Теорія ймовір. та матем. статист. Вип. 86, 2012, стор. 69–80
Teor. Ĭmovir. ta Matem. Statyst. No. 86, 2012, pp. 69–80

ГРАНИЧНІ ТЕОРЕМИ ДЛЯ ЕКСТРЕМАЛЬНИХ ЗАЛИШКІВ У ЛІНІЙНІЙ ТА НЕЛІНІЙНІЙ МОДЕЛЯХ РЕГРЕСІЇ

УДК 519.21

О. В. ІВАНОВ І І. К. МАЦАК

Анотація. У роботі отримано граничні теореми для максимальних залишків у моделях лінійної та нелінійної регресії. Наведено приклад застосування такого типу результатів до побудови критерію адекватності моделі регресії.

Abstract. In the paper the limit theorems for maximal residuals in linear and nonlinear regression models are obtained. An example of application of the results of such a type to construction of regression model adequacy test is given.

Аннотация. В работе получены теоремы для максимальных невязок в моделях линейной и нелинейной регресии. Приведен пример применения тпкого типа результатов к построению критерия адекватности модели регресии.

Класичні теореми регресійного аналізу описують різноманітні властивості залишкової суми квадратів, тобто суми квадратів відхилень спостережень від функції регресії, в яку замість невідомого параметра підставляється оцінка найменших квадратів (о.н.к.): див., наприклад, [1, 2].

Натомість в даній роботі вивчається асимптотична поведінка максимуму таких відхилень, а саме: доведено збіжність розподілу цього нормованого певним чином максимуму до граничного (при зростанні об'єму вибірки до нескінченності) розподілу нормованого максимуму однаково розподілених помилок спостережень. Розглянуто лінійну та нелінійну моделі регресії.

Отримані результати можна застосувати для побудови критеріїв адекватності моделей регресії(див. розділ 3).

1. Лінійна модель регресії

Розглянемо модель лінійної регресії

$$y_j = \sum_{i=1}^{q} \theta_i x_{ji} + \varepsilon_j, \qquad j = 1, \dots, n, \tag{1}$$

де ε_j — незалежні однаково розподілені (н.о.р.) випадкові величини (в.в.), $\mathsf{E}\,\varepsilon_j = 0$, $\mathsf{E}\,\varepsilon_j^2 = \sigma^2 < \infty$.

Нехай

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1q} \\ x_{21} & x_{22} & \dots & x_{2q} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nq} \end{pmatrix}$$

2010 *Mathematics Subject Classification.* Primary 60G70, 62J05.
Ключові слова і фрази. Модель регресії, екстремальні значення, критерії адекватності.

— $n \times q$-матриця плану регресійного експеременту, $\det(X^T X) \neq 0$. Тоді оцінка найменших квадратів (о.н.к.) невідомих параметрів $\theta^T = (\theta_1, \dots, \theta_q) \in \mathbf{R}^q$ за спостереженнями (1), тобто випадковий вектор $\hat{\theta}_n^T = (\hat{\theta}_{1n}, \dots, \hat{\theta}_{qn})$, що мінімізує функціонал

$$Q(\theta) = (Y - X\theta)^T (Y - X\theta) = \sum_{j=1}^{n} \left(y_j - \sum_{i=1}^{q} \theta_i x_{ji} \right)^2, \tag{2}$$

$Y^T = (y_1, \dots, y_n)$, задається формулою

$$\hat{\theta}_n = \left(X^T X \right)^{-1} X^T Y. \tag{3}$$

Покладемо

$$\begin{gathered} \hat{y_j} = \sum_{i=1}^{q} \hat{\theta}_{in} x_{ji}, \qquad \hat{\varepsilon}_j = y_j - \hat{y_j}, \qquad j = 1, \dots, n, \\ Z_n = \max_{1 \le j \le n} \varepsilon_j, \qquad \hat{Z}_n = \max_{1 \le j \le n} \hat{\varepsilon}_j, \\ d_{in}^2 = \sum_{j=1}^{n} x_{ji}^2, \qquad d_n = \operatorname{diag}(d_{in}, i = 1, \dots, q). \end{gathered} \tag{4}$$

Наступні дві умови є стандартними в регресійному аналізі при вивченні асимптотичних властивостей о.н.к.

(i) $$d_{in}^{-1} \max_{1 \le j \le n} |x_{ji}| \le k_i n^{-1/2}, \qquad i = 1, \dots, q. \tag{5}$$

Позначимо

$$J_n = d_n^{-1} X^T X d_n^{-1} = \left(d_{kn}^{-1} d_{ln}^{-1} \sum_{j=1}^{n} x_{jk} x_{jl} \right)_{k,l=1}^{q}, \tag{6}$$

$\lambda_{\min}(A)$ — найменше власне число додатно визначеної матриці A.

(ii) $$\lambda_{\min}(J_n) \ge \lambda_0 > 0 \quad \text{для } n > n_0. \tag{7}$$

Далі у роботі будуть використовуватись деякі результати класичної теорії екстремальних значень в.в. Нагадаємо відповідні твердження.

Асимптотична теорія екстремальних значень має справу з граничним розподілом Z_n при $n \to \infty$. Виявляється, що асимптотичний розподіл величини Z_n (нормованої відповідним чином) обов'язково повинен належати одній із трьох єдино можливих сімей розподілів. Зауважимо, що в загальній теорії екстремальних значень умови типу $\mathsf{E}\,\varepsilon = 0$, $\mathsf{E}\,\varepsilon^2 = \sigma^2 < \infty$ звичайно не накладаються.

Нехай в.в. (ε_j) мають функцію розподілу (ф.р.) $F(x)$. Припустимо, що для деяких констант $b_n > 0$, a_n при $n \to \infty$

$$b_n(Z_n - a_n) \xrightarrow{\mathbf{D}} \zeta, \tag{8}$$

і ζ має невироджену ф.р. $G(x) = \mathsf{P}(\zeta < x)$.

Якщо співвідношення (8) виконується, то кажуть, що ф.р. F належить області максимум притягнення закону G і пишуть $F \in D(G)$.

Теорема А (Б. В. Гнєденко [3]). *Якщо ф.р. F належить області максимум притягнення закону G, то G може мати лише один із наступних трьох типів розподілів:*

$$\begin{aligned} \textit{Тип I}: \quad \Phi_\alpha(x) &= \begin{cases} 0, & \textit{при} \quad x \le 0, \\ \exp(-x^{-\alpha}), & \textit{при}\ \alpha > 0,\ x > 0; \end{cases} \\ \textit{Тип II}: \quad \Psi_\alpha(x) &= \begin{cases} \exp(-(-x)^{\alpha}), & \textit{при}\ \alpha > 0,\ x \le 0, \\ 1, & \textit{при} \quad x > 0; \end{cases} \\ \textit{Тип III}: \quad \Lambda(x) &= \exp(-e^{-x}), \quad \textit{при}\ -\infty < x < \infty. \end{aligned} \tag{9}$$

І навпаки, будь-яка функція G, яка задається одною із рівностей (9), може бути граничною у співвідношенні (8), це буде так, наприклад, коли $F(x) = G(x)$.

Позначимо $x_F = \sup\{x\colon F(x) < 1\}$. Сформулюємо прості достатні умови приналежності ф.р. $F(x)$ областям максимум притягнення $D(\Phi_\alpha)$, $D(\Psi_\alpha)$ та $D(\Lambda)$ (див.[3]–[6]).

Теорема Б (Б. В. Гнєденко, Р. Мізес). *Для того, щоб F належала області максимум притягнення одного із трьох типів екстремальних розподілів, достатньо, щоб виконувались умови:*

$\textit{Тип I}\ (\Phi_\alpha)$:

$$x_F = \infty \quad \textit{i} \quad \forall x > 0 \quad \lim_{t\to\infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\alpha}, \quad \alpha > 0; \tag{10}$$

$\textit{Тип II}\ (\Psi_\alpha)$:

$$x_F < \infty \quad \textit{i} \quad \forall x > 0 \quad \lim_{h\downarrow 0} \frac{1 - F(x_F - xh)}{1 - F(x_F - h)} = x^{\alpha}, \quad \alpha > 0; \tag{11}$$

$\textit{Тип III}\ (\Lambda)$: $\exists F'(x) = f(x)$, $f(x)$ *має від'ємну похідну* $f^{'}(x)$ *в інтервалі* (x_0, x_F), $(x_F \le \infty)$, $f(x) = 0$ *для* $x > x_F$

$$\textit{i} \lim_{x\uparrow x_F} \frac{f^{'}(x)(1 - F(x)}{f^2(x)} = -1. \tag{12}$$

Достатні умови (10), (11) теореми **Б** будуть і необхідними, а умова (12) близька до необхідної.

Константи a_n, b_n в співвідношенні (8) для кожного із трьох типів можуть бути вибрані так:

$$\text{Тип I} \quad (\Phi_\alpha): \quad b_n = (\gamma_n)^{-1}, \quad a_n = 0; \tag{13}$$

$$\text{Тип II} \quad (\Psi_\alpha): \quad b_n = (x_F - \gamma_n)^{-1}, a_n = x_F; \tag{14}$$

При умові (12)

$$\text{Тип III} \quad (\Lambda): \quad a_n = \gamma_n, \quad b_n = r(\gamma_n), \tag{15}$$

де $\gamma_n = F^{-1}(1 - n^{-1}) = \inf(x\colon F(x) \ge 1 - n^{-1})$, $F(x) = 1 - \exp(-R(x))$, $R'(x) = r(x)$ (див. [5, с. 40]).

Приклад 1. Якщо (ε_j) — н.о.р. $N(0,1)$ в.в., то виконується умова (12), і при виборі констант

$$b_n = (2\ln n)^{1/2}, \qquad a_n = (2\ln n)^{1/2} - \frac{\ln\ln n + \ln(4\pi)}{2(2\ln n)^{1/2}} \qquad \text{при } n > 1, \tag{16}$$

справедливе співвідношення ([4]–[6])

$$\lim_{n\to\infty} \mathsf{P}\{b_n(Z_n - a_n) < x\} = \Lambda(x).$$

Теорема 1. *Нехай ф.р. F в.в. ε_j в моделі* (1) *належить області максимум притягнення закону G, виконуються умови (i), (ii) та*

$$b_n n^{-1/2} \to 0, \qquad n\to\infty. \tag{17}$$

Тоді

$$\lim_{n\to\infty} \mathsf{P}\{b_n(\hat{Z}_n - a_n) < x\} = G(x), \qquad x\in\mathbf{R}^1, \tag{18}$$

де G — один із трьох типів розподілів Φ_α, Ψ_α, Λ.

Доведення теореми 1. Із елементарної нерівності

$$\left|\max_{1\le j\le n} c_j - \max_{1\le j\le n} d_j\right| \le \max_{1\le j\le n} |c_j - d_j|$$

одержуємо

$$|\hat{Z}_n - Z_n| \le \max_{1\le j\le n} |\hat{\varepsilon}_j - \varepsilon_j| \le \sum_{i=1}^{q} |\hat{\theta}_{in} - \theta_i| \max_{1\le j\le n} |x_{ji}|,$$

а звідси

$$\Delta_n = \mathsf{E}\,|\hat{Z}_n - Z_n| \le \sum_{i=1}^{q} (\mathbf{D}\,\hat{\theta}_{in})^{1/2} \max_{1\le j\le n} |x_{ji}|. \tag{19}$$

З умови (ii) випливає існування матриці $J_n^{-1} = \mathcal{L}_n = (\mathcal{L}_n^{kl})_{k,l=1}^q$ та нерівності

$$\det J_n = \lambda_1 \dots \lambda_q \ge \lambda_{\min}^q \ge \lambda_0^q > 0,$$
$$(\det J_n)^{-1} \le \lambda_0^{-q}, \qquad n > n_0,$$

де $\lambda_1, \dots, \lambda_q$ — власні числа матриці J_n.

В матриці алгебраїчних доповнень до елементів матриці J_n кожний елемент складається з добутків елементів матриці J_n. Кожний елемент J_n за нерівністю Коші–Буняковського оцінюється одиницею. У відповідному мінорі всього $(q-1)!$ доданків. Таким чином, за умови (ii) кожний елемент матриці $\mathcal{L}_n$ оцінюється величиною

$$|\mathcal{L}_n^{kl}| \le \lambda_0^{-q}(q-1)!, \qquad k,l = 1,\dots,q. \tag{20}$$

Коваріаційна матриця нормованої о.н.к. $d_n(\hat{\theta}_n - \theta)$ має вигляд

$$\mathsf{E}\, d_n(\hat{\theta}_n - \theta)(\hat{\theta}_n - \theta)^T d_n = \sigma^2 \left(d_n^{-1} X^T X d_n^{-1}\right)^{-1} = \sigma^2 \mathcal{L}_n.$$

Тому для $n > n_0$ за оцінкою (20)

$$\mathbf{D}\hat{\theta}_{in} = \sigma^2 \mathcal{L}_n^{ii} d_{in}^{-2} \le \sigma^2 \lambda_0^{-q}(q-1)!\, d_{in}^{-2}. \tag{21}$$

З (19), (21) та умови (i) випливає, що

$$\Delta_n \le \sigma \lambda_0^{-q/2} \sqrt{(q-1)!} \left(\sum_{i=1}^{q} k_i\right) n^{-1/2}. \tag{22}$$

Нерівність (22) та ще не використана умова (17) теореми приводять до співвідношення

$$\mathsf{E}\,|b_n(\hat{Z}_n - a_n) - b_n(Z_n - a_n)| = b_n \Delta_n \to 0, \qquad n\to\infty. \quad \square$$

Наслідок 1. *Нехай виконуються умови (i), (ii). Тоді*

a) Якщо F задовольняє умову (10), *то вірна рівність* (18) *з* $G(x) = \Phi_\alpha$ *та* a_n, b_n *із* (13)*;*

b) Якщо F задовольняє умову (11) *при $\alpha > 2$, то вірна рівність* (18) *з* $G(x) = \Psi_\alpha(x)$ *та* a_n, b_n *із* (14)*;*

c) Якщо F задовольняє умову (12), *то вірна рівність* (18) *з* $G(x) = \Lambda(x)$ *та* a_n, b_n *із* (15).

Доведення наслідку 1 негайно випливає із теореми 1, якщо ми для кожного з випадків a)–c) установимо справедливість (17).

a) Згідно з умовами (10), (13) $x_F = \infty$, $\gamma_n \to \infty$, $b_n = \gamma_n^{-1} \to 0$, $n \to \infty$, тобто (17) вірно.

b) Умова (11) означає, що

$$1 - F(x_F - x) = x^\alpha L(x), \qquad x \downarrow 0,$$

де $L(x)$ повільно змінюється в нулі. Враховуючи рівність (14), маємо

$$\gamma_n = x_F - u_n, \qquad u_n^\alpha L(u_n) = \frac{1}{n}.$$

Останню рівність запишемо так

$$U_1\left(\frac{1}{u_n}\right) = n, \quad \text{де } U_1(x) = x^\alpha L_1(x),\ L_1(x) = \frac{1}{L(1/x)}.$$

Зрозуміло, що $L_1(x)$ повільно змінюється на нескінченості. Відомо (див. твердження 5 [7, с. 27]), що існує функція $L_2(x)$, яка повільно змінюється на нескінченості і для якої

$$U_1(U_2(x)) \sim U_2(U_1(x)) \sim x, \qquad x \to \infty,$$

де $U_2(x) = x^{1/\alpha} L_2(x)$. Тоді при $n \to \infty$

$$n^{1/\alpha} L_2(n) = U_2(n) = U_2(U_1(\frac{1}{u_n})) \sim \frac{1}{u_n}.$$

Таким чином

$$b_n = (x_F - \gamma_n)^{-1} = \frac{1}{u_n} \sim n^{1/\alpha} L_2(n),$$

і при $\alpha > 2$ умова (17) виконується.

Справедливість (17) у випадку *c*) випливає із наступної леми, яка, можливо, представляє і самостійний інтерес.

Лема 1. *Якщо виконується умова* (12) *і* a_n, b_n *визначені рівностями* (15), *то для будь-якого $\delta > 0$ існує $C = C_\delta < \infty$ таке, що*

$$a_n \le C n^\delta, \qquad b_n \le C n^\delta. \tag{23}$$

Доведення леми 1. Нехай

$$R(x) = -\ln(1 - F(x)), \qquad r(x) = R'(x) = \frac{F'(x)}{1 - F(x)}.$$

Звідси та (12)

$$-\frac{r'}{r^2} = \left(\frac{1}{r}\right)' = \left(\frac{1-F}{F'}\right)' = -1 - \frac{(1-F)F''}{(F')^2} \to 0, \qquad x \to x_F. \tag{24}$$

Тому для будь-якого $\delta > 0$ існує $x_1 < x_F$ таке, що для всіх $x \in (x_1, x_F)$

$$\frac{r'(x)}{r(x)} \le \delta r(x),$$

або, що еквівалентно,

$$(\ln r(x))' \le \delta R'(x).$$

Якщо вибрати $C = C_\delta$, для якого $\ln C > \ln r(x_1)$, то для $x \in (x_1, x_F)$

$$\ln r(x) \le \delta R(x) + \ln C,$$

а, отже,

$$r(x) \le C \exp(\delta R(x)).$$

Підставимо в останню нерівність $x = R^{-1}(\ln n)$. Тоді $r(R^{-1}(\ln n)) \le C n^\delta$. Залишається помітити, що в (15) $a_n = \gamma_n = R^{-1}(\ln n)$, $b_n = r(a_n)$, тобто друга оцінка в (23) вірна.

Щоб установити першу оцінку, знову скористаємось співвідношенням (24). Маємо для будь-якого $\delta > 0$, існує $x_2 < x_F$ таке, що для всіх $x \in (x_2, x_F)$

$$\frac{r'(x)}{r(x)} \ge -\delta r(x),$$

або

$$(\ln r(x))' \ge -\delta R'(x).$$

Звідси при умові $\ln C \le \ln r(x_2)$ одержуємо

$$\ln r(x) \ge -\delta R(x) + \ln C$$

і

$$r(x) \ge C \exp(-\delta R(x)).$$

Підставимо в останню нерівність $x = R^{-1}(y)$

$$r(R^{-1}(y)) \ge C \exp(-\delta y).$$

Якщо сюди додати рівність $(R^{-1}(y))' = (r(R^{-1}(y)))^{-1}$, то отримаємо

$$(R^{-1}(y))' \le C' \exp(\delta y),$$

або

$$R^{-1}(y) \le (C'/\delta) \exp(\delta y) + C'',$$

при $C'' \ge R^{-1}(y_2) = x_2$. Виберемо $y = \ln n$. Тоді

$$a_n = R^{-1}(\ln n) \le (C'/\delta) n^\delta + C''.$$

Зрозуміло, що цього досить для справедливості першої оцінки в (23). □

Приклад 1 (Продовження)**.** Якщо в моделі (1) (ε_j) — н.о.р. $N(0, \sigma^2)$ в.в. і виконуються умови (i), (ii), то

$$\lim_{n\to\infty} \mathsf{P}\left\{ b_n \left(\frac{\hat{Z}_n}{\sigma} - a_n \right) < x \right\} = \Lambda(x), \qquad x \in \mathbf{R}^1,$$

з нормуючими послідовностями a_n, b_n із (16).

Приклад 2. Нехай (ε_j) — рівномірно розподілені на $[-1, 1]$ в.в. Для них виконується умова (11) з $\alpha = 1$, а із (14) маємо $a_n = 1$, $b_n = \frac{n}{2}$

$$\lim_{n\to\infty} \mathsf{P}\left\{ \frac{n}{2}(Z_n - 1) < x \right\} = \Psi_1(x), \qquad x \in \mathbf{R}^1.$$

Оскільки $b_n n^{-1/2} \nrightarrow 0$, то умови теореми 1 не виконуються. Покажемо, що і рівність (18) не вірна.

Нехай модель спостережень (1) має найпростіший вид

$$y_j = \theta + \varepsilon_j, \qquad j = 1, \ldots, n.$$

Тоді о.н.к. $\hat{\theta} = \bar{y} = \theta + \bar{\varepsilon}$, і

$$\hat{y}_j = \bar{y}, \qquad \hat{\varepsilon}_j = y_j - \hat{y}_j = \varepsilon_j - \bar{\varepsilon}, \qquad \hat{Z}_n = \max_{1 \le j \le n} \varepsilon_j - \bar{\varepsilon} = Z_n - \bar{\varepsilon}.$$

Далі для нашого прикладу можна записати

$$\frac{n}{2}(\hat{Z}_n - 1) = \frac{n}{2}(Z_n - 1) - \frac{1}{2} \sum_{j=1}^{n} \varepsilon_j.$$

Оскільки $Z_n < 1$ майже напевно, то для $x > 0$ при великих n

$$\mathsf{P}\left\{\frac{n}{2}(\hat{Z}_n - 1) > x\right\} \ge \mathsf{P}\left\{-\frac{1}{2}\left(\frac{2n}{3}\right)^{1/2}\left(\frac{3}{2n}\right)^{1/2}\sum_{j=1}^{n}\varepsilon_j > x\right\}$$
$$\sim 1 - \Phi\left(\left(\frac{6}{n}\right)^{1/2} x\right) \sim \frac{1}{2},$$

де Φ — ф.р. $N(0,1)$ в.в. Це означає, що розподіл $\frac{n}{2}(\hat{Z}_n - 1)$ не збігається до Ψ_1, бо $1 - \Psi_1(x) = 0$ при $x > 0$.

Приклад 3. Припустимо, що в.в. ε_j мають розподіл Стьюдента з s ступенями волі, тобто

$$F_s(x) = K_s \int_{-\infty}^{x} \left(1 + \frac{u^2}{s}\right)^{-(s+1)/2} du, \qquad K_s = \frac{\Gamma((s+1)/2)}{\sqrt{\pi s}\,\Gamma(s/2)}.$$

За правилом Лопіталя для $x > 0$

$$\lim_{t \to \infty} \frac{1 - F_s(tx)}{1 - F_s(t)} = x^{-s},$$

тобто за (10) ф.р. F_s належить області максимум притягнення розподілу Φ_s, і при виконанні умов (i), (ii) справедливе твердження теореми 1.

2. Нелінійна модель регресії

Розглянемо нелінійну модель регресії

$$y_j = g(j, \theta) + \varepsilon_j, \qquad j = 1, \ldots, n, \tag{25}$$

в якій послідовність функцій $g(j, \cdot)$ визначена на деякій відкритій множині $\mathcal{O} \subset \mathbf{R}^q$ такій, що $\Theta^c \subset \mathcal{O}$, де Θ — обмежена відкрита опукла множина, яка містить невідоме істинне значення параметра θ. Вважатимемо, що функції $g(j, \cdot)$ двічі неперервно диференційовні в Θ^c. Відносно помилок спостережень припустимо, що ε_j, $j = 1, \ldots, n$, — н.о.р. в.в., $\mathsf{E}\,\varepsilon_j = 0$.

О.н.к. параметра θ, отриманою за спостереженнями (25), назвемо будь-який вектор $\hat{\theta}_n \in \Theta^c$ із властивістю

$$Q(\hat{\theta}_n) = \inf_{\tau \in \Theta^c} Q(\tau), \qquad Q(\tau) = \sum_{j=1}^{n} (y_j - g(j, \tau))^2.$$

На відміну від лінійного випадку, ми не можемо записати о.н.к. $\hat{\theta}_n$ у явному вигляді.

Введемо необхідні позначення. Нехай $\alpha = (\alpha_1, \ldots, \alpha_q)$ — мульті-індекс,

$$|\alpha| = \alpha_1 + \ldots + \alpha_q, \qquad d_n^{\alpha}(\theta) = d_{1n}^{\alpha_1}(\theta) \ldots d_{qn}^{\alpha_q}(\theta),$$

де $d_{in}^2(\theta) = \sum_{j=1}^{n} g_i^2(j, \theta)$, $g_i = \frac{\partial}{\partial \theta_i} g$, $i = 1, \ldots, q$.

Загальніше позначення

$$d_n(\alpha,\theta)=\sum_{j=1}^{n} g^{(\alpha)}(j,\theta)^2, \qquad g^{(\alpha)}=\frac{\partial^{|\alpha|}}{(\partial\theta_1)^{\alpha_1}\dots(\partial\theta_q)^{\alpha_q}}g.$$

Якщо $e_i\in\mathbf{R}^q$ — одиничний орт, то $d_n(e_i,\theta)=d_{in}(\theta)$, $i=1,\dots,q$.

Ми розглядаємо тільки випадок, коли

$$0<\liminf_{n\to\infty} n^{-1/2}d_{in}(\theta), \qquad i=1,\dots,q. \tag{26}$$

Позначимо $d_n(\theta)=\mathrm{diag}(d_{in}(\theta),i=1,\dots,q)$,

$$f^{(\alpha)}(j,u)=g^{(\alpha)}\left(j,\theta+n^{1/2}d_n^{-1}(\theta)u\right), \qquad U_n^c(\theta)=n^{-1/2}d_n(\theta)(\Theta^c-\theta),$$

$$\Phi_n^{(\alpha)}(u_1,u_2)=\sum_{j=1}^{n}\left(f^{(\alpha)}(j,u_1)-f^{(\alpha)}(j,u_2)\right)^2, \quad u_1,u_2\in U_n^c(\theta),$$

$$J_n(\theta)=\left(d_{kn}^{-1}(\theta)d_{ln}^{-1}(\theta)\sum_{j=1}^{n} g_k(j,\theta)g_l(j,\theta)\right)_{k,l=1}^{q},$$

$$\Lambda_n(\theta)=\left(\Lambda_n^{kl}(\theta)\right)_{k,l=1}^{q}=J_n^{-1}(\theta).$$

Позначимо також

$$S(r)=\{u\in\mathbf{R}^q\colon \|u\|<r\}, \quad r>0; \qquad u_n(\theta)=n^{-1/2}d_n(\theta)(\hat\theta_n-\theta).$$

У випадку нелінійної моделі регресії для доведення твердження аналогічного теоремі 1 доцільно використати деяку властивість консистентності о.н.к. $\hat\theta_n$ тому, що оцінка моментів $\hat\theta_n$ у нелінійної моделі вимагає жорсткіших умов на функцію регресії g, ніж отримання властивостей консистентності [2].

Наведемо умови, за яких о.н.к. $\hat\theta_n$ має потрібну властивість.

$\mathcal{A}_1$. $\mu_m=\mathsf{E}|\varepsilon_j|^m<\infty$ для деякого цілого $m\ge 3$.

$\mathcal{A}_2$. Для будь-якого $r>0$ існують константи $C_i(\alpha,r)$, $i=1,2,3$, такі, що

$$\sup_{u\in S^c(r)\cap U_n^c(\theta)} n^{(|\alpha|-1)/2}(d_n^{\alpha}(\theta))^{-1}d_n\left(\alpha,\theta+n^{1/2}d_n^{-1}(\theta)u\right)\le C_1(\alpha,r), \quad |\alpha|=1,2; \tag{27}$$

$$\sup_{u_1,u_2\in S^c(r)\cap U_n^c(\theta)} n(d_n^{\alpha}(\theta))^{-2}\Phi_n^{(\alpha)}(u_1,u_2)\|u_1-u_2\|^{-2}\le C_2(\alpha,r), \qquad |\alpha|=2; \tag{28}$$

$$\sup_{u\in S^c(r)\cap U_n^c(\theta)} n^{1/2}(d_n(\alpha,\theta))^{-1}\max_{1\le j\le n}\left|f^{(\alpha)}(j,u)\right|\le C_3(\alpha,r), \qquad |\alpha|=1,2. \tag{29}$$

$\mathcal{A}_3$. Якщо $g^{(\alpha)}(j,\theta)\ne 0$, то

$$\liminf_{n\to\infty} n^{1/2}(d_n^{\alpha}(\theta))^{-1}d_n(\alpha,\theta)>0, \qquad |\alpha|=2. \tag{30}$$

$\mathcal{A}_4$. Для $n>n_0$ $\quad\lambda_{\min}(J_n(\theta))\ge\lambda_0>0$.

$\mathcal{A}_5$. Для будь-якого $r>0$ і m із умови $\mathcal{A}_1$

$$\mathsf{P}\{\|u_n(\theta)\|\ge r)\}=O\left(n^{-(m-2)/2}\right). \tag{31}$$

Бачимо, що умова (29) узагальнює (i), а умова $\mathcal{A}_4$ аналогічна умові (ii) 1-го розділу статті. Співвідношення (31), яке узагальнює результат Малинво [8], є підсиленою властивістю слабкої консистентності о.н.к. $\hat\theta_n$. Умови, за яких вірне (31), та доведення цього факту міститься в теоремі 8 [2, стор. 30–32].

Наступний факт підсилює (31) і є переформулюванням теореми 19, стор. 91 книги [2]. Позначимо $\chi(\delta) = 2\lambda_0^{-1}\sigma q(m-2+\delta)^{1/2}$, де $\sigma^2 = \mu_2$.

Теорема 2. *Якщо виконано умови $\mathcal{A}_1 - \mathcal{A}_5$, то*

$$\mathsf{P}\left\{\|u_n(\theta)\| \geq \chi(\delta) n^{-1/2}\ln^{1/2} n\right\} = O\left(n^{-(m-2)/2}\right) \tag{32}$$

для будь-якого фіксованого $\delta > 0$.

Співвідношення (32) можна переписати у вигляді

$$\mathsf{P}\left\{\|d_n(\theta)(\hat{\theta}_n - \theta)\| \geq \chi(\delta)\ln^{1/2} n\right\} = O\left(n^{-(m-2)/2}\right). \tag{33}$$

Таким чином, теорема 2 — це результат про помірні відхилення нормованої о.н.к. $d_n(\theta)(\hat{\theta}_n - \theta)$.

Покладемо $\hat{y}_j = g(j, \hat{\theta}_n)$ і, використовуючи позначення 1- го розділу роботи, сформулюємо основний результат цього розділу.

Теорема 3. *Нехай ф.р. F в.в. ε_j в моделі (25) належить області максимум притягнення закону G, виконуються умови $\mathcal{A}_1 - \mathcal{A}_5$ та*

$$\left(\frac{\ln n}{n}\right)^{1/2} b_n \to 0, \qquad n \to \infty. \tag{34}$$

Тоді вірне співвідношення (18), у якому G — один із трьох типів розподілів Φ_α, Ψ_α, Λ.

Доведення теореми 3. Як і раніше,

$$|\hat{Z}_n - Z_n| \leq \max_{1\leq j\leq n} |\hat{\varepsilon}_j - \varepsilon_j| = \max_{1\leq j\leq n}\left|g(j,\hat{\theta}_n) - g(j,\theta)\right|.$$

З іншого боку,

$$g(j,\hat{\theta}_n) - g(j,\theta) = f(j, u_n(\theta)) - f(j,0) = \sum_{i=1}^{q} f_i(j, u_n^*(\theta)) n^{1/2} d_{in}^{-1}(\theta) u_{in}(\theta),$$

де $\|u_n^*(\theta)\| < \|u_n(\theta)\|$. Якщо $\|u_n(\theta)\| < 1$, то за умови (29)

$$\begin{aligned}|\hat{Z}_n - Z_n| &\leq \sum_{i=1}^{q} n^{1/2} d_{in}^{-1}(\theta) \max_{1\leq j\leq n} |f_i(j, u_n^*(\theta))| \cdot |u_{in}(\theta)| \\ &\leq \sum_{i=1}^{q} C_3(e_i, 1)|u_{in}(\theta)| \leq \|\mathbb{C}\| \cdot \|u_n(\theta)\|,\end{aligned}$$

де $\mathbb{C} = (C_3(e_1, 1)), \ldots, C_3(e_q, 1))$.

Маємо тепер для довільного $\varepsilon > 0$

$$\mathsf{P}\left\{b_n|\hat{Z}_n - Z_n| \geq \varepsilon\right\} \leq \mathsf{P}\left\{b_n\|u_n(\theta)\| \geq \frac{\varepsilon}{\|\mathbb{C}\|}\right\} + \mathsf{P}\{\|u_n(\theta)\| \geq 1\} = \mathcal{P}_1 + \mathcal{P}_2.$$

Завдяки $\mathcal{A}_5$, $\mathcal{P}_2 \to 0$. Далі оцінимо перший доданок

$$\begin{aligned}\mathcal{P}_1 &= \mathsf{P}\left\{b_n\|u_n(\theta)\| \geq \frac{\varepsilon}{\|\mathbb{C}\|}, \|u_n(\theta)\| \geq \chi(\delta)n^{-1/2}\ln^{1/2} n)\right\} \\ &\quad + \mathsf{P}\left\{b_n\|u_n(\theta)\| \geq \frac{\varepsilon}{\|\mathbb{C}\|}, \|u_n(\theta)\| < \chi(\delta)n^{-1/2}\ln^{1/2} n)\right\} \\ &= \mathcal{P}_3 + \mathcal{P}_4.\end{aligned}$$

За умовою (34) теореми 3, $\mathcal{P}_4 = 0$, починаючи з деякого номера, а $\mathcal{P}_3 = o(n^{-(m-2)/2})$ за теоремою 2. $\square$

Очевидно, присутність множника $\ln^{1/2} n$ в умові (34) не змінює доведення наслідку 1 теореми 1, і тому до теореми 3 ми можемо сформулювати

Наслідок 2. *Нехай для моделі* (25) *виконуються умови* $\mathcal{A}_1 - \mathcal{A}_5$. *Тоді вірні твердження* $a), b)$ *і* $c)$ *наслідку 1.*

Зробимо одне зауваження. Якщо ми візьмемо, зокрема, у моделі (25)

$$g(j,\theta) \equiv g(x_j,\theta), \qquad x_j = (x_{j1},\dots,x_{jl}) \in \mathbb{X} \subset \mathbf{R}^l, \qquad j=1,\dots,n,$$

де $\mathbb{X}$ — деяка область планування регресійного експерименту, то стає очевидним, що модель спотережень (25) є узагальненням моделі (1).

Припустимо, що $\mathbb{X}$ компактна множина, похідні за параметрами функції регресії $g(x,\theta)$ неперервні за сукупністю змінних $(x,\theta) \in \mathbb{X} \times \Theta^c$. Тоді, очевидно,

$$\sup_{\theta\in\Theta^c} d_n(\alpha,\theta) \le \bar{c}(\alpha) n^{1/2}, \qquad |\alpha| = 1,2, \tag{35}$$

де $\bar{c}(\alpha)$ — деякі константи. Ми розглядаємо тільки випадок (див. (26)), коли

$$d_{in}(\theta) \ge \underline{c}_i n^{1/2} \tag{36}$$

для деяких $\underline{c}_i > 0$, $i = 1,\dots,q$. А у випадку, коли $|\alpha| = 2$ і $g^{(\alpha)}(x,\theta) \ne 0$, вважатимемо, що

$$d(\alpha,\theta) \ge \underline{c}(\alpha) n^{1/2} \tag{37}$$

для деяких $\underline{c}(\alpha) > 0$.

За умов (35) та (36) без обмеження загальності ми можемо використовувати нормування $n^{1/2}\mathbb{I}_q$ замість $d_n(\theta)$. Відповідним чином спростяться й умови теореми 2. Так, умови (27), (29) і (30) випливають із (35)-(37), а умова (28) набуває вигляду

$$n^{-1} \sum_{j=1}^{n} \left(g^{(\alpha)}(x_j,\theta_1) - g^{(\alpha)}(x_j,\theta_2) \right)^2 \|\theta_1 - \theta_2\|^{-2} \le c(\alpha),$$

θ_1, $\theta_2 \in \Theta^c$, $|\alpha| = 2$.

3. Деякі застосування. Критерій адекватності моделі регресії

Досить часто на практиці виникає потреба в перевірці адекватності моделі регресії. Розглянемо на прикладі простої лінійної регресії застосування до цієї задачі теорем для екстремальних залишків.

Нехай

$$y_j = \theta_0 + \theta_1 x_j + \varepsilon_j, \qquad j = 1,\dots,n, \tag{38}$$

де ε_j — н.о.р. $N(0,\sigma^2)$ в.в. І нехай $\hat\theta_0, \hat\theta_1$ о.н.к. для параметрів θ_0, θ_1,

$$\hat y_j = \hat\theta_0 + \hat\theta_1 x_j, \quad \hat\varepsilon_j = y_j - \hat y_j,$$

$$\hat Z_n = \max_{1\le j\le n} \hat\varepsilon_j, \quad \bar x = \frac{1}{n}\sum_{j=1}^n x_j, \quad s^2 = \frac{1}{n}\sum_{j=1}^n (x_j - \bar x)^2.$$

Далі всюди константи a_n, b_n задаються рівностями (16). Так само, як і в теоремі 1, можна довести рівність

$$\lim_{n\to\infty} \mathsf{P}\left\{ b_n \left(\frac{\hat Z_n}{\sigma} - a_n \right) < x \right\} = \Lambda(x), \qquad x \in \mathbf{R}^1, \tag{39}$$

при умові

$$\frac{\ln n}{n} \frac{\max_{1\le j\le n} |x_j|^2}{s_n^2} \to 0, \qquad n \to \infty. \tag{40}$$

(Ця умова трохи слабкіша за умови теореми 1.)

Для наших цілей рівність (39) треба посилити наступним чином.

Теорема 4. *Для моделі* (38) *при умові* (40)

$$\lim_{n\to\infty} \mathsf{P}\left\{ b_n \left(\frac{\hat{Z}_n^*}{\hat{\sigma}_n} - a_n \right) < x \right\} = \Lambda^2(x), \qquad x \in \mathbf{R}^1, \tag{41}$$

де $\hat{Z}_n^* = \max_{1\le j\le n} |\hat{\varepsilon}_j|$, $\hat{\sigma}_n^2 = (n-2)^{-1}\sum_{j=1}^n \hat{\varepsilon}_j^2$.

Доведення теореми 4. Спочатку установимо співвідношення

$$\lim_{n\to\infty} \mathsf{P}\left\{ b_n \left(\frac{\hat{Z}_n^*}{\sigma} - a_n \right) < x \right\} = \Lambda^2(x), \tag{42}$$

при цьому, не обмежуючи загальності, можем вважати $\sigma = 1$.

Якщо позначимо

$$Z_n^* = \max_{1\le j\le n} |\varepsilon_j|, \qquad W_n = \min_{1\le j\le n} \varepsilon_j,$$

то при $u_n = x/b_n + a_n$, $n \to \infty$

$$\mathsf{P}\{b_n(Z_n^* - a_n) < x\} = \mathsf{P}\{Z_n^* < u_n\} = \mathsf{P}\{Z_n < u_n, W_n > -u_n\} \to \Lambda^2(x).$$

Останнє асимптотичне співвідношення випливає із теореми 1.8.2 ([6, с. 40]), бо для стандартної нормальної ф.р. $\Phi(x)$ маємо (див. [6, с. 25, 26])

$$n\Phi(-u_n) = n(1-\Phi(u_n)) \to \exp(-x) = -\ln(\Lambda(x)).$$

Щоб вивести звідси рівність (42), досить записати нерівність

$$|Z_n^* - \hat{Z}_n^*| \le \max_{1\le j\le n} |\varepsilon_j - \hat{\varepsilon}_j|$$

і скористатися міркуваннями із доведення теореми 1.

Таким чином, нам залишається довести імплікацію: (42) $\Rightarrow$ (41). Покладемо

$$\zeta_n = b_n \left(\frac{\hat{Z}_n^*}{\sigma} - a_n \right), \qquad \zeta_n^{'} = b_n \left(\frac{\hat{Z}_n^*}{\hat{\sigma}_n} - a_n \right).$$

Згідно рівності (42)

$$\zeta_n \xrightarrow{\mathbf{D}} \zeta, \qquad n \to \infty, \tag{43}$$

де $\mathsf{P}\{\zeta < x\} = \Lambda^2(x)$.

Безпосередньо із означення ζ_n та $\zeta_n^{'}$ маємо

$$\zeta_n - \zeta_n^{'} = b_n \frac{\hat{Z}_n^*}{\sigma} \left(\frac{\hat{\sigma}_n - \sigma}{\hat{\sigma}_n} \right) = \frac{(\zeta_n + a_n b_n)(\hat{\sigma}_n^2 - \sigma^2)}{\hat{\sigma}_n(\hat{\sigma}_n + \sigma)}. \tag{44}$$

Далі скористаємось відомою оцінкою (див., наприклад, [2])

$$\mathsf{E}\left|\hat{\sigma}_n^2 - \sigma^2\right| \le \sqrt{\mathbf{D}\,\hat{\sigma}_n^2} = O\left(\frac{1}{\sqrt{n}}\right). \tag{45}$$

Отже $\hat{\sigma}_n^2 \xrightarrow{\mathsf{P}} \sigma^2$, а враховуючи (43),

$$\zeta_n\left(\hat{\sigma}_n^2 - \sigma^2\right) \xrightarrow{\mathsf{P}} 0, \qquad n \to \infty.$$

Аналогічно (16) та (45) дають

$$\mathsf{E}\, a_n b_n \left|\hat{\sigma}_n^2 - \sigma^2\right| = a_n b_n O\left(\frac{1}{\sqrt{n}}\right) \to 0, \qquad n \to \infty.$$

Разом ці співвідношення та (44) дозволяють записати

$$\zeta_n - \zeta_n^{'} \xrightarrow{\mathsf{P}} 0, \qquad n \to \infty.$$

Звідси та (42) випливає рівність (41). □

Зрозуміло, що результат теореми 4 легко перенести в дусі доведення теореми 1 на множинну лінійну гаусівську регресію.

Асимптотичні рівності (41) і (42) можуть бути основою для перевірки гіпотези $H_0 = \{y = \theta_0 + \theta_1 x + \varepsilon, \varepsilon \sim N(0, \sigma^2)\}$ проти належним чином обраної альтернативи H_1. Треба вибрати статистику критерію у вигляді

$$T_n = b_n \left(\frac{\hat{Z}_n^*}{\hat{\sigma}_n} - a_n \right) \quad (\sigma \text{ невідоме}) \qquad \text{та} \qquad T_n^{'} = b_n \left(\frac{\hat{Z}_n^*}{\sigma} - a_n \right) \quad (\sigma \text{ відоме}).$$

Якщо вірна гіпотеза H_0, то при великих n статистики T_n та $T_n^{'}$ мають наближено розподіл $\Lambda^2(x)$.

Припустимо, що рівень значущості цього критерію дорівнює α . Тоді критична область є

$$K = \left(-\infty, -\ln\ln\sqrt{\frac{2}{\alpha}} \right) \cup \left(-\ln\ln\sqrt{\frac{2}{2-\alpha}}, \quad +\infty \right).$$

Так, наприклад, якщо $\alpha = 0.05$, то

$$K = (-\infty, -0.61218) \cup (4.36939, +\infty).$$

Наведемо лише два можливих приклади альтернативних гипотез H_1, які можуть бути поміченими критеріями на базі статистик T_n і $T_n^{'}$.

(i) $H_1 = \{y = \theta_0 + \theta_1 x + \varepsilon, \varepsilon$ має розподіл з хвостами, відмінними від нормальних$\}$.

(ii) Нехай $g(x, \theta) = \sum_{k=0}^{m} \theta_k x^k$ — поліном степеня $m \geq 2$. Розглянемо альтернативу $H_1 = \{y = g(x, \theta) + \varepsilon, \varepsilon \sim N(0, \sigma^2)\}$. В найцікавішому с практичної точки зору випадку, коли вірна гипотеза H_1, σ^2 — невідома, асимптотична поведінка T_n залишається незрозумілою. Ясно, що для точних тверджень тут потрібні подальші дослідження.

Література

1. Дж. Себер, *Линейный регрессионный анализ*, "Мир", Москва, 1980.
2. A. V. Ivanov, *Asymptotic Theory of Nonlinear Regression*, Kluwer Academic Press, Dordrecht, 1997.
3. B. V. Gnedenko, *Sur la distribution limit du terme maximum d'une serie aleatoire*, Ann. Math. **44** (1943), 423–453.
4. Я. Галамбош, *Асимптотическая теория экстремальных порядковых статистик*, "Наука", Москва, 1984.
5. S. I. Resnick, *Extreme Values, Regular Variation and Point Processes*, Springer, Berlin, 1987.
6. М. Лидбеттер, Г. Линдгрен, Х. Ротсен, *Экстемумы случайных последовательностей и процессов*, "Мир", Москва, 1989.
7. Е. Сенета, *Правильно меняющиеся функции*, "Наука", Москва, 1985.
8. E. Malinvaud, *The Consistency of Nonlinear Regression*, Ann. Math. Statist. **41** (1970), 953–969.

Національний технічний університет "Київський політехнічний інститут", пр. Перемоги, 37, Київ 03056

Адреса електронної пошти: alexntuu@gmail.com

Київський національний університет ім. Тараса Шевченка, пр. Глушкова 2, корп. 6, Київ 03127

Адреса електронної пошти: mik@unicyb.kiev.ua

Надійшла 09/10/2011