УДК 519.2

Е. С Агеева, чл.-корр. НАН Беларуси Ю.С. ХАРИН

СОСТОЯТЕЛЬНОСТЬ ОЦЕНКИ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ ПАРАМЕТРОВ МНОЖЕСТВЕННОЙ РЕГРЕССИИ ПО КЛАССИФИЦИРОВАННЫМ НАБЛЮДЕНИЯМ

НИИ прикладных проблем математики и информатики БГУ, Минск *Поступило*

Введение. В математической статистике и ее приложениях широко используется регрессионная модель. Ею описываются многие процессы в технике, экономике, медицине и других областях. Хорошо исследованы случаи, когда зависимые переменные наблюдаются с выбросами или с пропусками; при этом построены робастные (устойчивые) статистические выводы [5, 11, 12, 14]. Более сложной для анализа является ситуация, когда в выборке присутствуют цензурированные наблюдения: об их значениях известно только, что они попадают в некоторые интервалы ненулевой длины [6]. В статье рассматривается новая модель наблюдений нелинейной множественной регрессии, когда вместо истинных значений зависимой переменной наблюдаются номера классов (интервалов), в которые попадают эти значения.

Предлагаемая в статье новая модель множественной регрессии с классифицированными наблюдениями является обобщением известной модели с "округленными данными" (rounded data) [1, 3, 7, 8].

Регрессионная модель при наличии классификации наблюдений. Пусть на вероятностном пространстве $(\Omega,\mathcal{F},\mathbf{P})$ определена модель нелинейной множественной регрессии с N регрессорами:

$$Y_t=F(X_t;\theta^0)+\xi_t,\ t=1,\dots,n,\dots, \tag{1}$$

где n – объём выборки; $F(\cdot):\mathbf{X}\times\Theta\to\mathbb{R}^1$ – известная с точностью до векторного параметра функция регрессии; $\theta^0=(\theta_1^0,\dots,\theta_m^0)'\in\Theta\subseteq\mathbb{R}^m$ – неизвестный истинный вектор-столбец параметров функции регрессии; $X_t=(X_t^1,\dots,X_t^N)'\in\mathbf{X}\subseteq\mathbb{R}^m$ – наблюдаемый неслучайный вектор-столбец регрессоров; $Y_t\in\mathbb{R}^1$ – зависимая переменная; $\xi_t\in\mathbb{R}^1$ – ненаблюдаемая случайная величина ошибок с нормальным распределением вероятностей $\mathcal{N}(0,(\sigma^0)^2)$ с математическим ожиданием $\mathbf{E}\{\xi_t\}=0$ и дисперсией $0<\mathbf{E}\{\xi_t^2\}=(\sigma^0)^2<\infty$; штрих обозначает транспонирование матрицы. Предполагается, что $\{\xi_t\}_{t=1}^n$ независимы в совокупности.

Частными случаями модели (1), широко применяемыми в приложениях, являются множественная линейная регрессия ($m=N+1$):

$$Y_t=\theta_1^0X_t^1+\dots+\theta_{m-1}^0X_t^{m-1}+\theta_m^0+\xi_t,\ t=1,\dots,n,\dots, \tag{2}$$

и простая линейная регрессия ($m=N+1=2$):

$$Y_t=\theta_1^0X_t^1+\theta_2^0+\xi_t,\ t=1,\dots,n,\dots.$$

Пусть задана последовательность K непересекающихся борелевских множеств ($K\geq 2$):

$$A_1,\dots,A_K\in\mathcal{B}(\mathbb{R}^1),\ \cup_{k=1}^K A_k=\mathbb{R}^1,\ A_i\cap A_j=\emptyset,\ i\neq j.$$

Эта система борелевских множеств задает классификацию зависимой переменной Y_t:

$$Y_t \text{ относится к классу } \nu_t, \text{ если } Y_t\in A_{\nu_t},\ \nu_t\in\mathbf{K}=\{1,\dots,K\}. \tag{3}$$

В дальнейшем условимся полагать, что множества $A_1,\dots,A_K\in\mathcal{B}(\mathbb{R}^1)$ являются интервалами:

$$A_k=(a_{k-1},a_k],\ k\in\mathbf{K}, \tag{4}$$

где $-\infty=a_0<a_1<\ldots<a_{K-1}<a_K<\infty$ – заданный упорядоченный набор границ интервалов.

Вместо точных значений зависимой переменной $Y_1,\ldots,Y_n$ наблюдаются лишь соответствующие номера классов $\nu_1,\ldots,\nu_n\in\mathbf{K}$. Задача заключается в том, чтобы по классифицированным наблюдениям $\nu_1,\ldots,\nu_n$ и значениям регрессоров $X_1,\ldots,X_n$ построить статистические оценки для неизвестного вектора параметров $\alpha^0=(\theta^0,(\sigma^0)^2)'\in W\subseteq\mathbb{R}^{m+1}$.

Общий вид оценок максимального правдоподобия (ОПМ) для $\boldsymbol{\alpha^0=(\theta^0,(\sigma^0)^2)'}$. Используя модельные предположения (1), (3), примем обозначения:

$$P_X(k;\alpha)=\mathbf{P}_{X,\alpha}\{Y\in A_k\}=\frac{1}{\sqrt{2\pi}\sigma}\int_{A_k}e^{-\frac{(z-F(X,\theta))^2}{2\sigma^2}}dz,\ k\in\mathbf{K},\ \alpha=(\theta,\sigma^2)'\in W, \tag{5}$$

где $\mathbf{P}_{X,\alpha}\{\cdot\}$ – вероятностная мера, порожденная нормальным распределением вероятностей $\mathcal{N}(F(X,\theta),\sigma^2)$ случайной величины Y при фиксированных значениях регрессора X и вектора параметров α. Обозначим $\Phi(\cdot)$ – функцию распределения вероятностей стандартного нормального закона $\mathcal{N}(0,1)$.

Лемма 1. *Если имеет место модель наблюдения (1), (3), (4), то логарифмическая функция правдоподобия допускает представление:*

$$l(\alpha)=\sum_{t=1}^{n}\ln\left(\Phi\left(\frac{a_{\nu_t}-F(X_t;\theta)}{\sigma}\right)-\Phi\left(\frac{a_{\nu_t-1}-F(X_t;\theta)}{\sigma}\right)\right). \tag{6}$$

Доказательство. В силу независимости $\{\nu_t\}_{t=1}^{n}$ с учетом (5) логарифмическая функция правдоподобия имеет вид $l(\alpha)=\sum_{t=1}^{n}\ln P_{X_t}(\nu_t;\alpha)$. Из (4), (5) имеем:

$$P_{X_t}(\nu_t;\alpha)=\frac{1}{\sqrt{2\pi}\sigma}\int_{a_{\nu_t-1}}^{a_{\nu_t}}e^{-\frac{(z-F(X_t,\theta))^2}{2\sigma^2}}dz=\Phi\left(\frac{a_{\nu_t}-F(X_t;\theta)}{\sigma}\right)-\Phi\left(\frac{a_{\nu_t-1}-F(X_t;\theta)}{\sigma}\right). \tag{7}$$

Подставляя (7) в предыдущее выражение, получаем (6). □

Максимизируя функцию $l(\alpha)$ по α, найдем ОМП:

$$\hat{\alpha}^n=(\hat{\theta}^n,(\hat{\sigma}^n)^2)':\ l(\hat{\alpha}^n)=\max_{\alpha\in W}l(\alpha). \tag{8}$$

Для решения нелинейной экстремальной задачи (6), (8) целесообразно применять численные методы [10].

Состоятельность ОМП $\boldsymbol{\hat{\alpha}^n=(\hat{\theta}^n,(\hat{\sigma}^n)^2)'}$. Модель наблюдения (1), (3), (4) является моделью с независимыми, но неодинаково распределенными наблюдениями, поэтому при выполнении ряда предположений к ней применима теорема о состоятельности по вероятности, доказанная Ходли в [4]. Сформулируем эту теорему применительно к модели наблюдений (1), (3), (4). Определим вспомогательные функции:

$$P_X(k;\alpha,\rho)=\sup_{|\alpha'-\alpha|\le\rho}P_X(k;\alpha'),\ \rho>0;\ \psi_X(k;r)=\sup_{|\alpha'|\ge r}P_X(k;\alpha'),\ k\in\mathbf{K},\ r>0;$$

$$R_X(k;\alpha^0,\alpha)=\ln\frac{P_X(k;\alpha)}{P_X(k;\alpha^0)},\ R_X(k;\alpha^0,\alpha,\rho)=\ln\frac{P_X(k;\alpha,\rho)}{P_X(k;\alpha^0)},\ V_X(k;\alpha^0,r)=\ln\frac{\psi_X(k;r)}{P_X(k;\alpha^0)},$$

а также дискретную случайную величину ν с распределением вероятностей (5): $\mathbf{P}_{X,\alpha}\{\nu=k\}=P_X(k;\alpha)$, $k\in\mathbf{K}$; обозначим $\mathbf{E}_{X,\alpha}\{\cdot\}$ и $\mathbf{D}_{X,\alpha}\{\cdot\}$ – математическое ожидание и дисперсию по этому распределению вероятностей. Для любой случайной величины ζ определим так называемую «усечённую» снизу случайную величину (для некоторой константы $B\ge0$):

$$\zeta^{(B)}=\begin{cases}\zeta, & \text{если } \zeta\ge-B;\\ -B, & \text{если } \zeta<-B.\end{cases}$$

Теорема 1. *Пусть для модели (1), (3), (4) выполнены следующие условия:*

C1. $\alpha\in W$, *где W – замкнутое подмножество $\mathbb{R}^{m+1}$.*

C2. $P_{X_t}(\nu_t;\alpha)$ *– функция, полунепрерывная сверху по* α, *равномерно по t.*

C3. *Существуют* $\rho^*=\rho^*(\theta)>0$ *и* $r>0$ *такие, что для некоторых* $\delta>0$ *и* $M>0$ *сразу для всех* t, $t=1,\ldots,n$, *выполнено*

(i) $\mathbf{E}_{X_t,\alpha^0}\{R_{X_t}^{(0)}(\nu_t;\alpha^0,\alpha,\rho)\}^{1+\delta}\le M$, 0≤ρ≤ρ*;

(ii) $\mathbf{E}_{X_t,\alpha^0}\{V_{X_t}^{(0)}(\nu_t;\alpha^0,r)\}^{1+\delta}\le M$.

С4. *Существует* B>0*, для которого*

(i) $\overline{\lim_{n\to\infty}}(\frac{1}{n}\sum_{t=1}^{n}\mathbf{E}_{X_t,\alpha^0}\{R_{X_t}(\nu_t;\alpha^0,\alpha,\rho)\})^{(B)}<0$, α≠α^0;

(ii) $\overline{\lim_{n\to\infty}}(\frac{1}{n}\sum_{t=1}^{n}\mathbf{E}_{X_t,\alpha^0}\{V_{X_t}(\nu_t;\alpha^0,r)\})^{(B)}<0$.

С5. $R_{X_t}(\nu_t;\alpha^0,\alpha,\rho)$, $V_{X_t}(\nu_t;\alpha^0,r)$ *– борелевские функции от* ν_t.

Тогда ОМП $\hat{\alpha}^n$*, определяемая (8), состоятельна по вероятности:*

$$\hat{\alpha}^n\xrightarrow[n\to\infty]{\mathbf{P}}\alpha^0.$$

В [2] предложены более жесткие условия, при которых оценка максимального правдоподобия $\hat{\alpha}^n$ является сильно состоятельной оценкой для вектора параметров α^0. В частности, вместо условия С3(i) требуется, чтобы $\mathbf{E}_{X_t,\alpha^0}\{e^{sR_{X_t}^{(0)}(\nu_t;\alpha^0,\alpha,\rho)}\}\le M$ для некоторых s>0, ρ<ε(α); условия С3(ii) и С4(ii) заменяются на условие: $\frac{P_X(\nu;\alpha)}{P_X(\nu;\alpha^0)}\xrightarrow[|\alpha|\to\infty]{\mathbf{P}=1}0$.

Предложим ещё один вариант условий сильной состоятельности ОМП, который представим в виде следующей теоремы.

Теорема 2. *Пусть для модели (1), (3), (4) выполнены условие* С1, *а также следующие условия:*

У1. *Существуют* ρ>0 *и* r>0 *такие, что для некоторого* M>0 *сразу для всех* t=1,…,n *выполняются неравенства:*

(i) $\mathbf{D}_{X_t,\alpha^0}\{R_{X_t}(\nu_t;\alpha^0,\alpha,\rho)\}\le M$;

(ii) $\mathbf{D}_{X_t,\alpha^0}\{V_{X_t}(\nu_t;\alpha^0,r)\}\le M$.

У2. *Для* ρ>0 *и* r>0*, определённых в* У1, *справедливы соотношения:*

(i) $\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{t=1}^{n}\mathbf{E}_{X_t,\alpha^0}\{R_{X_t}(\nu_t;\alpha^0,\alpha,\rho)\}<0$, α≠α^0;

(ii) $\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{t=1}^{n}\mathbf{E}_{X_t,\alpha^0}\{V_{X_t}(\nu_t;\alpha^0,r)\}<0$.

Тогда ОМП $\hat{\alpha}^n$ *сильно состоятельна:*

$$\hat{\alpha}^n\xrightarrow[n\to\infty]{\mathbf{P}=1}\alpha^0. \qquad (9)$$

Доказательство. Рассмотрим произвольное ε>0. Обозначим W^ε={α∈W: |α-α^0|≥ε}.

Пусть r и ρ определяются из условия У1, W_1={α∈W^ε: |α|≤r}. С учетом С1 множество W_1 замкнуто и ограничено. Следовательно, существует конечное число точек α^1,…,α^h∈W_1 таких, что $W_1\subseteq\bigcup_{j=1}^{h}S(\alpha^j,\rho)$, где S(α,ρ) – шар радиуса ρ с центром в точке α.

Пусть α^{h+1}∈W^ε\\W_1. Определим T_X(α^j), j=1,…,h+1, следующим образом:

$$T_X(\alpha^j)=\mathbf{E}_{X,\alpha^0}\left\{R_X(\nu;\alpha^0,\alpha,\rho)\right\},\ j=1,\dots,h;\ T_X(\alpha^{h+1})=\mathbf{E}_{X,\alpha^0}\{V_X(\nu;\alpha^0,r)\}.$$

С учетом У1 справедливы следующие оценки:

$$\sum_{t=1}^{\infty}\frac{\mathbf{D}_{X_t,\alpha^0}\{R_{X_t}(\nu_t;\alpha^0,\alpha,\rho)\}}{t^2}\le\sum_{t=1}^{\infty}\frac{M}{t^2}<\infty,\ j=1,\dots,h;\ \sum_{t=1}^{\infty}\frac{\mathbf{D}_{X_t,\alpha^0}\{V_{X_t}(\nu_t;\alpha^0,r)\}}{t^2}\le\sum_{t=1}^{\infty}\frac{M}{t^2}<\infty.$$

Следовательно, применима первая теорема Колмогорова [9]:

$$\frac{1}{n}\sum_{t=1}^{n}(R_{X_t}(\nu_t;\alpha^0,\alpha,\rho)-T_{X_t}(\alpha^j))\xrightarrow[n\to\infty]{\mathbf{P}=1}0,\ j=1,\dots,h, \qquad (10)$$

$$\frac{1}{n}\sum_{t=1}^{n}(V_{X_t}(\nu_t;\alpha^0,r)-T_{X_t}(\alpha^{h+1}))\xrightarrow[n\to\infty]{\mathbf{P}=1}0. \qquad (11)$$

Условие У2 утверждает, что

$$\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{t=1}^{n}T_{X_t}(\alpha^j)=T(\alpha^j)<0,\ j=1,\ldots,h+1. \qquad (12)$$

Из (10), (11) и (12) следует, что с вероятностью 1 имеют место следующие соотношения:

$$\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{t=1}^{n}R_{X_t}(\nu_t;\alpha^0,\alpha,\rho)<0,\ j=1,\ldots,h;\ \ \overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{t=1}^{n}V_{X_t}(\nu_t;\alpha^0,r)<0.$$

Значит,

$$\mathbf{P}_{\{X_t\}_{t=1}^n,\alpha^0}\left\{\overline{\lim_{n\to\infty}}\prod_{t=1}^{n}\frac{P_{X_t}(\nu_t;\alpha^j,\rho)}{P_{X_t}(\nu_t;\alpha^0)}=0\right\}=1,\ j=1,\ldots,h;\ \mathbf{P}_{\{X_t\}_{t=1}^n,\alpha^0}\left\{\overline{\lim_{n\to\infty}}\prod_{t=1}^{n}\frac{\psi_{X_t}(\nu_t;r)}{P_{X_t}(\nu_t;\alpha^0)}=0\right\}=1.$$

Очевидно, что

$$0\le\sup_{\alpha\in W^\varepsilon}\prod_{t=1}^{n}P_{X_t}(\nu_t;\alpha)\le\sum_{j=1}^{h}\prod_{t=1}^{n}P_{X_t}(\nu_t;\alpha^j,\rho)+\prod_{t=1}^{n}\psi_{X_t}(\nu_t;r).$$

Тогда

$$0\le\overline{\lim_{n\to\infty}}\frac{\sup\limits_{\alpha\in W^\varepsilon}\prod_{t=1}^{n}P_{X_t}(\nu_t;\alpha)}{\prod_{t=1}^{n}P_{X_t}(\nu_t;\alpha^0)}\le\sum_{j=1}^{h}\overline{\lim_{n\to\infty}}\prod_{t=1}^{n}\frac{P_{X_t}(\nu_t;\alpha^j,\rho)}{\prod_{t=1}^{n}P_{X_t}(\nu_t;\alpha^0)}+\overline{\lim_{n\to\infty}}\prod_{t=1}^{n}\frac{\psi_{X_t}(\nu_t;r)}{\prod_{t=1}^{n}P_{X_t}(\nu_t;\alpha^0)}.$$

Следовательно,

$$\mathbf{P}_{\{X_t\}_{t=1}^n,\alpha^0}\left\{\overline{\lim_{n\to\infty}}\frac{\sup\limits_{\alpha\in W^\varepsilon}\prod_{t=1}^{n}P_{X_t}(\nu_t;\alpha)}{\prod_{t=1}^{n}P_{X_t}(\nu_t;\alpha^0)}=0\right\}=1. \qquad (13)$$

Заметим, что для любого n справедлива следующая вложенность событий

$$\{|\hat\alpha^n-\alpha^0|\ge\varepsilon\}\subseteq\left\{\sup_{\alpha\in W^\varepsilon}\prod_{t=1}^{n}P_{X_t}(\nu_t;\alpha)=\prod_{t=1}^{n}P_{X_t}(\nu_t;\hat\alpha^n)\right\}. \qquad (14)$$

В силу определения ОМП (8) имеем: $\prod_{t=1}^{n}P_{X_t}(\nu_t;\hat\alpha^n)\ge\prod_{t=1}^{n}P_{X_t}(\nu_t;\alpha^0)$. Тогда на основании (13), (14) для любого $\varepsilon>0$ существует n_0 такое, что для $n\ge n_0$ выполнено $\mathbf{P}_{\{X_t\}_{t=1}^n,\alpha^0}\{|\hat\alpha^n-\alpha^0|<\varepsilon\}=1$, что и доказывает сильную состоятельность ОМП. □

Заметим, что если W – замкнутое ограниченное подмножество $\mathbb{R}^{m+1}$, то условия У1(ii) и У2(ii) опускаются. В доказательстве можно в качестве W_1 взять множество W, положив $r=\max_{\alpha\in W}|\alpha|$, и опустить все рассуждения, связаные с $|\alpha|>r$.

Сформулируем далее ряд достаточных условий сильной состоятельности ОМП, проще проверяемых при решении прикладных задач.

Достаточные условия сильной состоятельности ОМП $\hat{\boldsymbol\alpha}^{\mathbf{n}}=(\hat{\boldsymbol\theta}^{\mathbf{n}},(\hat{\boldsymbol\sigma}^{\mathbf{n}})^2)'$. Приведем вспомогательные утверждения.

Лемма 2. *Пусть для любого фиксированного значения* $\theta\in\Theta$ *функция* $F(X;\theta)$ *ограничена на* $\mathbf{X}\subseteq\mathbb{R}^N$. *Тогда если* $K<+\infty$, *то существует константа* $c(\alpha)>0$ *такая, что* $P_X(k;\alpha)\ge c(\alpha)$ *сразу для всех* $X\in\mathbf{X}\subseteq\mathbb{R}^N$, $k\in\mathbf{K}$.

Доказательство. По условию данной леммы существуют ограниченные функции $-\infty<M_1(\theta)\le M_2(\theta)<+\infty$ такие, что для любого $X\in\mathbf{X}\subseteq\mathbb{R}^N$

$$M_1(\theta)\le F(X;\theta)\le M_2(\theta).$$

Тогда

$$(F(X;\theta)-z)^2\le\max_{i=1,2}(M_i(\theta)-z)^2=M(z;\theta),\ z\in\mathbb{R}.$$

В силу (4), (5) для любого $X\in\mathbf{X}\subseteq\mathbb{R}^N$

$$P_X(k;\alpha)=\frac{1}{\sqrt{2\pi}\sigma}\int_{a_{k-1}}^{a_k}e^{-\frac{(z-F(X,\theta))^2}{2\sigma^2}}dz\ge\frac{1}{\sqrt{2\pi}\sigma}\int_{a_{k-1}}^{a_k}e^{-\frac{M(z,\theta)}{2\sigma^2}}dz=c_k(\alpha)>0.$$

В силу конечности K выберем $c(\alpha)=\min_{k\in\mathbf{K}}c_k(\alpha)$. □

Лемма 3. *Для* $\alpha\neq\alpha^0$ *справедливо неравенство:*

$$\mathbf{E}_{X,\alpha^0}\{\ln P_X(\nu;\alpha)\}\le \mathbf{E}_{X,\alpha^0}\{\ln P_X(\nu;\alpha^0)\},\ X\in\mathbf{X}\subseteq\mathbb{R}^N,$$

причем

$$\mathbf{E}_{X,\alpha^0}\{\ln P_X(\nu;\alpha)\}=\mathbf{E}_{X,\alpha^0}\{\ln P_X(\nu;\alpha^0)\}\Leftrightarrow \sum_{k=1}^{K}|P_X(k;\alpha)-P_X(k;\alpha^0)|=0.$$

Доказательство. Воспользуемся неравенством Йенсена [9] для выпуклой вверх функции $y=\ln x$ и условием нормировки:

$$\mathbf{E}_{X,\alpha^0}\left\{\ln\frac{P_X(\nu;\alpha)}{P_X(\nu;\alpha^0)}\right\}\le \ln\mathbf{E}_{X,\alpha^0}\left\{\frac{P_X(\nu;\alpha)}{P_X(\nu;\alpha^0)}\right\}=\ln\sum_{k=1}^{K}\frac{P_X(k;\alpha)}{P_X(k;\alpha^0)}P_X(k;\alpha^0)=\ln\sum_{k=1}^{K}P_X(k;\alpha)=\ln 1=0.$$

В силу неравенства Йенсена

$$\mathbf{E}_{X,\alpha^0}\{\ln P_X(\nu;\alpha)\}=\mathbf{E}_{X,\alpha^0}\{\ln P_X(\nu;\alpha^0)\}\Leftrightarrow \sum_{k=1}^{K}|P_X(k;\alpha)-P_X(k;\alpha^0)|=0.$$

□

Лемма 4. *Пусть* $K>2$. *Тогда для любого* $\delta>0$ *существует* $d<0$ *такое, что для любого* $\alpha\neq\alpha^0$ *если в точке* $X\in\mathbf{X}\subseteq\mathbb{R}^N$ *выполнено* $|F(X;\theta^0)-F(X;\theta)|\ge\delta$ *или* $|\sigma^2-(\sigma^0)^2|\ge\delta$, *то* $E_{X,\alpha^0}\left\{\ln\frac{P_X(\nu;\alpha)}{P_X(\nu;\alpha^0)}\right\}\le d$.

Доказательство. Проведем доказательство от противного. Пусть существуют $\delta>0$, $\alpha\neq\alpha^0$ и $X\in\mathbf{X}$ такие, что из $|F(X;\theta^0)-F(X;\theta)|\ge\delta$ или $|\sigma^2-(\sigma^0)^2|\ge\delta$ следует, что $E_{X,\alpha^0}\left\{\ln\frac{P_X(\nu;\alpha)}{P_X(\nu;\alpha^0)}\right\}=0$. В силу леммы 3 равенство $E_{X,\alpha^0}\left\{\ln\frac{P_X(\nu;\alpha)}{P_X(\nu;\alpha^0)}\right\}=0$ равносильно тому, что $P_X(k;\alpha)=P_X(k;\alpha^0)$, $k\in\mathbf{K}$. Из (5) и свойств функции Лапласа получаем $(\sigma-\sigma^0)a_{K-1}=(\sigma-\sigma^0)a_1=\sigma F(X;\theta^0)-\sigma^0F(X;\theta)$. Если $\sigma-\sigma^0\neq0$, то $a_{K-1}=a_1$. Это противоречит тому, что $K>2$. Если же $\sigma-\sigma^0=0$, то возникает противоречие с предположением о том, что $|F(X;\theta^0)-F(X;\theta)|\ge\delta$. □

Теорема 3. *Пусть для модели (1), (3), (4) выполнены следующие условия:*

У1*. *Число классов ограничено:* $2<K<+\infty$.

У2*. Θ *– замкнутое подмножество* $\mathbb{R}^m$; *известно такое* $\bar\sigma^2>0$, *что* $\bar\sigma^2\le(\sigma^0)^2$.

У3*. *Множество возможных значений регрессора* $\mathbf{X}\subseteq\mathbb{R}^N$ *– компакт.*

У4*. *Функция* $F(X;\theta)$ *непрерывна на* $\mathbf{X}\times\Theta$.

У5*. *Для любого фиксированного значения* $\theta\in\Theta$ *функция* $F(X;\theta)$ *ограничена на* $\mathbf{X}\subseteq\mathbb{R}^N$.

У6*. *Для любого* $\varepsilon>0$ *существует* $\delta=\delta(\varepsilon)>0$ *такое, что для любого* $\alpha\in W^\varepsilon$, *где* $W^\varepsilon=\{\alpha\in\Theta\times[\bar\sigma^2,\infty): |\alpha-\alpha^0|\ge\varepsilon\}$, *существует нижний предел*

$$\underline{\lim}_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\mathbf{I}\{|F(X_t;\theta^0)-F(X_t;\theta)|\ge\delta\}=b,$$

где $0<b<b(\theta,\theta^0,\delta,F(\cdot))\le1$, $\mathbf{I}\{A\}\in\{0,1\}$ *– индикатор истинности события A.*

У7*. *Для любого* $R>0$ *существует* $r>0$ *такое, что*

$$\underline{\lim}_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\mathbf{I}\left\{\inf_{|\theta|\ge r}|F(X_t;\theta)|\ge R\right\}=q,\ 0<q<q(R,F(\cdot))\le1.$$

Тогда ОМП $\hat\alpha^n$ *сильно состоятельна, т.е. выполняется (9).*

Доказательство. Выберем произвольное $\varepsilon>0$. Существует $\delta_1=2\delta>0$, для которого выполняется условие У6*, т.е.

$$\underline{\lim}_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\mathbf{I}\{|F(X_t;\theta^0)-F(X_t;\theta)|\ge2\delta\}=b>0.$$

Из условия У7* следует, что для $R=\max_{\mathbf{X}}F(X;\theta^0)+2\varepsilon$ существует $r>0$ такое, что $\underline{\lim}_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\mathbf{I}\left\{\inf_{|\theta|\ge r}|F(X_t;\theta)|\ge R\right\}=q>0$. Пусть $r_0=\sqrt{2}\max\{r,(\sigma^0)^2+2\varepsilon\}$ и $W_1=\{\alpha\in W^\varepsilon: |\alpha|\le r_0\}$.

В силу условий У3* и У4* функция $F(X,\theta)$ равномерно непрерывна на $\mathbf{X}\times\{\theta: |\theta|\le r_0\}$. Тогда для δ существует $\rho>0$ такое, что если $|X_1-X_2|\le\rho$ и $|\theta_1-\theta_2|\le\rho$, то $|F(X_1;\theta_1)-F(X_2;\theta_2)|<\delta$. С учетом У2* множество W_1 замкнуто и ограничено. Следовательно, существует конечное число точек $\alpha^1,\dots,\alpha^h \in W_1$ таких, что $W_1 \subseteq \bigcup_{j=1}^{h} S(\alpha^j,\rho)$.

Пусть $\alpha^{h+1} \in W^\varepsilon \backslash W_1$. Определим $T_X(\alpha^j)$, $j=1,\dots,h+1$, следующим образом:

$$T_X(\alpha^j) = \mathbf{E}_{X,\alpha^0}\{\ln P_X(\nu;\alpha^j,\rho)\} - \mathbf{E}_{X,\alpha^0}\{\ln P_X(\nu;\alpha^0)\},\ j=1,\dots,h;$$

$$T_X(\alpha^{h+1}) = \mathbf{E}_{X,\alpha^0}\{\ln \psi_X(\nu;r_0)\} - \mathbf{E}_{X,\alpha^0}\{\ln P_X(\nu;\alpha^0)\}.$$

Из условия У2* следует выполнение условия С1.

На основании леммы 2 с учетом У5* для любого α существует $c(\alpha)>0$ такое, что $P_X(\nu;\alpha) \ge c(\alpha)$ для любого $X \in \mathbf{X}$. Тогда

$$c(\alpha^j) \le \frac{P_X(\nu;\alpha^j)}{P_X(\nu;\alpha^0)} \le \frac{P_X(\nu;\alpha^j,\rho)}{P_X(\nu;\alpha^0)} \le \frac{1}{c(\alpha^0)},\ j=1,\dots,h;\ c(\alpha^{h+1}) \le \frac{P_X(\nu;\alpha^{h+1})}{P_X(\nu;\alpha^0)} \le \frac{\psi_X(\nu;r_0)}{P_X(\nu;\alpha^0)} \le \frac{1}{c(\alpha^0)}.$$

Тогда справедливы следующие оценки:

$$\mathbf{D}_{X_t,\alpha^0}\left\{\ln \frac{P_{X_t}(\nu_t;\alpha^j,\rho)}{P_{X_t}(\nu_t;\alpha^0)}\right\} \le \max\{(\ln c(\alpha^0))^2, (\ln c(\alpha^j))^2\} < \infty,\ j=1,\dots,h,$$

$$\mathbf{D}_{X_t,\alpha^0}\left\{\ln \frac{\psi_{X_t}(\nu_t;r_0)}{P_{X_t}(\nu_t;\alpha^0)}\right\} \le \max\{(\ln c(\alpha^0))^2, (\ln c(\alpha^{h+1}))^2\} < \infty.$$

Следовательно, выполнено условие У1.

Докажем, что выполняется условие У2, т.е.

$$\overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{t=1}^{n} T_{X_t}(\alpha^j) = T(\alpha^j) < 0,\ j=1,\dots,h+1.$$

По лемме 4 с учетом У1* для δ существует такое $d_\delta<0$, что если для $\alpha \ne \alpha^0$ в точке $X_t \in \mathbf{X}$ выполнено $|F(X_t;\theta^0)-F(X_t;\theta)| \ge \delta$, то $\mathbf{E}_{X_t,\alpha^0}\left\{\ln \frac{P_{X_t}(\nu_t;\alpha)}{P_{X_t}(\nu_t;\alpha^0)}\right\} \le d_\delta$. Если выполнено $|F(X_t;\theta^0)-F(X_t;\theta^j)| \ge 2\delta$, то для любого θ такого, что $|\theta-\theta^j| \le \rho$, выполнено $|F(X_t;\theta^0)-F(X_t;\theta)| \ge \delta$. Получаем, что тогда для любого $\alpha \in S(\alpha^j,\rho)$ выполнено $\mathbf{E}_{X_t,\alpha^0}\left\{\ln \frac{P_{X_t}(\nu_t;\alpha)}{P_{X_t}(\nu_t;\alpha^0)}\right\} \le d_\delta$. Значит, если выполнено $|F(X_t;\theta^0)-F(X_t;\theta^j)| \ge 2\delta$, то $T_{X_t}(\alpha^j) \le d_\delta$. Следовательно,

$$\underline{\lim_{n\to\infty}} \frac{1}{n}\sum_{t=1}^{n} \mathbf{I}\{T_{X_t}(\alpha^j) \le d_\delta\} \ge \underline{\lim_{n\to\infty}} \frac{1}{n}\sum_{t=1}^{n} \mathbf{I}\{|F(X_t;\theta^0) - F(X_t;\theta^j)| \ge 2\delta\} = b^j > 0,\ j=1,\dots,h.$$

Тогда

$$\overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{t=1}^{n} T_{X_t}(\alpha^j) \le d_\delta \underline{\lim_{n\to\infty}} \frac{1}{n}\sum_{t=1}^{n} \mathbf{I}\{T_{X_t}(\alpha^j) \le d_\delta\} \le d_\delta b^j < 0,\ j=1,\dots,h.$$

Пусть в точке $X_t \in \mathbf{X}$ справедливо неравенство: $\inf_{|\theta|\ge r} |F(X_t;\theta)| \ge R$. Тогда при любом θ, $|\theta| \ge r$, имеем: $|F(X_t;\theta^0)-F(X_t;\theta)| \ge 2\varepsilon$. На $W^\varepsilon \backslash W_1$ или $|\theta| \ge r$, или $|\sigma^2-(\sigma^0)^2| \ge 2\varepsilon$. Значит для любого $\alpha \in W^\varepsilon \backslash W_1$ или $|F(X_t;\theta^0)-F(X_t;\theta)| \ge 2\varepsilon$, или $|\sigma^2-(\sigma^0)^2| \ge 2\varepsilon$. Тогда по лемме 4 с учетом У1* существует $d_{2\varepsilon}<0$ такое, что $\mathbf{E}_{X_t,\alpha^0}\left\{\ln \frac{P_{X_t}(\nu_t;\alpha)}{P_{X_t}(\nu_t;\alpha^0)}\right\} \le d_{2\varepsilon}$. Т.к. это верно для любого $\alpha \in W^\varepsilon \backslash W_1$, то $T_{X_t}(\alpha^{h+1}) \le d_{2\varepsilon}$. Следовательно,

$$\underline{\lim_{n\to\infty}} \frac{1}{n}\sum_{t=1}^{n} \mathbf{I}\{T_{X_t}(\alpha^{h+1}) \le d_{2\varepsilon}\} \ge \underline{\lim_{n\to\infty}} \frac{1}{n}\sum_{t=1}^{n} \mathbf{I}\left\{\inf_{|\theta|\ge r} |F(X_t;\theta)| \ge R\right\} = q > 0.$$

Тогда

$$\overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{t=1}^{n} T_{X_t}(\alpha^{h+1}) \le d_{2\varepsilon} \underline{\lim_{n\to\infty}} \frac{1}{n}\sum_{t=1}^{n} \mathbf{I}\{T_{X_t}(\alpha^{h+1}) \le d_{2\varepsilon}\} \le d_{2\varepsilon} q < 0.$$

Таким образом, выполнены условия С1, У1 и У2. Следовательно, применима теорема 2, и ОМП $\hat{\alpha}^n$ сильно состоятельна. □

Заметим, что если W – замкнутое ограниченное подмножество $\mathbb{R}^{m+1}$, то условия У7* опускается. В доказательстве можно в качестве W_1 взять множество W, положив $r=\max_{\alpha\in W}|\alpha|$, и опустить все рассуждения, связаные с $|\alpha|>r$.

Поясним содержательный смысл условия У6*. Оно означает, что ε-уклонение параметра значимо для функции регрессии $F(\cdot)$ хотя бы на множестве значений регрессоров $\mathcal{X}_n(\delta)=\{X_t\in\{X_1,\ldots,X_n\}: |F(X_t;\theta)-F(X_t;\theta^0)|\geq\delta\}$, мощность которого растёт вместе с ростом n:

$$|\mathcal{X}_n(\delta)|/n \xrightarrow[n\to\infty]{} b,\ 0<b\leq 1.$$

Применим теорему 3 для случая множественной линейной регрессии (2), где $\Theta=Q^m$, $\mathbf{X}=T^N$, $T=[t_1,T_1]$, $Q=[q_1,Q_1]$, $t_1,T_1,q_1,Q_1\in\mathbb{R}$.

Теорема 4. *Пусть справедлива модель множественной линейной регрессии (2), где* $\Theta=Q^m$, $\mathbf{X}=T^N$, $T=[t_1,T_1]$, $Q=[q_1,Q_1]$, $t_1,T_1,q_1,Q_1\in\mathbb{R}$, *и выполнены следующие предположения:*

П1. *Число классов ограничено:* $2<K<+\infty$.

П2. *Известно такое* $\bar{\sigma}^2>0$, *что* $\bar{\sigma}^2\leq(\sigma^0)^2$.

П3. *Для любого* $\varepsilon>0$ *существует* $\delta=\delta(\varepsilon)>0$ *такое, что для любых* $\alpha\in W^\varepsilon$, *где* $W^\varepsilon=\{\alpha\in\Theta\times[\bar{\sigma}^2,\infty)$: $|\alpha-\alpha^0|\geq\varepsilon\}$, *выполнено*

$$\varliminf_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\mathbf{I}\{|(X_t'\vdots 1)(\theta^0-\theta)|\geq\delta\}=b,\ 0<b<b(\theta,\theta^0,\delta)\leq 1.$$

Тогда ОМП $\hat{\alpha}^n$ *сильно состоятельна, т.е. выполняется (9).*

Доказательство. Предположения П1, П3 совпадают с условиями У1*, У6*. Предположение П2 и вид множества Θ обеспечивают выполнение условия У2*. Заметим, что $W=\Theta\times[\bar{\sigma}^2,\infty)$ – замкнутое ограниченное множество, т.е. условие У7* можно опустить. Выполнение условий У3*, У4*, У5* обеспечивается видом функции множественной линейной регрессии и множества $\mathbf{X}$. □

Результаты компьютерного моделирования. Проиллюстрируем теоретические результаты компьютерными экспериментами. Для компьютерного моделирования в качестве функции регрессии использовалась производственная функция Кобба-Дугласа, широко применяемая в эконометрических приложениях [15]:

$$Y_t=\theta_1^0(X_t^1)^{\theta_2^0}(X_t^2)^{\theta_3^0}+\xi_t,\ t=1,\ldots,n,\ldots,$$

где $m=3$; $\theta_1^0=1$; $\theta_2^0=3$; $\theta_3^0=4$; $(\sigma^0)^2=9$; $K=4$; $A_1=(-\infty,10]$; $A_2=(10,40]$; $A_3=(40,60]$; $A_4=(60,\infty)$. Значения регрессоров $\{X_t^1,X_t^2\}_{t=1}^n$ представляют собой узлы равномерной сетки $[0,2]\times[0,2]$. Для нахождения ОМП использовался метод градиентного спуска решения экстремальной задачи (6), (8) [10]. По методу Монте-Карло для каждого значения объема выборки n проводилось $Q=1000$ экспериментов и вычислялась среднеквадратичная погрешность оценивания параметров θ^0 и $(\sigma^0)^2$:

$$V_{\theta^0}^n=\frac{1}{Q}\sum_{q=1}^{Q}\|\hat{\theta}^n-\theta^0\|^2,\ V_{\sigma^{0^2}}^n=\frac{1}{Q}\sum_{q=1}^{Q}(\hat{\sigma}^{n,q^2}-\sigma^{0^2})^2.$$

На рисунке 1 представлен график зависимости статистики $V_{\theta^0}^n$ от объема выборки n, иллюстрирующий состоятельность оценки $\hat{\theta}^n$. На рисунке 2 представлен график зависимости статистики $V_{\sigma^{0^2}}^n$ от объема выборки n, иллюстрирующий состоятельность оценки $(\hat{\sigma}^n)^2$.

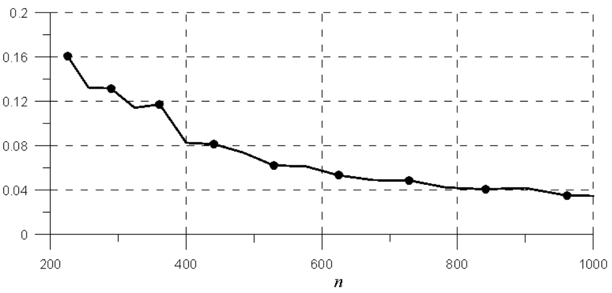

Рис. 1 График зависимости $V^n_{\theta^0}$ от *n*

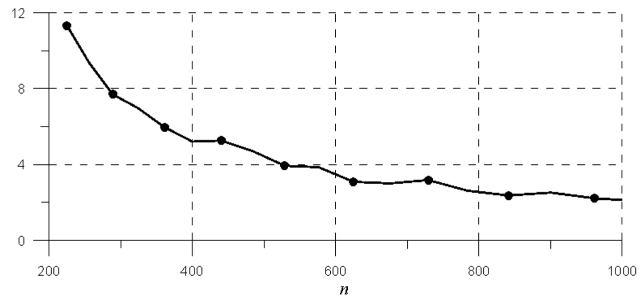

Рис. 2 График зависимости $V^n_{\sigma^{0^2}}$ от *n*

Заключение. В статье исследована модель множественной нелинейной регрессии (1), в которой зависимые данные наблюдаются не полностью: вместо их точных значений известны только номера классов (интервалов), в которые они попадают. Построены оценки максимального правдоподобия для параметров модели. Найдены условия состоятельности по вероятности и сильной состоятельности ОМП параметров функции регрессии при наличии классификации наблюдений. Теоретические результаты согласуются с результатами компьютерного моделирования.

Литература

1. Bai, Z., Zheng, S., Zhang, B., Hu, Z. // J. Statist. Plann. Inferense. 2009. 139, no. 8. P. 2526–2542.
2. Chao, M. T. // Dr. Y. W. Chen's 60-year Memorial Volume. Academia Sinica, Taipei, 1970.
3. Dempster, A.P. and Rubin, D.B. // J. Roy. Statist. Soc. 1983. Ser. B., 45. p 51–59.
4. Hoadley B. // Ann. Math. Statist. 1971. Vol. 42, no. 4. p 1977–1991.
5. Kharin, Yu. // Communications in Statistics – Theory and Methods. 2011. 40, no 16. p 2893–2906.
6. Nelson, W., Hahn, G.J. // Technometrics. 1972. Vol. 14. p 247–269.
7. Sen Roy, S., Guriab, S. // Statistics. 2009. 43, no. 6. p 531–539.
8. Sheppard, W. F. // Proc. London Math. Soc. 2009. 29. p 353–380.
9. Боровков, А.А. Теория вероятностей. М.: Наука, 1986.–432 с.
10. Калитин Н.Н. Численные методы. М.: Наука, 1978.–512 с.
11. Литтл, Р. Дж. А., Рубин Д.Б. Статистический анализ данных с пропусками. М.: Финансы и статистика, 1990.–336 с.
12. Харин, Ю.С. Оптимальность и робастность в статистическом прогнозировании. Мн.: БГУ, 2008.–263 с.
13. Харин, Ю.С., Жук Е.Е. Математическая и прикладная статистика. Мн.:БГУ, 2005.–279 с.
14. Хьюбер, Дж. П. Робастность в статистике. М.: Мир, 1984.–304 с.
15. Greene, W. H. Econometric analysis. N.Y.: MacMillan, 2000.–720 p.