4. **Zinovyev A. Y., Gorban A. N., Sumner N. R.** Topological grammars for data approximation / Appl. Math. Lett. 2007. Vol. 20. N 4. P. 382 – 386.
5. **Chi-Hyuck Jun, Il-Gyo Chong**. Performance of some variable selection methods when multicollinearity is present / Chemometrics and Intelligent Laboratory Systems. 2005. Vol. 78. N 1, 2. P. 103 – 112.
6. **Jiang Guohua, Wang Hansheng, Li Guodong**. Robust regression shrinkage and consistent variable selection through the LAD-lasso / J. Business Econ. Stat. 2008. Vol. 25. P. 347 – 355.
7. **Herzog F., Hildmann M.** Robust calculation and parameter estimation of the hourly price forward curve / 17th Power Systems Computation Conference. Stockholm. 2011. P. 1 – 7.
8. **Efron B., Hastie T., Johnstone I., Tibshirani R.** Least angle regression / The Annals of Statistics. 2004. Vol. 32. N 3. P. 407 – 499.
9. **Stepashko V. S., Ivakhnenko A. G.** Pomekhoustoichivost' modelirovaniya [Noise Immunity of modeling]. — Kiev: Naukova dumka, 1985. — 216 p. [in Russian].
10. **Smith H., Draper N. R.** Appied regression analysis. — New York: John Wihley and Sons, 1998. — 736 p.
11. **Grant P. M., Chen S., Cowan S. F. N.** Orthogonal least squares learning algorithm for radial basis function network / Neural Networks. 1991. Vol. 2. N 2. P. 302 – 309.
12. **Belsley A. D.** Conditioning Diagnostics: Collinearity and Weak Data in Regression. — New York: John Wiley and Sons, 1991. — 396 p.
13. **Abdolkhalig A.** Optimized calculation of hourly price forward curve (HPFC) / Int. J. Electr. Comp. Electronics Comm. Eng. 2008. Vol. 2. N 9. P. 840 – 850.
14. **Caro G., Hildmann M.** What makes a good hourly price forward curve? / European Energy Market, IEEE 10th International Conference, 2013. Stockholm. P. 1 – 7.
15. **Kachapova F., Kachapov I.** Orthogonal projection in teaching regression and financial mathematics / J. Stat. Education. 2010. Vol. 18. N 1. P. 1 – 18.
16. Time series with electricity prices: https://svn.code.sf.net/p/dmba/code/data/germanspotprice.csv

УДК 519.24

ЛОКАЛЬНАЯ АСИМПТОТИЧЕСКАЯ НОРМАЛЬНОСТЬ СТАТИСТИЧЕСКИХ ЭКСПЕРИМЕНТОВ И ЕЕ РОЛЬ В ТЕОРИИ ОЦЕНИВАНИЯ И ПРОВЕРКИ ГИПОТЕЗ

© **А. А. Абдушукуров, Н. С. Нурмухамедова**[1]

Статья поступила 18 июня 2015 г.

Основной задачей теории оценивания является нахождение оптимальных оценок для неизвестных параметров. Существуют два подхода к решению этих задач. Первый основан на выборке конечного объема, второй — асимптотический — на выборке с растущим объемом. Асимптотический подход может обладать свойствами оптимальности при $n \to \infty$. Он базируется на понятии асимптотической минимаксности оценок. Локальная асимптотическая минимаксность оценок опирается на асимптотическое поведение последовательности статистических экспериментов при сближающихся последовательностях альтернативных гипотез. В данной работе рассмотрена асимптотическая нормальность оценок байесовского типа и асимптотически минимаксная эффективность оценок максимального правдоподобия с использованием свойства локальной асимптотической нормальности статистики отношения правдоподобия в модели случайного цензурирования с двух сторон.

Ключевые слова: локальная асимптотическая нормальность; статистика отношения правдоподобия; асимптотическая минимаксная эффективность; случайное цензурирование.

Статистика отношения правдоподобия (СОП) играет фундаментальную роль в теории принятия решений, особенно в теории проверки статистических гипотез. Пусть $(\mathbf{X}^{(n)}, \mathbf{B}^{(n)}, \mathbf{P}^{(n)})$ — статистическая модель, соответствующая повторной независимой выборке наблюдений $X^{(n)} = (X_1, ..., X_n)$ случайной величины (с.в.) X с распределением $Pl_\theta \in \mathbf{P} = \{P_\theta, \theta \in \Theta\}$, где Θ — открытое множество в R^1, θ — неизвестный параметр. При общей постановке задач теории проверки гипотез предполагается, что неизвестный параметр исходного распределения P_θ принадлежит заданному подмножеству $\Theta_0 \in \Theta$ (гипотеза H_0) множества возможных значений параметра θ. Дополнительное предположение (альтернативная гипотеза к основной гипотезе H_0) подразумевает, что $\theta \in \Theta_1 = \Theta \backslash \Theta_0$ (гипотеза H_1). Основная задача теории состоит в проверке соответствия реальных экспериментальных данных предполагаемой гипотезе на основе статистического критерия, т.е. процедуры, позволяющей принимать или отвергать данную гипотезу. Критерии дают возможность утверждать, что результаты наблюдений не противоречат принятой гипотезе, т.е. статистические выводы формулируются в следующем виде: экспериментальные данные согласуются с данной гипотезой (или

[1] Национальный университет Узбекистана им. М. Улугбека, г. Ташкент, Узбекистан; e-mail: a_abdushukurov@rambler.ru

противоречат ей). Однако процедура проверки гипотезы сопряжена с ошибками двух родов: отвергнуть гипотезу H_0, когда она верна (ошибка I рода); принять гипотезу H_0, когда она неверна (ошибка II рода). Естественно, возникает задача о построении критериев проверки гипотез, которые позволяют минимизировать значения (или вероятности) ошибок обоих родов. Однако на практике часто невозможно построение статистических критериев со сколь угодно малыми ошибками I и II родов. Это и составляет основу статистических задач теории проверки гипотез. Для данной пары гипотез H_0 и H_1 существуют разнообразные точные или асимптотические (при $n \to \infty$) критерии их проверки. Задача заключается в выборе самого оптимального критерия (например, несмещенного, наиболее мощного) среди них для проверки H_0 против H_1. Среди разнообразных критериев следует особо выделить основанные на СОП. Они играют фундаментальную роль в теории проверки гипотез. Согласно лемме Неймана – Пирсона критерии, основанные на СОП, являются оптимальными по сравнению с другими критериями, построенными на основе других статистик.

Интерес представляет случай, когда гипотеза H_1 становится близкой к основной гипотезе H_0 с ростом n, т.е. $H_1 = H_{1n} \to H_0$ при $n \to \infty$. Предположим, что основная гипотеза H_0 является простой: H_0:$\theta = \theta_0$, а альтернатива — сложной, но близкой к H_0, т.е. — H_{1n}:$\theta_n = \theta_0 + \Gamma_n(\theta_0)^{-1}h_n$, где $h_n \to h \in \Theta$ при $n \to \infty$. Зададим СОП

$$l_n(h_n) = \frac{\mathrm{d}P^{(n)}_{\theta_0+\Gamma_n^{-1}h_n}}{\mathrm{d}P^{(n)}_{\theta_0}}. \tag{1.1}$$

На базе асимптотических свойств статистики (1.1) создана целая теория, играющая важную роль в математической статистике. Асимптотическая теория оценивания и проверки гипотез всецело опирается на свойства этой статистики. Самое важное свойство СОП (1.1) — локальная асимптотическая нормальность (ЛАН), которая дает возможность развития асимптотической теории оценок максимального правдоподобия (ОМП) и байесовских оценок, а также контигуальности вероятностных мер [1 – 10].

Нахождение оптимальных оценок для неизвестных параметров является основной задачей теории оценивания. Существуют два подхода к решению этих задач: первый основан на выборке конечного объема, второй — асимптотический — на выборке с растущим объемом. Первый подход зависит лишь от модели и объема выборки и, следовательно, является относительно ограничительным. С другой стороны, асимптотический подход может обладать свойством оптимальности при $n \to \infty$. Данное свойство называют локальной асимптотической минимаксностью оценок, и оно опирается именно на свойство ЛАН последовательности статистических экспериментов. В данной работе определяется свойство ЛАН статистических моделей в случае полных наблюдений. С его использованием исследуются асимптотические свойства оценки байесовского типа для неизвестного параметра, а также ОМП и ее асимптотическое свойство в модели случайного цензурирования с двух сторон.

Локальная асимптотическая нормальность статистических экспериментов

Пусть $\mathbf{E}^{(n)} = (\mathbf{X}^{(n)}, \mathbf{B}^{(n)}, \{P^{(n)}_\theta, \theta \in \Theta\})$, $\Theta \subset R^1$, — последовательность статистических экспериментов, соответствующая независимой повторной выборке $\mathbf{X}^{(n)} = (X_1, ..., X_n)$. Рассмотрим СОП при θ_0, $\theta_1 \in \Theta$, $\theta_0 \ne \theta_1$:

$$l_n(\mathbf{X}^{(n)}) = \frac{\mathrm{d}P_{\theta_1}(X^{(n)})}{\mathrm{d}P_{\theta_0}(X^{(n)})} \text{ —} \tag{2.1}$$

производную абсолютно непрерывной компоненты меры $P^{(n)}_{\theta_1}$ по $P^{(n)}_{\theta_0}$ на наблюдении $\mathbf{X}^{(n)}$. Статистика (2.1) играет важную роль как в теории оценивания, так и проверке гипотез. Из леммы Неймана – Пирсона [2] известно, что критерий проверки гипотезы H_0: $\theta = \theta_0$ против альтернативы H_1: $\theta = \theta_1$ является наиболее мощным. Если $\theta_1 = \hat{\theta}_n$ ОМП для истинного значения θ_0 неизвестного параметра, то статистика $2\log l_n$ имеет асимптотическое хи-квадрат распределение со степенью свободы s и этот результат можно использовать для построения доверительного интервала для θ_0 или же проверки гипотез о параметре θ [3]. Еще одно важное свойство СОП (2.1) проявляется в случае, когда альтернатива θ_1 является «близкой» при $n \to \infty$ к θ_0. Это свойство называется ЛАН, остановимся на нем подробней. При $\theta_1 = \theta_0 + \Gamma_n(\theta_0)^{-1}h_n$ рассмотрим СОП (1.1)

$$l_n(h_n) = \frac{\mathrm{d}P^{(n)}_{\theta_0+\Gamma_n^{-1}h_n}(X^{(n)})}{\mathrm{d}P^{(n)}_{\theta_0}(X^{(n)})}.$$

Определение [1,8]. Семейство $\{P^{(n)}_\theta, \theta \in \Theta\}$ удовлетворяет условию ЛАН в точке $\theta_0 \in \Theta$ при $n \to \infty$, если для некоторой $\Gamma(\theta)$ и последовательности $h_n \to h \in \Theta$ справедливо представление

$$l_n(h_n) = (1 + \varepsilon_n(h))\exp\left\{h\Delta_n(\theta_0) - \frac{1}{2}h^2\Gamma(\theta_0)\right\}, \tag{2.2}$$

где для любого $h \in \Theta$: $\varepsilon_n(h) \to 0$ по $P^{(n)}_{\theta_0}$-вероятности; $L(\Delta_n(\theta_0)/P^{(n)}_{\theta_0}) \to N(0; \Gamma(\theta_0))$ при $n \to \infty$; $\Gamma(\theta_0) \ne 0$ и не зависит от h.

Рассмотрим случай независимых и одинаково распределенных регулярных экспериментов с невырожденной фишеровской информацией $I(\theta)$ в точке θ_0. Положим $\Gamma_n = n^{-1/2}$, $h_n = h \in \Theta$, $\Gamma(\theta) = I(\theta)$ и

$$\Delta_n(\theta_0) = n^{-1/2} \sum_{i=1}^{n} \frac{\partial \log f(X_i, \theta_0)}{\partial \theta}, \qquad (2.3)$$

где $f(x, \theta) = dP_\theta/d\nu$ — плотность «одномерного» эксперимента $\mathbf{E}_i = (\mathbf{X}, \underline{\mathbf{B}}, \{P_\theta, \theta \in \Theta\})$, соответствующего наблюдению X_i, $i = \overline{1, n}$. При этих выборах Γ_n, h_n, $\Gamma(\theta)$ и $\Delta_n(\theta_0)$ имеет место следующее утверждение Гаека.

Теорема [1,5]. Пусть $\Theta \subset R^1$, а плотность вероятности $f(x; \theta)$ экспериментов $\mathbf{E}_i$ удовлетворяет условиям регулярности:

I) функция $f(x, \theta)$ абсолютно непрерывна по θ в некоторой окрестности U_{θ_0} точки θ_0 для всех $x \in \mathbf{X}$;

II) производная $\partial f(x, \theta)/\partial\theta$ существует при каждом $\theta \in U_{\theta_0}$ для ν — почти всех $x \in \mathbf{X}$;

III) информация Фишера

$$I(\theta) = M_\theta \left[\frac{\partial \log f(X_i, \theta)}{\partial \theta} \right]^2 \text{ —}$$

непрерывна и положительна при $\theta = \theta_0$.

Тогда семейство мер $\{P_\theta^{(n)}, \theta \in \Theta\}$, соответствующее последовательности независимых экспериментов $\mathbf{E}^{(n)} = \mathbf{E}_1 \times ... \times \mathbf{E}_n$ с плотностью f, удовлетворяет условию ЛАН (2.2).

Отметим одно важное применение ЛАН в теории проверки гипотез. Известно [2, стр. 84], что для семейства экспоненциальных плотностей существует равномерно наиболее мощный критерий $\varphi_n = \varphi_n(X^{(n)})$ уровня α. Поскольку ЛАН позволяет аппроксимировать СОП экспоненциальной плотностью $\exp\{\theta\Delta_n(\theta_0) - A(\theta; \theta_0)\}$, то при помощи статистики $\Delta_n(\theta_0)$ можно построить асимптотически равномерно наиболее мощный критерий $\psi_n = \psi_n(X^{(n)})$ того же уровня α, т.е. среди всех критериев ψ_n таких, что

$$\limsup_{n\to\infty} \sup_{\theta \le \theta_0} M_{\theta_0} \psi_n \le \alpha,$$

справедливо неравенство

$$\limsup_{n\to\infty} \left[\sup_{\theta > \theta_0} (M_\theta \varphi_n - M_\theta \psi_n) \right] \le 0.$$

Критерий ψ_n задается как

$$\psi_n = \begin{cases} 1, \Delta_n(\theta_0) > c_n, \\ \gamma_n, \Delta_n(\theta_0) = c_n, \\ 0, \Delta_n(\theta_0) < c_n, \end{cases}$$

где c_n и γ_n определяются равенством $M_{\theta_0} \psi_n = \alpha$.

Используя свойство ЛАН, можно найти мощность критерия. Рассмотрим задачу проверки гипотезы H_0: $\theta = \theta_0$ против альтернативы H_1: $\theta = \theta_0 + u/\sqrt{n}$, $u > 0$. Согласно фундаментальной лемме Неймана – Пирсона наиболее мощный критерий существует и задается с критической областью $S_{n,\alpha} = \{x^{(n)}: l_n > c_\alpha\}$, где константа c_α определяется из условия $P_{H_0}(l_n > c_\alpha) = \alpha$. Обозначим через β_n^* мощность критерия при альтернативе H_1: $\beta_n^* = P_{H_1}(l_n > c_\alpha)$ и найдем ее предел при $n \to \infty$. Согласно свойству ЛАН

$$P_{H_1}(l_n < x) \underset{n\to\infty}{\to} N\left(-\frac{1}{2} u^2 I(\theta_0); u^2 I(\theta_0) \right).$$

Отсюда следует, что

$$\beta_n^* \underset{n\to\infty}{\to} \beta^* = 1 - \Phi\left(\frac{c_\alpha}{u\sqrt{I(\theta_0)}} + \frac{1}{2} u \sqrt{I(\theta_0)} \right),$$

где $\Phi(x)$ — функция распределения стандартного нормального закона.

Как следует из вышеизложенного, свойства ЛАН для СОП и связанные с ней свойства статистических оценок для случая полной выборки к настоящему времени довольно подробно исследованы. Значительный интерес представляет исследование аналогичных свойств и в случае неполных цензурированных наблюдений. В работах [11 – 17] установлены свойства ЛАН и асимптотические представления для СОП в моделях конкурирующих рисков с неполными наблюдениями. Рассмотрим асимптотическую эффективность оценок байесовского типа с использованием свойства ЛАН для СОП в модели случайного цензурирования с двух сторон.

Асимптотическая эффективность оценок байесовского типа. Неполные наблюдения

В медико-биологических исследованиях индивидуумов на выживаемость, в испытаниях технических устройств на надежность может возникнуть ситуация, когда испытываемые объекты попадают под наблюдение по истечении некоторого случайного времени после начала испытаний. Такое явление называют входом с задержкой (delayed entry), или цензурированием слева. При этом время жизни (или безотказной работы) может подвергаться также и правостороннему случайному цензурированию. Опишем соответствующую математическую модель.

Пусть X — с.в., означающая наработку объекта с функцией выживания $1 - F(x, \theta)$, зависящей от неизвестного параметра θ, и с плотностью $f(x, \theta)$, $\theta \in \Theta \subseteq R^1$. Предположим, что с.в. X подвергается случайному цензурированию с двух сторон с.в. L и Y с функциями распределения K и G и с плотностями k и g соответственно, которые не зависят от θ.

Пусть $\{(L_i, X_i, Y_i),\ i \ge 1\}$ — последовательность независимых реализаций вектора $(\mathbf{L}, \mathbf{X}, \mathbf{Y})$ с независимыми компонентами. Наблюдается выборка

$$\{\widetilde{Z}_i = (Z_i; \Delta_i^{(0)}, \Delta_i^{(1)}, \Delta_i^{(2)}), 1 \le i \le n\} = V^{(n)},$$

где $Z_i = \max\ (L_i, \min\ (X_i, Y_i)) = L_i \vee (X_i \wedge Y_i)$, $\Delta_i^{(0)} = I(X_i \wedge Y_i < L_i)$, $\Delta_i^{(1)} = I(L_i \le X_i \le Y_i)$, $\Delta_i^{(2)} = I(L_i \le Y_i < X_i)$. Обозначим $\widetilde{Z}^{(n)} = (\widetilde{Z}_1, \widetilde{Z}_2, ..., \widetilde{Z}_n)$ и пусть $\{\mathbf{Y}^{(n)}, \mathbf{U}^{(n)}, Q_\theta^{(n)}\}$ — последовательность статистических экспериментов, порожденная наблюдениями $\widetilde{Z}^{(n)}$, где $\mathbf{Y}^{(n)} = \{\mathbf{X} \times \{0,1\}^{(3)}\}^{(n)}$, $\mathbf{U}^{(n)} = \sigma(\mathbf{Y}^{(n)})$ и $Q_\theta^{(n)}$ — вероятностная мера на $(\mathbf{Y}^{(n)}, \mathbf{U}^{(n)})$ с одномерным распределением

$$Q_\theta(x, y^{(0)}, y^{(1)}, y^{(2)}) =$$

$$= P(Z_i < x, \Delta_i^{(0)} = y^{(0)}, \Delta_i^{(1)} = y^{(1)}, \Delta_i^{(2)} = y^{(2)});$$

$y^{(m)} \in \{0, 1\}$, $m = 0, 1, 2$. Пусть $\varepsilon_{y^{(m)}}$ — считающая мера, сосредоточенная в точке $y^{(m)}$ и $dv(\widetilde{z}_i) = \varepsilon_{y^{(m)}} \times dz_i$, $i = \overline{1, n}$. Тогда распределение $Q_\theta^{(n)}$ является абсолютно непрерывным относительно меры $v^{(n)}(\widetilde{z}^{(n)}) = v(\widetilde{z}_1) \times ... \times v(\widetilde{z}_n)$, ее плотность на $\mathbf{Y}^{(n)}$ при каждом $\theta \in \Theta$ определяется формулой

$$p_n(\widetilde{Z}^{(n)}; \theta) = \frac{dQ_\theta^{(n)}(\widetilde{Z}^{(n)})}{dv^{(n)}(\widetilde{z}(n))} =$$

$$= \prod_{m=1}^{n} k(Z_m)[1 - G(Z_m))(1 - F(Z_m; \theta))]^{\Delta_m^{(0)}} \times$$

$$\times K(Z_m)(1 - G(Z_m))f(Z_m; \theta)^{\Delta_m^{(1)}} \times$$

$$\times K(Z_m)g(Z_m)(1 - F(Z_m; \theta))^{\Delta_m^{(2)}}. \qquad (3.1)$$

При $u \in R^1$ определим «близкую альтернативу» $\theta + \frac{u}{\sqrt{n}} = \theta \in \Theta$ где θ_0 — истинное значение параметра θ, и зададимся логарифмом СОП $l_n(u) = \log \frac{dQ_{\theta_n}^{(n)}(\widetilde{Z}^{(n)})}{dQ_{\theta_0}^{(n)}(\widetilde{Z}^{(n)})}$.

Сформулируем условия регулярности, при справедливости которых будем иметь ЛАН для семейства распределений $\{\widetilde{Q}_\theta^{(n)}, \theta \in \Theta\}$:

1) носитель $N_f = \{x: f(x, \theta) > 0\}$ не зависит от θ;

2) для любых двух точек θ_1, $\theta_2 \in \Theta$, $\theta_1 \ne \theta_2$ и $x \in N_f$, $f(x, \theta_1) \ne f(x, \theta_2)$;

3) существуют и конечны для всех x производные $\partial^m f(x, \theta)/\partial\theta^m$, $m = 1, 2$; при этом

$$\int_{-\infty}^{\infty} \left| \frac{\partial^m f(x, \theta)}{\partial \theta^m} \right| \mathrm{d}x < \infty, \ m = 1, 2;$$

4) функция $\frac{\partial \log f(x, \theta_0)}{\partial \theta}$ является функцией ограниченной вариации;

5) информация Фишера

$$J(\theta) = \int_{-\infty}^{\infty} \left(\frac{\partial \log f(x, \theta)}{\partial \theta} \right)^2 \mathrm{d}T^{(1)}(x, \theta) +$$

$$+ \int_{-\infty}^{\infty} \left(\frac{\partial \log(1 - (1 - G(x))(1 - F(x, \theta)))}{\partial \theta} \right)^2 \mathrm{d}T^{(0)}(x, \theta) +$$

$$+ \int_{-\infty}^{\infty} \left(\frac{\partial \log(1 - F(x, \theta))}{\partial \theta} \right)^2 \mathrm{d}T^{(2)}(x, \theta)$$

является конечной и положительной в точке $\theta = \theta_0$.

В работе [15] при этих условиях регулярности установлено следующее представление при каждом $u \in R^1$:

$$\ln(u) = uJ^{1/2}(\theta_0)\zeta - \frac{u^2}{2} J(\theta_0) + R_n(u), \qquad (3.2)$$

где $\zeta^D = N(0, 1)$; $R_n(u) \to 0$ при $n \to \infty$ по $Q_{\theta_0}^{(n)}$-вероятности.

Пусть $\{\pi(u), u \in \Theta\}$ — неотрицательная измеримая функция и $l(d, \theta) = (d - \theta)^2$ — функция потерь на множестве $D \times \Theta$, где D — множество возможных оценок для θ. Рассмотрим оценки $\theta_n \in D$, определяемые соотношением

$$\hat{\theta}_n = \arg \min_{d \in D} \frac{\int_\Theta l(d, \theta) p_n(\widetilde{Z}^{(n)}, \theta) \pi(\theta) \mathrm{d}\theta}{\int_\Theta P_n(\widetilde{Z}^{(n)}, \theta) \pi(\theta) \mathrm{d}\theta}. \qquad (3.3)$$

Заметим, что если θ — с.в. с априорной плотностью π, то $\hat{\theta}_n$ является байесовской оценкой для θ. Покажем асимптотическую нормальность $\hat{\theta}_n$. Пусть $N(x, \theta) = 1 - (1 - \mathrm{G}(x))(1 - F(x, \theta)) = P_\theta(X_i \wedge Y_i < x)$. Определим субраспределения

$$T^{(0)}(x, \theta) = P_\theta(X_i \wedge Y_i < L_i; L_i < x),$$

$$T^{(1)}(x, \theta) = P(L_i \le X_i \le Y_i; X_i < x),$$

$$T^{(2)}(x, \theta) = P_\theta(L_i \le Y_i < X_i; Y_i < x),$$

для которых $T^{(0)}(x, \theta) + T^{(1)}(x, \theta) + T^{(2)}(x, \theta) = P_\theta(Z_i < x)$ для всех $(x, \theta) \in R^1 \times \Theta$.

Теорема 3.1 [14]. Пусть в условиях регулярности (1 – 4) $\pi(\theta_0) \neq 0$. Тогда при $n \to \infty$

$$L\left(\frac{\sqrt{n}(\theta_n - \theta_0)}{Q_{\theta_0}^{(n)}}\right) \to N(0, J^{-1}(\theta_0)).$$

Замечание. Из теоремы 3. 1 согласно определению Фишера [1, с. 127] оценку $\hat{\theta}_n$ можно назвать асимптотически эффективной.

Асимптотически минимаксная эффективность оценок максимального правдоподобия. Неполные наблюдения

Фишер [1, § 1.9] ввел понятие эффективной оценки для обозначения состоятельных, асимптотически нормальных оценок с асимптотически минимальной дисперсией. Этот подход подразумевал асимптотически наилучшую оценку, удовлетворяющую следующим двум условиям:

1) если θ_n ОМП для θ, то в условиях регулярности I – III

$$L(\sqrt{n}(\theta_n - \theta)) \underset{n\to\infty}{\to} N(0, I^{-1}(\theta));$$

2) если T_n — асимптотически нормальная последовательность оценок, то

$$\lim_{n\to\infty} M_\theta(\sqrt{n}(T_n - \theta))^2 \geq I^{-1}(\theta),\ \theta \in \Theta \subset R^1.$$

Однако [1, § 1.9] существуют и суперэффективные относительно квадратических функций потерь оценки и поэтому такое определение асимптотически наилучшей оценки является несостоятельной. Гаек [6] (см. также [1, с. 223]) установил асимптотически минимаксную границу рисков любых статистических оценок, если известны асимптотические свойства байесовских оценок. Он доказал этот результат с использованием свойства ЛАН для СОП. В работе [15] в общей модели конкурирующих рисков при случайном цензурировании с двух сторон в условиях регулярности установлено свойство ЛАН для СОП. Исследуем свойство асимптотической минимаксной эффективности ОМП. В частности, сформулируем теорему Гаека о минимаксной нижней границе качества различных оценок для широкого класса функций потерь при определенных условиях регулярности, в которой даются необходимое и достаточное условия достижимости этой границы. Затем эту теорему используем для установления асимптотической минимаксной эффективности ОМП.

Пусть функция потерь $w(u)$ удовлетворяет условиям:

$w(u)$ неотрицательная функция на R^1, $w(0) = 0$ и $w(u)$ непрерывна в точке $u = 0$;

функция $w(u)$ симметрична: $w(-u) = w(u)$;

множества $\{u\colon w(u) < c\}$ выпуклы для всех $c > 0$ и ограничены для достаточно малых $c > 0$;

$$\int_{R^1} w(u)\exp\left(-\frac{u^2}{2}\right)\mathrm{d}u < \infty.$$

Класс функций потерь, удовлетворяющих этим условиям, обозначим W.

Теорема 4.1 [1]. Пусть семейство $\{P_\theta^{(n)},\ \theta \in \Theta\}$ удовлетворяет условию ЛАН в точке $\theta_0 \in \Theta$ с нормировкой Γ_n, причем $\Gamma_n^2 \to 0$ при $n \to \infty$. Тогда для любого семейства оценок T_n и функции потерь $w \in W$ справедливо неравенство

$$\lim_{\delta\to 0}\liminf_{n\to\infty}\sup_{|\theta-\theta_0|<\delta} M_\theta\{w[\Gamma_n^{-1}(T_n - \theta)]\} \geq$$
$$\geq \frac{1}{\sqrt{2\pi}}\int_{R^1} w(u)e^{-u^2/2}\mathrm{d}u = Mw(\xi_0),$$

где $L(\xi_0) = n(0;\ 1)$. Кроме того, для любой непостоянной функции w нижняя граница

$$\lim_{\delta\to 0}\lim_{n\to\infty}\sup_{|\theta-\theta_0|<\delta} M_\theta\{w[\Gamma_n^{-1}(T_n - \theta)]\} = Mw(\xi_0) \quad (4.1)$$

достигается тогда и только тогда, когда при $n \to \infty$ разность

$$\Gamma_n^{-1}(T_n - \theta_0) - \Delta_{n,\theta_0} \to 0 \text{ по } P_{\theta_0}^{(n)}\text{-вероятности.} \quad (4.2)$$

Если (4.1) имеет место, то оценка T_n называется асимптотически минимаксно эффективной. Таким образом, (4.2) является необходимым и достаточным условием асимптотической минимаксной эффективности оценки T_n.

Рассмотрим теперь справедливость этой теоремы для ОМП для одномерного параметра θ в модели случайного цензурирования с двух сторон. Пусть $\hat{\theta}_n$ — ОМП, т.е. решение уравнения

$$\frac{\partial}{\partial\theta}\log p_n(\tilde{Z}^{(n)}, \theta) = 0, \quad (4.3)$$

где $p_n(\tilde{Z}^{(n)}, \theta)$ — функция правдоподобия, определенная в (3.1).

Следовательно, в условиях регулярности 1 – 5

$$\frac{1}{n}\frac{\partial\log p_n(\tilde{Z}^{(n)}, \theta)}{\partial\theta} = \int_{-\infty}^{\infty}\frac{\partial\log f(Z_m, \theta)}{\partial\theta}\mathrm{d}T_n^{(1)}(x) +$$
$$+\int_{-\infty}^{\infty}\frac{\partial\log(1-(1-G(Z_m))(1-F(Z_m, \theta)))}{\partial\theta}\mathrm{d}T_n^{(0)}(x) +$$
$$+\int_{-\infty}^{\infty}\frac{\partial\log(1-F(Z_m, \theta))}{\partial\theta}\mathrm{d}T_n^{(2)}(x). \quad (4.4)$$

Теорема 4.2 [17]. Пусть $\hat{\theta}_n$ — единственное решение уравнения (4.3) и имеют место условия регулярности 1 – 5. Тогда ОМП $\hat{\theta}_n$ является асимптотически минимаксной эффективной для параметра θ.

Замечание. Из теоремы 4.2 также следует, что

$$L\left(\frac{\sqrt{n}(\hat{\theta}_n - \theta_0)}{\widetilde{Q}^{(n)}_{\theta_0}}\right) \underset{n\to\infty}{\to} N(0; J^{-1}(\theta_0)),$$

т.е. ОМП и оценка байесовского типа (3.3) являются асимптотически эквивалентными.

Из вышеизложенного материала следует, что СОП играет важную роль при фиксированном объеме выборки согласно лемме Неймана – Пирсона для нахождения наиболее мощного критерия, а также при растущем объеме выборки для установления оптимальных свойств ОМП и оценок байесовского типа как в случае полных, так и неполных наблюдений.

ЛИТЕРАТУРА

1. **Ибрагимов И. А., Хасьминский Р. З.** Асимптотическая теория оценивания. — М.: Наука, 1979. — 527 с.
2. **Леман Э.** Проверка статистических гипотез. — М.: Наука, 1979. — 408 с.
3. **Леман Э.** Теория точечного оценивания. — М.: Наука, 1991. — 444 с.
4. **Русас Дж.** Контигуальность вероятностных мер. — М.: Мир, 1975. — 254 с.
5. **Hajek J.** A characterization of limiting distributions of regular estimates / Z. Wahrscheinlichkeits theorie und Verw. Gebiete. 1970. Vol. 14. P. 323 – 330.
6. **Hajek J.** Local asymptotic minimax and admissibility in estimation / Proc. Sixth. Berkeley Symp. on Math. Statist. and Prob. 1972. Vol. 1. P. 175 – 194.
7. **Le Cam L.** On some asymptotic properties of the maximum likelihood estimates and related Bayes estimates / Univ. California Publ. Statist. 1953. Vol. 1. P. 277 – 330.
8. **Le Cam L.** Locally asymptotically normal families of distributions / Univ. Calif. Publ. Statist. 1960. Vol. 3. P. 37 – 98.
9. **Le Cam L.** On the assumptions used to prove asymptotic normality of maximum likelihood estimates / Ann. Math. Statist. 1970. Vol. 41. N 3. P. 802 – 828.
10. **Van der Vaart A. W.** Asymptotic Statistics. — Cambridge Univ. Press, 1998. — 443 p.
11. **Абдушукуров А. А.** Статистика неполных наблюдений. — Ташкент: Университет, 2009. — 296 с.
12. **Абдушукуров А. А., Нурмухамедова Н. С.** Локальная асимптотическая нормальность в модели конкурирующих рисков / Узб. матем. ж-л. 2012. № 2. С. 5 – 12.
13. **Нурмухамедова Н. С.** Результаты аппроксимаций для статистик отношения правдоподобия в обобщенной модели случайного цензурирования с двух сторон / Узб. матем. ж-л. 2012. № 1. С. 95 – 106.
14. **Нурмухамедова Н. С.** Локальная асимптотическая нормальность и ее роль при оценивании неизвестного параметра / Вести НУУз. 2012. № 1. С. 166 – 169.
15. **Abdushukurov A. A., Nurmukhamedova N. S.** Local approximate normality of likelihood ratio statistics in competing risks model under random censorship from both sides / Far East J. Theor. Stat. 2013. Vol. 42. N 2. P. 107 – 122.
16. **Abdushukurov A. A., Nurmukhamedova N. S.** Asymptotics of the generalized statistics for testing the hypothesis under random censoring / Int. J. Res. Rev. Applied Sci. 2012. Vol. 13. Issue 2. P. 567 – 573.
17. **Abdushukurov A. A., Nurmukhamedova N. S.** Asymptotic minimax efficiency of maximum likelihood estimates in competing risks model / Acta NUUz. 2014. N 1. P. 3 – 8.

REFERENCES

1. **Ibragimov I. A., Has'minskii R. Z.** Asimptoticheskaya teoriya otsenivaniya [Asymptotic Theory of Estimation]. — Moscow: Nauka, 1979. — 527 p.
2. **Lehmann E.** Proverka statisticheskih gipotez. [Testing Statistical Hypotheses]. — Moscow: Nauka, 1979. — 408 p. [Russian translation].
3. **Lehmann E.** Teoriya tochechnogo osenivaniya. [Theory of Point Estimation]. — Moscow: Nauka, 1991. — 444 p. [Russian translation].
4. **Roussas G.** Contiguity of probability measures. — Moscow: Mir, 1975. — 254 p. [Russian translation].
5. **Hajek J.** A characterization of limiting distributions of regular estimates / Z. Wahrscheinlichkeits theorie und Verw. Gebiete. 1970. Vol. 14. P. 323 – 330.
6. **Hajek J.** Local asymptotic minimax and admissibility in estimation / Proc. Sixth. Berkeley Symp. on Math. Statist. and Prob. 1972. Vol. 1. P. 175 – 194.
7. **Le Cam L.** On some asymptotic properties of the maximum likelihood estimates and related Bayes estimates / Univ. California Publ. Statist. 1953. Vol. 1. P. 277 – 330.
8. **Le Cam L.** Locally asymptotically normal families of distributions / Univ. Calif. Publ. Statist. 1960. Vol. 3. P. 37 – 98.
9. **Le Cam L.** On the assumptions used to prove asymptotic normality of maximum likelihood estimates / Ann. Math. Statist. 1970. Vol. 41. N 3. P. 802 – 828.
10. **Van der Vaart A. W.** Asymptotic Statistics. — Cambridge Univ. Press, 1998. — 443 p.
11. **Abdushukurov A. A.** Statistika nepolnykh nablyudenii. [Statistics incomplete observations]. — Tashkent: University, 2009. — 296 p. [in Russian].
12. **Abdushukurov A. A., Nurmukhamedova N. S.** Lokal'naya asimptoticheskaya normal'nost' v modeli konkuriruyushchikh riskov [Locally asymptotically normality in competing risks model] / Uzb. Matem. J. 2012. N 2. P. 5 – 12 [in Russian].
13. **Nurmukhamedova N. S.** Rezul'taty approksimatsii dlya statistik otnosheniya pravdopodobiya v obobshchennoi modeli sluchainogo tsenzurirovaniya s dvukh storon [Approximation results for likelihood ratio statistics in generalized model of random censorship from both sides] / Uzb. Matem. J. 2012. N 1. P. 95 – 106 [in Russian].
14. **Nurmukhamedova N. S.** Lokal'naya asimptoticheskaya normal'nost' i ee rol' pri otsenivanii neizvestnogo parametra [Locally asymptotically normality and its role in estimation of unknown parameter] / Vesti NUUz. 2012. N 1. P. 166 – 169 [in Russian].
15. **Abdushukurov A. A., Nurmukhamedova N. S.** Local approximate normality of likelihood ratio statistics in competing risks model under random censorship from both sides / Far East J. Theor. Stat. 2013. Vol. 42. N 2. P. 107 – 122.
16. **Abdushukurov A. A., Nurmukhamedova N. S.** Asymptotics of the generalized statistics for testing the hypothesis under random censoring / Int. J. Res. Rev. Applied Sci. 2012. Vol. 13. Issue 2. P. 567 – 573.
17. **Abdushukurov A. A., Nurmukhamedova N. S.** Asymptotic minimax efficiency of maximum likelihood estimates in competing risks model / Acta NUUz. 2014. N 1. P. 3 – 8.