range of the absolute errors and the root mean squared errors (RMSE) for the estimators of $g(t)$.

We study the questions that $G^*(u)$ and $B(u)$ approximate to $G(u)$. To achieve this purpose, we run a small simulation study, and data generated from the model (3.1) with $\beta = 0.75$. Simulations were run with sample sizes $n = 90$, $M_n = [0.8n^{1/5}]$. we choose two weights, one is the exponential mean 1 weights; another is that W_i has the following density function,

$$p(w) = 2^{-\frac{1}{4}}\left(w - 1 + \frac{\sqrt{2}}{2}\right)^{-\frac{1}{2}} e^{-(w-1+\frac{\sqrt{2}}{2})} I_{(1-\frac{\sqrt{2}}{2},\,\infty)}(w),$$

where $I_{(1-\frac{\sqrt{2}}{2},\,\infty)}(w)$ is the indicator function of $(1 - \frac{\sqrt{2}}{2}, \infty)$. We call the chi-squared weights, because its density function is obtained by transforming the chi-square distribution with 1 degree of freedom. The run results are drawn in Figure 3.

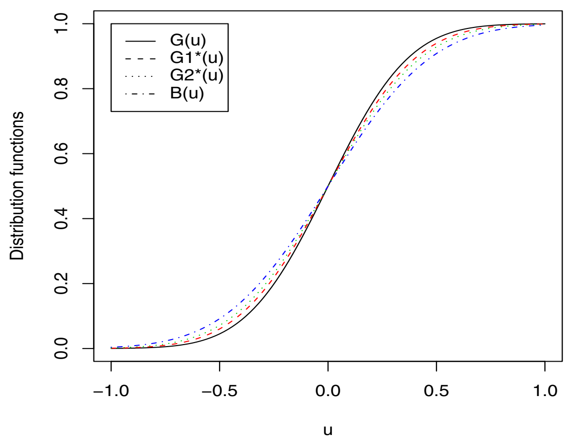

Figure 3. Comparison of the random weighted distributions with the true and the asymptotic normal distributions

In Figure 3, the $G1^*(u)$ and $G2^*(u)$ are the random weighted distributions when choose exponential weights and chi-squared weights, respectively. Figure 3 compares the different between the random weighted distributions and the asymptotic normal distributions. It is clear in this example that the random weighted distribution gives a more accurate approximation to the true distribution than the asymptotic normal distribution do. But, the exponential weights are best than both the chi-squared weights and the asymptotic normal distribution. Hence, we can use $G1^*(u)$ to obtain the quantile of order p.

3.3 Accuracy of the Confidence Intervals

To study the accuracy of estimators, we consider the lengths and the coverage probabilities for $\widehat{I}_2$ in (2.7). We also use the model (3.1) with $\beta = 0.75$ and $\sigma = 0.3$, 0.6, where $\sigma^2 = \text{var}(e_1)$. Simulations were run with sample sizes $n = 30, 90, 180$. We also take $M_n = [0.8n^{1/5}]$, and W_i's are the exponential mean 1 weights and the chi-squared weights. The average lengths and the coverage probabilities of the confidence intervals with nominal level 0.95, were computed by using 2000 simulation runs. The simulation results are presented in Tables 1.

Table 1. Lengths and coverage probabilities of the confidence intervals of β when α=0.05 and use the exponential mean 1 weights and the chi-squared weights.

		Average lengths			Coverage probabilities		
σ	n	EW	CSW	NA	EW	CSW	NA
0.3	30	0.120	0.125	0.148	0.905	0.896	0.884
	90	0.116	0.119	0.140	0.927	0.921	0.900
	180	0.103	0.107	0.135	0.941	0.935	0.916
0.6	30	0.134	0.138	0.158	0.892	0.887	0.890
	90	0.125	0.127	0.152	0.916	0.910	0.902
	180	0.119	0.124	0.148	0.935	0.928	0.908

Note. EW: the exponential mean 1 weights, CSW: the chi-squared weights, NA: the normal approximation.

It is clear from Tables 1 that the random weighting method consistency gives shorter intervals and higher levels than the normal approximation. We also see from Tables 1 that the lengths and the coverage probabilities of confidence intervals have relations with error variance σ^2. The larger the error variance is, the larger the length of confidence intervals.

References

[1] Bennett, G. Probability inequalities for sums of independent random variables. *J. Amer. Statist. Assoc.*, 57: 33–45 (1962)
[2] Chen, H. Convergence rates for parametric components in a partly linear model. *Ann. Statist.*, 16: 136–146 (1988)
[3] Dvoretzky, A. Central limit theorems for dependent random variables. In *Proceedings Sixth Berkley Symp. Math. Statist. Prob.*, Univ of California Press, 513–555 (1972)
[4] Efron, B. Bootstrap methods: another look at the jackknife. *Ann. Statist.*, 7: 1–26 (1979)
[5] Eggleston, H.G. Convexity cambridge tracts in mathematics and mathematical physics. Cambridge University Press, New York, 1958
[6] Engel, R., Granger, C. Rice, J. et al. Semiparametric estimation of the relation between weather and electricity sales. *J. Amer. Statist. Assoc.*, 81: 310–320 (1986)
[7] Rubin, Donald B. The Bayesian bootstrap. *Ann. Statist.* 9: 130–134 (1981)
[8] Speckman, P. Kernel Smoothing in partial linear models. *J. Roy. Statist. Soc.* (Series B), 50: 413–436 (1988)
[9] Shi, P.D., Li, G.Y. Asymptotic normality of L_1-norm estimates for parametric components in a party linear model. *China Ann. Math.* (Series A), 15: 478–484 (1994)
[10] Shi, P.D., Li, G.Y. A note on the convergence rates of M-Estimates for partly linear model. *Statistics*, 26: 27–47 (1995)
[11] Xue, L.G., Zhu, L.X. L_1-norm estimation and random weighting method in a semiparametric model. *Technical report*, College of Applied Sciences, Beijing University of Technology, Beijing 2004