Huang et al. *Journal of Inequalities and Applications* (2016) 2016:125
DOI 10.1186/s13660-016-1064-6

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

The large deviation for the least squares estimator of nonlinear regression model based on WOD errors

Xufeng Huang[1], Xufei Tang[2], Xin Deng[2] and Xuejun Wang[2*]

*Correspondence: wxjahdx2000@126.com
[2]School of Mathematical Sciences, Anhui University, Hefei, 230601, P.R. China
Full list of author information is available at the end of the article

Abstract

As a kind of dependent random variables, the widely orthant dependent random variables, or WOD for short, have a very important place in dependence structures for the intricate properties. And so its behavior and properties in different statistical models will be a major part in our research interest. Based on WOD errors, the large deviation results of the least squares estimator in the nonlinear regression model are established, which extend the corresponding ones for independent errors and some dependent errors.

MSC: Primary 62J02; secondary 62F12

Keywords: widely orthant dependent random variables; large deviation; nonparametric regression model; least squares estimator

1 Introduction

Many researchers have paid attention to the study of the probability limit theorem and its applications for the independent random variables, while the fact is that most of the random variables found in real practice are dependent, which just motivates the authors' interests in how well the dependent random variables will behave in some cases.

One of the important dependence structures is the widely orthant dependence structure. The main purpose of the paper is to study the large deviation for the least squares estimator of the nonlinear regression model based on widely orthant dependent errors.

1.1 Brief review

Consider the following nonlinear regression model:

$$X_i = f_i(\theta) + \xi_i, \quad i \geq 1, \tag{1.1}$$

where $\{X_i\}$ is observed, $\{f_i(\theta)\}$ is a known sequence of continuous functions possibly nonlinear in $\theta \in \Theta$, Θ denotes a closed interval on the real line, and $\{\xi_i\}$ is a mean zero sequence of random errors. Denote

$$Q_n(\theta) = \frac{1}{2}\sum_{i=1}^{n} \omega_i^2 \bigl(x_i - f_i(\theta)\bigr)^2, \tag{1.2}$$

where $\{\omega_i\}$ is a known sequence of positive numbers. An estimator θ_n is said to be a least squares estimator of θ if it minimizes $Q_n(\theta)$ over $\theta \in \Theta$, *i.e.* $Q_n(\theta_n) = \inf_{\theta\in\Theta} Q_n(\theta)$.

Noting that $Q(x_1, x_2, \ldots, x_n; \theta) = Q_n(\theta)$ is defined on $\mathbf{R}^n \times \Theta$, where Θ is compact. Furthermore, $Q(x;\theta)$, where $x = (x_1, x_2, \ldots, x_n)$, is a Borel measurable function of x for any fixed $\theta \in \Theta$ and a continuous function of θ for any fixed $x \in \mathbf{R}^n$. Lemma 3.3 of Schmetterer [1] shows that there exists a Borel measurable map $\theta_n : \mathbf{R}^n \to \Theta$ such that $Q_n(\theta_n) = \inf_{\theta\in\Theta} Q_n(\theta)$. In the following, we will consider this version as the least squares estimator θ_n.

Let θ_0 be the true parameter and assume that $\theta_0 \in \Theta$. Ivanov [2] established the following large deviation result for independent and identically distributed (i.i.d.) random variables.

Theorem 1.1 *Let $\{\xi_i, i \geq 1\}$ be i.i.d. with $E|\xi_i|^p < \infty$ for some $p > 2$. Suppose that there exist some constants $0 < c_1 < c_2 < \infty$ such that*

$$c_1(\theta_1 - \theta_2)^2 \leq \frac{1}{n}\sum_{i=1}^{n}\bigl(f_i(\theta_1) - f_i(\theta_2)\bigr)^2 \leq c_2(\theta_1 - \theta_2)^2, \quad \forall n \geq 1, \tag{1.3}$$

for all $\theta_1, \theta_2 \in \Theta$. Then, for every $\rho > 0$, it has

$$P\bigl(n^{1/2}|\theta_n - \theta_0| > \rho\bigr) \leq c\rho^{-p}, \quad \forall n \geq 1, \tag{1.4}$$

where c is a positive constant independent of n and ρ.

Hu [3] also got the result (1.4) and gave its application to martingale difference, φ-mixing sequence and negatively associated (NA, in short) sequence. In addition, Hu [4] proved the following large deviation result:

$$P\bigl(n^{1/2}|\theta_n - \theta_0| > \rho\bigr) \leq cn^{1-\rho/2}\rho^{-p}, \tag{1.5}$$

under the condition that $\sup_{n\geq 1} E|\xi_n|^p < \infty$ for some $1 < p \leq 2$, and Hu gave its application to the martingale difference, the φ-mixing sequence, the NA sequence, and the weakly stationary linear process. Recently, Yang and Hu [5] obtained some large deviation results based on $\tilde{\rho}$-mixing, asymptotically almost negatively associated, negatively orthant dependent and L_p-mixingales random errors. For more details as regards the nonlinear regression model, one can refer to Ibregimov and Has'minskii [6], Ivanov and Leonenko [7], Ivanov [8], and so on. In this paper, the large deviation results for the least squares estimator of the nonlinear regression model based on the WOD error will be investigated.

Inspired by the above literature, we will establish the large deviation results based on widely orthant dependent errors.

1.2 Concept of widely orthant dependence structure

In this section, we will present the widely orthant dependence structure, which was introduced by Wang *et al.* [9].

Definition 1.1 For the random variables $\{X_n, n \geq 1\}$, if there exists a finite real sequence $\{f_U(n), n \geq 1\}$ satisfying for each $n \geq 1$ and for all $x_i \in \mathbf{R}$, $1 \leq i \leq n$,

$$P(X_1 > x_1, X_2 > x_2, \dots, X_n > x_n) \leq f_U(n) \prod_{i=1}^{n} P(X_i > x_i), \tag{1.6}$$

then we say that the $\{X_n, n \geq 1\}$ are widely upper orthant dependent (WUOD, in short); if there exists a finite real sequence $\{f_L(n), n \geq 1\}$ satisfying for each $n \geq 1$ and for all $x_i \in \mathbf{R}$, $1 \leq i \leq n$,

$$P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) \leq f_L(n) \prod_{i=1}^{n} P(X_i \leq x_i), \tag{1.7}$$

then we say that the $\{X_n, n \geq 1\}$ are widely lower orthant dependent (WLOD, in short); if they are both WUOD and WLOD, then we say that the $\{X_n, n \geq 1\}$ are widely orthant dependent (WOD, in short), and $f_U(n), f_L(n)$, $n \geq 1$ are called dominating coefficients.

An array $\{X_{ni}, i \geq 1, n \geq 1\}$ of random variables is called a row-wise WOD if for every $n \geq 1$, $\{X_{ni}, i \geq 1\}$ is a sequence of WOD random variables.

As mentioned above, Wang *et al.* [9] first introduced the concept of WOD random variables. Their properties and applications have been studied consequently. For instance, WOD random variables include some common negatively dependent random variables, some positively dependent random variables and others, which were shown in the examples provided by Wang *et al.* [9] and the uniform asymptotic for the finite-time ruin probability of a new dependent risk model with a constant interest rate was also investigated in the same work. He *et al.* [10] established the asymptotic lower bounds of precise large deviations with non-negative and dependent random variables. The uniform asymptotic for the finite time ruin probabilities of two types of non-standard bidimensional renewal risk models with constant interest forces and diffusion generated by Brownian motions was proposed by Chen *et al.* [11]. The Bernstein type inequality for WOD random variables and its applications were studied by Shen [12]. Wang *et al.* [13] investigated the complete convergence for WOD random variables and gave its applications to nonparametrics regression models, and so forth.

As is well known, the class of WOD random variables contains END random variables, NOD random variables, NSD random variables, NA random variables, and independent random variables as special cases. Hence, it is meaningful to extend the results of Yang and Hu [5] to WOD errors.

Throughout this paper, let $\{\xi_i, i \geq 1\}$ be a sequence of WOD random variables with dominating coefficients $f_U(n), f_L(n)$, $n \geq 1$. Denote $f(n) = \max\{f_U(n), f_L(n)\}$. Let C denote a positive constant, which may vary in different spaces. Let $\lfloor x \rfloor$ be the integer part of x.

The main results and their proofs are presented in Section 3 and for the convenience of the reader, some useful lemmas relating to the proofs are listed in Section 2.

2 Preliminary lemmas

In this section, we provide some important lemmas will be used to prove the main results of the paper. The first one is the basic property for WOD random variables, which was established by Wang *et al.* [13].

Lemma 2.1 *Let $\{X_n, n \geq 1\}$ be a sequence of WOD random variables.*

(i) *If $\{h_n(\cdot), n \geq 1\}$ are all non-decreasing (or all non-increasing), then $\{h_n(X_n), n \geq 1\}$ are still WOD.*

(ii) *For each $n \geq 1$ and any $s \in \mathbf{R}$,*

$$E\exp\left\{s\sum_{i=1}^{n}X_i\right\} \leq f(n)\prod_{i=1}^{n}E\exp\{sX_i\}. \tag{2.1}$$

The next lemma is very useful to prove the main results of the paper, which can be found in Hu [4].

Lemma 2.2 *Let $\{\Omega, \mathcal{F}, P\}$ be a probability space, $[T_1, T_2]$ be a closed interval on the real line. Assume that $V(\theta) = V(\omega, \theta)$ $(\theta \in [T_1, T_2], \omega \in \Omega)$ is a stochastic process such that $V(\omega, \theta)$ is continuous for all $\omega \in \Omega$. If there exist numbers $\alpha > 0$, $r > 0$ and $C = C(T_1, T_2) < \infty$ such that*

$$E\left|V(\theta_1) - V(\theta_2)\right|^r \leq C|\theta_1 - \theta_2|^{1+\alpha}, \quad \forall \theta_1, \theta_2 \in [T_1, T_2],$$

then for any $\epsilon > 0$, $a > 0$, $\theta_0, \theta_0 + \epsilon \in [T_1, T_2]$, $\gamma \in (2, 2+\alpha)$, one has

$$P\Big(\sup_{\theta_0 \leq \theta_1, \theta_2 \leq \theta_0 + \epsilon}\left|V(\theta_1) - V(\theta_2)\right| \geq a\Big) \leq \frac{8C}{(\alpha - \gamma + 2)(\alpha - \gamma + 3)}\left(\frac{8\gamma}{\gamma - 2}\right)^r\frac{\epsilon^{\alpha+1}}{a^r}. \tag{2.2}$$

The following are the Marcinkiewicz-Zygmund type inequality and Rosential-type inequality for WOD random variables, which play an important role in the proof.

Lemma 2.3 (*cf.* Wang *et al.* [13]) *Let $p \geq 1$ and $\{X_n, n \geq 1\}$ be a sequence of WOD random variables with $EX_n = 0$ and $E|X_n|^p < \infty$ for each $n \geq 1$. Then there exist positive constants $C_1(p)$ and $C_2(p)$ depending only on p such that, for $1 \leq p \leq 2$,*

$$E\left|\sum_{i=1}^{n}X_i\right|^p \leq \left[C_1(p) + C_2(p)f(n)\right]\sum_{i=1}^{n}E|X_i|^p, \tag{2.3}$$

and for $p > 2$,

$$E\left|\sum_{i=1}^{n}X_i\right|^p \leq C_1(p)\sum_{i=1}^{n}E|X_i|^p + C_2(p)f(n)\left(\sum_{i=1}^{n}E|X_i|^2\right)^{p/2}. \tag{2.4}$$

3 Main results and their proofs

Based on the useful inequalities in Section 2, we now study the large deviation results for the least squares estimator of the nonlinear regression model based on WOD errors.

Theorem 3.1 *Consider the model* (1.1). *Assume that there exist positive constants c_1, c_2, c_3, c_4 such that*

$$c_1|\theta_1 - \theta_2| \leq \left|f_i(\theta_1) - f_i(\theta_2)\right| \leq c_2|\theta_1 - \theta_2|, \quad \forall \theta_1, \theta_2 \in \Theta, i \geq 1, \tag{3.1}$$

and

$$c_3 \le \omega_i \le c_4, \quad \forall i \ge 1. \tag{3.2}$$

Let $\{\xi_i, i \ge 1\}$ *be a sequence of mean zero WOD random variables with* $E|\xi_i|^p < \infty$ *for some* $p > 2$. *Denote*

$$\Gamma_{p,n} = \sum_{i=1}^{n} E|\xi_i|^p, \quad n \ge 1, \tag{3.3}$$

and

$$\Delta_{p,n} = \left(\sum_{i=1}^{n} \bigl(E|\xi_i|^p\bigr)^{2/p} \right)^{p/2}, \quad n \ge 1. \tag{3.4}$$

Then there exists a positive constant $C(p)$ *such that*

$$P\bigl(n^{1/2}|\theta_n - \theta_0| > \rho\bigr) \le C(p)\bigl(\Gamma_{p,n} + f(n)\Delta_{p,n}\bigr)n^{-p/2}\rho^{-p} \tag{3.5}$$

for every $\rho > 0$ *and all* $n \ge 1$.

Proof Denote

$$\Psi_n(\theta_1, \theta_2) = \frac{1}{n} \sum_{i=1}^{n} \omega_i^2 \bigl(f_i(\theta_1) - f_i(\theta_2)\bigr), \qquad V_n(\theta) = \frac{1}{n^{1/2}} \sum_{i=1}^{n} \xi_i \bigl(f_i(\theta) - f_i(\theta_0)\bigr)^2$$

and

$$U_n(\theta) = \frac{V_n(\theta)}{n^{1/2}\Psi_n(\theta, \theta_0)}, \quad \theta \ne \theta_0.$$

Without loss of generality, we assume that $\omega_i = 1$ for all i. The general case can be obtained similarly in view of (3.2). It follows from (3.1) that

$$c_1^2(\theta_1 - \theta_2)^2 \le \Psi_n(\theta_1, \theta_2) \le c_2^2(\theta_1 - \theta_2)^2 \tag{3.6}$$

for all $\theta_1, \theta_2 \in \theta$ and $n \ge 1$. Denote $A_{n\epsilon} = [|\theta_n - \theta_0| > \epsilon]$. If $\epsilon \in A_{n\epsilon}$, then $\theta_n \ne \theta_0$ and

$$\begin{aligned}
\sum_{i=1}^{n} \xi_i^2 &\ge \sum_{i=1}^{n} \bigl(X_i - f_i(\theta_n)\bigr)^2 \\
&= \sum_{i=1}^{n} \bigl(X_i - f_i(\theta_0)\bigr)^2 + 2\sum_{i=1}^{n} \bigl(X_i - f_i(\theta_0)\bigr)\bigl(f_i(\theta_0) - f_i(\theta_n)\bigr) \\
&\quad + \sum_{i=1}^{n} \bigl(f_i(\theta_0) - f_i(\theta_n)\bigr)^2 \\
&= \sum_{i=1}^{n} \xi_i^2 - 2nU_n(\theta_n)\Psi_n(\theta_n, \theta_0) + n\Psi_n(\theta_n, \theta_0),
\end{aligned}$$

which implies that

$$\Psi_n(\theta_n,\theta_0)\bigl(1-2U_n(\theta_n)\bigr)\le 0. \tag{3.7}$$

Noting that $\theta_n \neq \theta_0$, we have by (3.6) that $\Psi_n(\theta_n,\theta_0) > 0$, which together with (3.7) shows that $U_n(\theta_n) \ge 1/2$. Thus, $A_{n\epsilon} = [|\theta_n - \theta_0| > \epsilon] \subset [U_n(\theta_n) \ge 1/2]$, and, for any $\epsilon > 0$,

$$P\bigl(|\theta_n-\theta_0|>\epsilon\bigr)\le P\Bigl(\sup_{|\theta_n-\theta_0|>\epsilon} U_n(\theta_n)\ge 1/2\Bigr). \tag{3.8}$$

Putting $\epsilon = \rho n^{-1/2}$ in (3.8), we have

$$\begin{aligned} P\bigl(n^{1/2}|\theta_n-\theta_0|>\rho\bigr) &\le P\Bigl(\sup_{|\theta_n-\theta_0|>\rho}\bigl|U_n(\theta)\bigr|\ge 1/2\Bigr)\\ &\quad + P\Bigl(\sup_{\rho n^{-1/2}<\left|\theta_n-\theta_0\right|\le\rho}\bigl|U_n(\theta)\bigr|\ge 1/2\Bigr) \end{aligned} \tag{3.9}$$

for every $\rho > 0$. It follows from (3.6) again that

$$\begin{aligned} \sup_{|\theta-\theta_0|>\rho}\frac{|V_n(\theta)|}{n^{1/2}\Psi_n(\theta,\theta_0)} &= \sup_{|\theta-\theta_0|>\rho}\frac{|V_n(\theta)|}{n^{1/2}\Psi_n^{1/2}(\theta,\theta_0)\Psi_n^{1/2}(\theta,\theta_0)}\\ &\le \sup_{|\theta-\theta_0|>\rho}\frac{|V_n(\theta)|}{n^{1/2}\Psi_n^{1/2}(\theta,\theta_0)\rho c_1}. \end{aligned} \tag{3.10}$$

Hence,

$$\begin{aligned} P\Bigl(\sup_{|\theta-\theta_0|>\rho}\bigl|U_n(\theta)\bigr|\ge 1/2\Bigr) &= P\Bigl(\sup_{|\theta-\theta_0|>\rho}\frac{|V_n(\theta)|}{n^{1/2}\Psi_n(\theta,\theta_0)}\ge 1/2\Bigr)\\ &\le P\Bigl(\sup_{|\theta-\theta_0|>\rho}\frac{|V_n(\theta)|}{n^{1/2}\Psi_n^{1/2}(\theta,\theta_0)}\ge\frac{c_1\rho}{2}\Bigr). \end{aligned} \tag{3.11}$$

Cauchy's inequality yields

$$\left(\frac{|V_n(\theta)|}{n^{1/2}\Psi_n^{1/2}(\theta,\theta_0)}\right)^2 = \left(\frac{1}{n}\sum_{i=1}^n \xi_i\frac{f_i(\theta)-f_i(\theta_0)}{\Psi_n^{1/2}(\theta,\theta_0)}\right)^2 \le \frac{1}{n}\sum_{i=1}^n \xi_i^2,\quad \forall\theta\neq\theta_0. \tag{3.12}$$

Noting hat $p > 2$, we have by Minkowski's inequality

$$\left(E\left(\sum_{i=1}^n \xi_i^2\right)^{p/2}\right)^{2/p}\le\sum_{i=1}^n\bigl(E|\xi_i|^p\bigr)^{2/p}. \tag{3.13}$$

Hence, we can obtain by Markov's inequality, (3.11)-(3.13) and $f(n) \ge 1$ (from the definition of WOD random variables)

$$\begin{aligned} P\Bigl(\sup_{|\theta-\theta_0|>\rho}\bigl|U_n(\theta)\bigr|\ge 1/2\Bigr) &\le P\left(\frac{1}{n}\sum_{i=1}^n \xi_i^2\ge\left(\frac{1}{2}c_1\rho\right)^2\right)\\ &\le\left(\frac{4}{nc_1^2\rho^2}\right)^{p/2}E\left(\sum_{i=1}^n\xi_i^2\right)^{p/2} \end{aligned}$$

$$\begin{aligned}
&\le \left(\frac{4}{nc_1^2\rho^2}\right)^{p/2}\left(\sum_{i=1}^{n}\bigl(E|\xi_i|^p\bigr)^{2/p}\right)^{p/2} \\
&\le C_1(p)n^{-p/2}\rho^{-p}\Delta_{p,n} \\
&\le C_1(p)n^{-p/2}\rho^{-p}f(n)\Delta_{p,n}. \qquad (3.14)
\end{aligned}$$

For $m = 0,1,2,\dots,\lfloor n^{1/2}\rfloor$, denote $\theta(m) = \theta_0 + \frac{\rho}{n^{1/2}} + \frac{m\rho}{\lfloor n^{1/2}\rfloor}$, $\rho_m = \theta(m) - \theta_0(> 0)$. It follows from (3.6) again that

$$\begin{aligned}
\sup_{\rho_m\le\theta-\theta_0\le\rho_{m+1}}\bigl|U_n(\theta)\bigr| &\le \sup_{\rho_m\le\theta-\theta_0\le\rho_{m+1}}\frac{|V_n(\theta)|}{n^{1/2}c_1^2(\theta-\theta_0)^2} \\
&\le \sup_{\rho_m\le\theta-\theta_0\le\rho_{m+1}}\frac{|V_n(\theta)|}{n^{1/2}c_1^2\rho_m^2}, \qquad (3.15)
\end{aligned}$$

and thus

$$\begin{aligned}
P\Bigl(\sup_{\rho n^{-1/2}\le\theta-\theta_0\le\rho}\bigl|U_n(\theta)\bigr| \ge \frac{1}{2}\Bigr) &\le \sum_{m=0}^{\lfloor n^{1/2}\rfloor-1} P\Bigl(\sup_{\rho_m\le\theta-\theta_0\le\rho_{m+1}}\bigl|U_n(\theta)\bigr| \ge \frac{1}{2}\Bigr) \\
&\le \sum_{m=0}^{\lfloor n^{1/2}\rfloor-1} P\Bigl(\sup_{\rho_m\le\theta-\theta_0\le\rho_{m+1}}\bigl|V_n(\theta)\bigr| \ge \frac{1}{2}c_1^2\rho_m^2 n^{1/2}\Bigr).
\end{aligned}$$

Noting that

$$\sup_{\rho_m\le\theta-\theta_0\le\rho_{m+1}}\bigl|V_n(\theta)\bigr| \le \bigl|V_n\bigl(\theta(m)\bigr)\bigr| + \sup_{\theta(m)\le\theta_1,\theta_2\le\theta(m+1)}\bigl|V_n(\theta_2)-V_n(\theta_1)\bigr|,$$

we have

$$\begin{aligned}
&P\Bigl(\sup_{\rho_m\le\theta-\theta_0\le\rho_{m+1}}\bigl|V_n(\theta)\bigr| \ge \frac{1}{2}c_1^2\rho_m^2 n^{1/2}\Bigr) \\
&\quad\le P\Bigl(\bigl|V_n\bigl(\theta(m)\bigr)\bigr| \ge \frac{1}{4}c_1^2\rho_m^2 n^{1/2}\Bigr) \\
&\qquad + P\Bigl(\sup_{\theta(m)\le\theta_1,\theta_2\le\theta(m+1)}\bigl|V_n(\theta_2)-V_n(\theta_1)\bigr| \ge \frac{1}{4}c_1^2\rho_m^2 n^{1/2}\Bigr). \qquad (3.16)
\end{aligned}$$

In view of the definition of $V_n(\theta)$, it is easy to check that

$$V_n\bigl(\theta(m)\bigr) = \frac{1}{n^{1/2}}\sum_{i=1}^{n}\xi_i\bigl(f_i\bigl(\theta(m)\bigr)-f_i(\theta_0)\bigr),\quad V_n(\theta_2)-V_n(\theta_1) = \frac{1}{n^{1/2}}\sum_{i=1}^{n}\xi_i\bigl(f_i(\theta_2)-f_i(\theta_1)\bigr).$$

By Markov's inequality, (2.4) in Lemma 2.3, (3.1), and Hölder's inequality, we get

$$\begin{aligned}
&P\Bigl(\bigl|V_n\bigl(\theta(m)\bigr)\bigr| \ge \frac{1}{4}c_1^2\rho_m^2 n^{1/2}\Bigr) \\
&\quad\le \left(\frac{4}{c_1^2\rho_m^2 n^{1/2}}\right)^{p} E\bigl|V_n\bigl(\theta(m)\bigr)\bigr|^p
\end{aligned}$$

$$
\begin{aligned}
&\le \left(\frac{4}{c_1^2\rho_m^2 n^{1/2}}\right)^p \frac{1}{n^{p/2}}\left\{C_1(p)\sum_{i=1}^n E|\xi_i|^p\big|f_i(\theta(m))-f_i(\theta_0)\big|^p\right.\\
&\quad\left.+C_2(p)f(n)\left(\sum_{i=1}^n E\xi_i^2\big(f_i(\theta(m))-f_i(\theta_0)\big)^2\right)^{p/2}\right\}\\
&\le \left(\frac{4c_2}{c_1^2\rho_m^2 n^{1/2}}\right)^p \frac{1}{n^{p/2}}\big|\theta(m)-\theta_0\big|^p\left\{C_1(p)\sum_{i=1}^n E|\xi_i|^p+C_2(p)f(n)\left(\sum_{i=1}^n E\xi_i^2\right)^{p/2}\right\}\\
&\le \left(\frac{4c_2}{c_1^2\rho_m^2 n^{1/2}}\right)^p \frac{1}{n^{p/2}}\big|\theta(m)-\theta_0\big|^p\left\{C_1(p)\sum_{i=1}^n E|\xi_i|^p+C_2(p)f(n)\left(\sum_{i=1}^n \big(E|\xi_i|^p\big)^{2/p}\right)^{p/2}\right\}\\
&\le C_3(p)\rho_m^{-p}n^{-p}\big(\Gamma_{p,n}+f(n)\Delta_{p,n}\big). \qquad (3.17)
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
&E\big|V_n(\theta_2)-V_n(\theta_1)\big|^p\\
&\quad\le \frac{1}{n^{p/2}}\left\{C_4(p)\sum_{i=1}^n E|\xi_i|^p\big|f_i(\theta_2)-f_i(\theta_1)\big|^p+C_5(p)f(n)\left(\sum_{i=1}^n E\xi_i^2\big(f_i(\theta_2)-f_i(\theta_1)\big)^2\right)^{p/2}\right\}\\
&\quad\le \frac{c_2}{n^{p/2}}\left(C_4(p)\sum_{i=1}^n E|\xi_i|^p+C_5(p)f(n)\left(\sum_{i=1}^n E\xi_i^2\right)^{p/2}\right)|\theta_2-\theta_1|^p\\
&\quad\le \frac{C_6(p)}{n^{p/2}}\left(\sum_{i=1}^n E|\xi_i|^p+f(n)\left(\sum_{i=1}^n \big(E|\xi_i|^p\big)^{2/p}\right)^{p/2}\right)|\theta_2-\theta_1|^p\\
&\quad= \frac{C_6(p)}{n^{p/2}}\big(\Gamma_{n,p}+f(n)\Delta_{p,n}\big)|\theta_1-\theta_2|^p\\
&\quad=: C(n,p)|\theta_1-\theta_2|^p. \qquad (3.18)
\end{aligned}
$$

For $\forall\,\theta_1,\theta_2\in\Theta$ and $n\ge 1$, applying Lemma 2.2 with $r=1+\alpha=p$, $C=C(n,p)$, $\epsilon=\rho/\lfloor n^{-1/2}\rfloor$, $a=\frac{1}{4}c_1^2\rho_m^2 n^{1/2}$, and $\gamma\in(2,p+1)$, we can obtain

$$
\begin{aligned}
&P\Big(\sup_{\theta(m)\le\theta_1,\theta_2\le\theta(m+1)}\big|V_n(\theta_2)-V_n(\theta_1)\big|\ge\frac{1}{4}c_1^2\rho_m^2 n^{1/2}\Big)\\
&\quad=P\Big(\sup_{\theta(m)\le\theta_1,\theta_2\le\theta(m)+\rho/\lfloor n^{-1/2}\rfloor}\big|V_n(\theta_2)-V_n(\theta_1)\big|\ge\frac{1}{4}c_1^2\rho_m^2 n^{1/2}\Big)\\
&\quad\le \frac{8C_6(p)n^{-p/2}(\Gamma_{p,n}+f(n)\Delta_{p,n})}{(p+1-\gamma)(p+2-\gamma)}\left(\frac{8\gamma}{\gamma-2}\right)^p\left(\frac{\rho}{\lfloor n^{-1/2}\rfloor}\right)^p\left(\frac{4}{c_1^n\rho_m^2 n^{1/2}}\right)^p\\
&\quad\le C_7(p)\rho^p\big(\Gamma_{p,n}+f(n)\Delta_{p,n}\big)n^{-3p/2}\rho_m^{-2p}. \qquad (3.19)
\end{aligned}
$$

Noting that $\rho_0=\rho n^{-1/2}$, $\rho_m>m\rho n^{-1/2}$, $p>2$, we have by (3.16), (3.17), and (3.19) that

$$
\begin{aligned}
&P\Big(\sup_{\rho n^{-1/2}\le\theta-\theta_0\le\rho}\big|U_n(\theta)\big|\ge\frac{1}{2}\Big)\\
&\quad\le\sum_{m=0}^{\lfloor n^{1/2}\rfloor-1}\big\{C_3(p)\rho_m^{-p}n^{-p}\big(\Gamma_{p,n}+f(n)\Delta_{p,n}\big)+C_7(p)\rho_m^{-2p}n^{-3p/2}\rho^p\big(\Gamma_{p,n}+f(n)\Delta_{p,n}\big)\big\}
\end{aligned}
$$

$$\begin{aligned}
&\leq C_3(p)n^{-p/2}\rho^{-p}\bigl(\Gamma_{p,n}+f(n)\Delta_{p,n}\bigr)+C_7(p)n^{-p/2}\rho^{-p}\bigl(\Gamma_{p,n}+f(n)\Delta_{p,n}\bigr)\\
&\quad+n^{-p/2}\rho^{-p}\bigl(\Gamma_{p,n}+f(n)\Delta_{p,n}\bigr)\sum_{m=1}^{\lfloor n^{1/2}\rfloor-1}\left(\frac{C_3(p)}{m^p}+\frac{C_7(p)}{m^{2p}}\right)\\
&\leq C_8(p)n^{-p/2}\rho^{-p}\bigl(\Gamma_{p,n}+f(n)\Delta_{p,n}\bigr). \qquad (3.20)
\end{aligned}$$

Similarly, we have

$$P\Bigl(\sup_{\rho n^{-1/2}\leq\theta_0-\theta\leq\rho}\bigl|U_n(\theta)\bigr|\geq\frac{1}{2}\Bigr)\leq C_9(p)n^{-p/2}\rho^{-p}\bigl(\Gamma_{p,n}+f(n)\Delta_{p,n}\bigr). \qquad (3.21)$$

Therefore, the desired result (3.5) follows from (3.9), (3.14), (3.20), and (3.21) immediately. This completes the proof of the theorem. □

Inspired by Theorem 3.1, we will consider the case $p\in(1,2]$ and establish the following result.

Theorem 3.2 *Consider the model* (1.1). *Let the conditions* (3.1) *and* (3.2) *in Theorem* 3.1 *hold, and* $E|\xi_i|^p<\infty$ *for some* $p\in(1,2]$. *Denote*

$$\Lambda_{p,n}=\sum_{i=1}^{n}E|\xi_i|^p,\quad n\geq 1. \qquad (3.22)$$

Then there exists a positive constant $C(p)$ *such that*

$$P\bigl(n^{1/2}|\theta_n-\theta_0|>\rho\bigr)\leq C(p)f(n)\Lambda_{p,n}n^{-p/2}\rho^{-p} \qquad (3.23)$$

for every $\rho>0$ *and all* $n\geq 1$.

Proof Similar to the above proof, we have by the C_r inequality, (3.1), and (3.6)

$$\begin{aligned}
\left|\frac{V_n(\theta)}{n^{1/2}\Psi_n^{1/2}(\theta,\theta_0)}\right|^p &= \left|\frac{1}{n}\sum_{i=1}^{n}\xi_i\frac{f_i(\theta)-f_i(\theta_0)}{\Psi_n^{1/2}(\theta,\theta_0)}\right|^p\\
&\leq\frac{1}{n^p}n^{p-1}\sum_{i=1}^{n}|\xi_i|^p\frac{|f_i(\theta)-f_i(\theta_0)|^p}{\Psi_n^{p/2}(\theta,\theta_0)}\\
&\leq\frac{C_1(p)}{n}\sum_{i=1}^{n}|\xi_i|^p,\quad\forall\theta\neq\theta_0, \qquad (3.24)
\end{aligned}$$

which implies that

$$\begin{aligned}
P\Bigl(\sup_{|\theta-\theta_0|>\rho}\bigl|U_n(\theta)\bigr|\geq\frac{1}{2}\Bigr) &\leq P\left(\frac{C_1(p)}{n}\sum_{i=1}^{n}|\xi_i|^p\geq\left(\frac{1}{2}c_1\rho\right)^p\right)\\
&\leq\left(\frac{2}{c_1\rho}\right)^p\frac{C_1(p)}{n}\sum_{i=1}^{n}E|\xi_i|^p\\
&\leq C_2(p)n^{-1}\rho^{-p}\Lambda_{p,n}. \qquad (3.25)
\end{aligned}$$

Similar to the proof of (3.17), we see by Markov's inequality, (2.3) in Lemma 2.3, and (3.1) that

$$\begin{aligned} &P\left(\left|V_n(\theta(m))\right| \geq \frac{1}{4}c_1^2\rho_m^2 n^{1/2}\right) \\ &\quad\leq \left(\frac{4c_2}{c_1^2\rho_m^2 n^{1/2}}\right)^p \frac{1}{n^{p/2}}\left|\theta(m)-\theta_0\right|^p \bigl(C_3(p)+C_4(p)f(n)\bigr)\Lambda_{p,n} \\ &\quad\leq C_5(p)\bigl(1+f(n)\bigr)\rho_m^{-p}n^{-p}\Lambda_{p,n}. \end{aligned} \tag{3.26}$$

On the other hand, for all θ_1, θ_2 and $n \geq 1$, we have

$$E\left|V_n(\theta_2)-V_n(\theta_1)\right|^p \leq \frac{C_6(p)}{n^{p/2}}\bigl(1+f(n)\bigr)\Lambda_{p,n}|\theta_1-\theta_2|^p =: C(n,p)|\theta_1-\theta_2|^p. \tag{3.27}$$

In view of the proof of (3.19) and noting that $f(n) \geq 1$, we have

$$\begin{aligned} &P\left(\sup_{\theta(m)\leq\theta_1,\theta_2\leq\theta(m+1)}\left|V_n(\theta_2)-V_n(\theta_1)\right| \geq \frac{1}{4}c_1^2\rho_m^2 n^{1/2}\right) \\ &\quad\leq \frac{8C_6(p)n^{-p/2}(1+f(n))\Lambda_{p,n}}{(p+1-\gamma)(p+2-\gamma)}\left(\frac{8\gamma}{\gamma-2}\right)^p\left(\frac{\rho}{\lfloor n^{1/2}\rfloor}\right)^p\left(\frac{4}{c_1^2\rho_m^2 n^{1/2}}\right)^p \\ &\quad\leq C_7(p)\rho^p f(n)\Lambda_{p,n}n^{-3p/2}\rho_m^{-2p}. \end{aligned} \tag{3.28}$$

Following a similar way, we can get the proof below:

$$\begin{aligned} &P\left(\sup_{\rho n^{-1/2}\leq\theta-\theta_0\leq\rho}\left|U_n(\theta)\right| \geq \frac{1}{2}\right) \\ &\quad\leq \sum_{m=0}^{\lfloor n^{\frac{1}{2}}\rfloor-1}\left\{C_5(p)\bigl(1+f(n)\bigr)\rho_m^{-p}n^{-p}\Lambda_{p,n} + C_7(p)\rho^p\bigl(1+f(n)\bigr)\Lambda_{p,n}n^{-3p/2}\rho_m^{-2p}\right\} \\ &\quad\leq C_5(p)\bigl(1+f(n)\bigr)n^{-p/2}\Lambda_{p,n}\rho^{-p} + C_7(p)\bigl(1+f(n)\bigr)n^{-p/2}\Lambda_{p,n}\rho^{-p} \\ &\qquad + n^{-p/2}\rho^{-p}\Lambda_{p,n}\bigl(1+f(n)\bigr)\sum_{m=1}^{\lfloor n^{\frac{1}{2}}\rfloor-1}\left(\frac{C_5(p)}{m^p}+\frac{C_7(p)}{m^{2p}}\right) \\ &\quad\leq C_8(p)n^{-p/2}f(n)\Lambda_{p,n}\rho^{-p}, \end{aligned} \tag{3.29}$$

and thus

$$P\left(\sup_{\rho n^{-1/2}\leq\theta_0-\theta\leq\rho}\left|U_n(\theta)\right| \geq \frac{1}{2}\right) \leq C(p)n^{-p/2}f(n)\Lambda_{p,n}\rho^{-p}. \tag{3.30}$$

Therefore, the desired result (3.23) follows from (3.9), (3.25), (3.29), and (3.30) immediately. The proof is completed. □

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Author details
[1]Wendian College, Anhui University, Hefei, 230601, P.R. China. [2]School of Mathematical Sciences, Anhui University, Hefei, 230601, P.R. China.

Acknowledgements
This work was supported by the National Natural Science Foundation of China (11501004, 11501005, 11526033), the Natural Science Foundation of Anhui Province (1508085J06), the Students Science Research Training Program of Anhui University (KYXL2014017), and the Research Teaching Model Curriculum of Anhui University (xjyjkc1407).
The authors are most grateful to the editor and anonymous referees for careful reading of the manuscript and valuable suggestions, which helped in improving an earlier version of this paper.

Received: 8 October 2015 Accepted: 28 March 2016 Published online: 23 April 2016

References
1. Schmetterer, L: Introduction to Mathematical Statistics. Springer, Berlin (1974)
2. Ivanov, AV: An asymptotic expansion for the distribution of the least squares estimator of the nonlinear regression parameter. Theory Probab. Appl. **21**(3), 557-570 (1976)
3. Hu, SH: The rate of convergence for the least squares estimator in nonlinear regression model with dependent errors. Sci. China Ser. A **45**(2), 137-146 (2002)
4. Hu, SH: Consistency for the least squares estimator in nonlinear regression model. Stat. Probab. Lett. **67**(2), 183-192 (2004)
5. Yang, WZ, Hu, SH: Large deviation for a least squares estimator in a nonlinear regression model. Stat. Probab. Lett. **91**, 135-144 (2014)
6. Ibragimov, IA, Has'minskii, RZ: Statistical Estimation: Asymptotic Theory. Springer, New York (1981). Translated by Samuel Kotz
7. Ivanov, AV, Leonenko, NN: Statistical Analysis of Random Fields. Kluwer Academic Publishers, Dordreht/Boston/London (1989)
8. Ivanov, AV: Asymptotic Theory of Nonlinear Regression. Kluwer Academic Publishers, Dordreht/Boston/London (1997)
9. Wang, K, Wang, Y, Gao, Q: Uniform asymptotics for the finite-time ruin probability of a new dependent risk model with a constant interest rate. Methodol. Comput. Appl. Probab. **15**(1), 109-124 (2013)
10. He, W, Cheng, DY, Wang, YB: Asymptotic lower bounds of precise large deviations with nonnegative and dependent random variables. Stat. Probab. Lett. **83**, 331-338 (2013)
11. Chen, Y, Wang, L, Wang, YB: Uniform asymptotics for the finite-time ruin probabilities of two kinds of nonstandard bidimensional risk models. J. Math. Anal. Appl. **401**(1), 114-129 (2013)
12. Shen, AT: Bernstein-type inequality for widely dependent sequence and its application to nonparametric regression models. Abstr. Appl. Anal. **2013**, Article ID 862602 (2013)
13. Wang, XJ, Xu, C, Hu, TC, Volodin, A, Hu, SH: On complete convergence for widely orthant-dependent random variables and its applications in nonparametrics regression models. Test **23**(3), 607-629 (2014)