А. Л. РУХИН, *мл. науч. сотр.*
МИ АН СССР, Ленинград

О ДОПУСТИМОСТИ ОЦЕНОК МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Пусть $(X, \mathfrak{A}, P_\theta)$ — измеримое пространство с семейством вероятностных мер на нем, где $\theta \in \Theta \subset R_p$. Пусть Π обозначает класс всех оценок параметра θ, распределение которых имеет непрерывную плотность, т. е. таких, что мера

$$Q_\theta^f(A) = P_\theta(f(x) \in A),\ f \in \Pi$$

абсолютно непрерывна и

$$Q_\theta^f(A) = \int_A q_\theta^f(t)\,dt$$

с непрерывной функцией q_θ^f. Введем в классе Π следующую меру качества. Будем говорить, что оценка f не хуже чем g в смысле максимального правдоподобия, если при всех $\theta \in \Theta$

$$q_\theta^f(\theta) \geqslant q_\theta^g(\theta). \tag{1}$$

Другими словами, мы считаем, что мера качества оценки, связанная со степенью вырожденности ее распределения в точке истинного значения параметра, характеризуется значением плотности указанного распределения в этой точке.

Отметим также, что определение качества (1) можно несколько вольно формулировать и на обычном для математической статистики языке функций потерь, если в качестве последних допускать обобщенные функции.

Действительно, пусть функция потерь $W(f, \theta)$ зависит лишь от разности $f - \theta$. Тогда

$$\int_X W(f(x) - \theta)\,dP_\theta(x) = \int_\Theta W(t - \theta)\,dQ_\theta^f(t) = \int_\Theta W(t - \theta)\,q_\theta^f(t)\,dt. \tag{2}$$

Если в (2) положить $W(t) = -\delta(t)$, где δ — обобщенная функция Дирака, то мера качества, порождаемая такой функцией потерь, совпадает с (1).

Нетрудно также видеть, что байесовскими оценками θ, отвечающими равномерному распределению на Θ, являются обычные оценки максимального правдоподобия.

Займемся оцениванием в смысле максимального правдоподобия вещественного параметра сдвига $\theta \in R_1$ при $X = R_n$ из семейства $\{P_\theta : P_\theta(A) = P(A - \theta),\ A \in \mathfrak{A}\}$ по данным повторной выборки $x_1, \ldots, x_n$ из указанной совокупности. Предположим, что мера P имеет непрерывную плотность π и оценка максимального правдоподобия $\widehat{\theta} = \widehat{\theta}(x_1, \ldots, x_n)$ определяется из соотношения

$$\prod_1^n \pi(x_j - \widehat{\theta}) = \sup_{\theta \in R_1} \prod_1^n \pi(x_j - \theta) \tag{3}$$

почти для всех $(x_1, \ldots, x_n) \in R_n$ единственным образом. Пусть $x = \widehat{\theta}(x_1, \ldots, x_n)$, $y_2 = x_2 - x_1, \ldots, y_n = x_n - x_1$. Тогда плотность совместного распределения x и $y = (y_2, \ldots, y_n)$ относительно произведения лебеговой меры на R_1 и меры ν, порождаемой распределением y на R_{n-1}, имеет вид $p(x, y)$ причем в силу (3) для всех $t \in R_1$ и почти всех $y \in R_{n-1}$

$$p(0, y) \geqslant p(t, y).$$

Теорема. Если оценка максимального правдоподобия $\widehat{\theta}$ параметра сдвига θ такова, что плотность распределения $\widehat{\theta}$ и $yp(t, y)$ непрерывна и одновершинна по t при почти всех $y \in R$, то θ абсолютно допустима в смысле (1) в классе Π.

Доказательство. Введем последовательность функций выигрыша

$$W_\varepsilon(x) = \begin{cases} \dfrac{1}{2\varepsilon}, & |x| \leqslant \varepsilon \\ 0, & |x| > \varepsilon \end{cases}$$

и заметим, что

$$\iint W_\varepsilon(x - \theta)\, p(x - \theta, y)\, dx\, d\nu(y) \underset{\varepsilon \to 0}{\to} \int p(0, y)\, d\nu(y).$$

Установим сначала, что байесовская оценка, отвечающая равномерному распределению на R_1, абсолютно почти допустима в классе всех оценок θ. Действительно, эта оценка $f_\varepsilon(x, y) = x + \delta_\varepsilon(y)$ определяется из соотношения

$$\int W_\varepsilon(f_\varepsilon(x, y) - \theta)\, p(x - \theta, y)\, d\theta = \sup_{t \in R_1} \int W_\varepsilon(t - \theta)\, p(x - \theta, y)\, d\theta =$$

$$= \frac{1}{2\varepsilon} \sup_{t \in R_1} \int_{t-\varepsilon}^{t+\varepsilon} p(x - \theta, y)\, d\theta$$

и

$$p(x - f_\varepsilon(x, y) - \varepsilon, y) = p(x - f_\varepsilon(x, y) + \varepsilon, y)$$

или

$$p(-\delta_\varepsilon - \varepsilon, y) = p(-\delta_\varepsilon + \varepsilon, y). \tag{4}$$

В силу наших предположений δ_ε находится из (4) однозначным образом и из результатов работы [1] следует, что $f_\varepsilon(x, y)$ абсолютно почти допустима. Другими словами, если $g\ (x, y)$ — произвольная оценка параметра θ, то неравенство

$$\iint W_\varepsilon(f_\varepsilon(x, y) - \theta)\, p(x - \theta, y)\, dx d\nu(y) <$$
$$< \iint W_\varepsilon(g(x, y) - \theta)\, p(x - \theta, y)\, dx d\nu(y)$$

не может иметь места при θ из множества положительной лебеговой меры.

Покажем теперь, что

$$\iint W_\varepsilon(f_\varepsilon(x, y) - \theta)\, p(x - \theta, y)\, dx d\nu(y) =$$
$$= \iint W_\varepsilon(\delta_\varepsilon(y) + x)\, p(x, y)\, dx d\nu(y) \underset{\varepsilon \to 0}{\to} \int p(0, y)\, d\nu(y).$$

Действительно,

$$I_\varepsilon = \iint W_\varepsilon(\delta_\varepsilon(y) + x)\, p(x, y)\, dx d\nu(y) - \iint W_\varepsilon(x)\, p(x, y)\, dx d\nu(y) =$$
$$= \frac{1}{2\varepsilon}\left\{ \int\limits_{\{\delta_\varepsilon > 0\}} d\nu(y) \left[\int\limits_{-\varepsilon-\delta_\varepsilon}^{-\varepsilon} p(x, y)\, dx - \int\limits_{\varepsilon-\delta_\varepsilon}^{\varepsilon} p(x, y)\, dx \right] + \right.$$
$$\left. + \int\limits_{\{\delta_\varepsilon < 0\}} d\nu(y) \left[\int\limits_{\varepsilon}^{\varepsilon-\delta_\varepsilon} p(x, y)\, dx - \int\limits_{-\varepsilon}^{-\varepsilon-\delta_\varepsilon} p(x, y)\, dx \right] \right\}.$$

В силу одновершинности функции $p\ (x, y)$ и условия (4) выводим для $\delta_\varepsilon > 0$

$$0 \leqslant \int\limits_{-\varepsilon-\delta_\varepsilon}^{-\varepsilon} p(x, y)\, dx - \int\limits_{\varepsilon-\delta_\varepsilon}^{\varepsilon} p(x, y)\, dx \leqslant [p(-\varepsilon, y) - p(\varepsilon, y)]\, \delta_\varepsilon$$

и для $\delta_\varepsilon < 0$

$$0 \leqslant \int\limits_{\varepsilon}^{\varepsilon-\delta_\varepsilon} p(x, y)\, dx - \int\limits_{\varepsilon}^{-\varepsilon-\delta_\varepsilon} p(x, y)\, dx \leqslant [p(\varepsilon, y) - p(-\varepsilon, y)]\, |\delta_\varepsilon|.$$

Итак,

$$I_\varepsilon \leqslant \frac{1}{2\varepsilon} \int |\delta_\varepsilon(y)|\, |p(\varepsilon, y) - p(-\varepsilon, y)|\, d\nu(y) \leqslant$$
$$\leqslant \frac{1}{2} \int |p(\varepsilon, y) - p(-\varepsilon, y)|\, d\nu(y) \underset{\varepsilon \to 0}{\to} 0.$$

Так как

$$\iint W_\varepsilon(x)\, p(x, y)\, dx d\nu(y) \underset{\varepsilon \to 0}{\to} \int p(0, y)\, d\nu(y),$$

а

$$\iint W_\varepsilon(g(x, y) - \theta)\, p(x - \theta, y)\, dx d\nu(y)$$

стремится при $\varepsilon \to 0$ к значению плотности распределения g в точке θ, то в (1) при $f = x$ имеет место знак равенства при почти всех θ. Из предположенной непрерывности указанных плотностей следует, что в (1) имеет место знак равенства при всех θ, и оценка $\hat{\theta}$, таким образом, абсолютно допустима в смысле максимального правдоподобия.

Замечание 1. Функции выигрыша W_ε, введенные в доказательстве теоремы, отвечают доверительному оцениванию параметра θ с помощью интервалов постоянной длины 2ε. Нетрудно показать, что при $\varepsilon \to 0$ соответствующие байесовские оценки сходятся к оценке максимального правдоподобия. Таким образом, мера качества, задаваемая (1), может рассматриваться как соответствующая предельному случаю доверительного оценивания.

Замечание 2. Доказанная теорема неверна в случае оценивания многомерного параметра сдвига. Так, если $p \geqslant 3$ и

$$\pi(x_1 - \theta_1, \ldots, x_p - \theta_p) = \frac{1}{(2\pi)^{\frac{p}{2}}} \exp\left\{-\frac{1}{2}\sum_{j=1}^{p}(x_j - \theta_j)^2\right\}$$

плотность p-мерного нормального закона с единичной ковариационной матрицей, то оценка максимального правдоподобия $\hat{\theta}(x) = (\bar{x}_1, \ldots, \bar{x}_p)$, имеющая риск в смысле (1), равный $(n/2\pi)^{\frac{p}{2}}$, недопустима. Действительно, предположим для простоты, что объем выборки $n = 1$ (к этому случаю сводится общий за счет перехода к достаточным статистикам) и рассмотрим оценку параметра θ, предложенную Ч. Стейном [2, 3],

$$\hat{\theta}_1(x) = \left(x_1\left(1 - \frac{a^2}{|x|^2}\right), \ldots, x_p\left(1 - \frac{a^2}{|x|^2}\right)\right),$$

где $a^2 = p - 2$, $|x|^2 = \sum_{j=1}^{p} x_j^2$.

Вычислим плотность ее распределения. Пусть

$$y_j = x_j\left(1 - \frac{a^2}{|x|^2}\right), \qquad j = 1, \ldots, p.$$

Тогда в случае $|x| > a$

$$|y| = \frac{|x|^2 - a^2}{|x|}, \quad |x| = \frac{1}{2}|y| + \sqrt{\frac{1}{4}|y|^2 + a^2},$$

$$x_j = y_j\left(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{a^2}{|y|^2}}\right).$$

Если же имеет место случай $|x| < a$, то

$$|y| = \frac{a^2 - |x|^2}{|x|},$$

$$|x| = -\frac{|y|}{2} + \sqrt{\frac{1}{4}|y|^2 + a^2},$$

$$x_j = y_j\left(\frac{1}{2} - \sqrt{\frac{1}{4} + \frac{a^2}{|y|^2}}\right).$$

Вычисляя соответствующие якобианы, видим, что

$$I_1 = \left(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{a^2}{|y|^2}}\right)^{p-1}\left(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{a^2}{|y|^2}} - \right.$$

$$\left. - \frac{a^2}{|y|^2}\left(\frac{1}{4} + \frac{a^2}{|y|^2}\right)^{-\frac{1}{2}}\right),$$

$$I_2 = \left(-\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{a^2}{|y|^2}}\right)^{p-1}\left(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{a^2}{|y|^2}} + \right.$$

$$\left. + \frac{a^2}{|y|^2}\left(\frac{1}{4} + \frac{a^2}{|y|^2}\right)^{-\frac{1}{2}}\right). \qquad (5)$$

Итак, плотность распределения оценки $\widehat{\theta}_1(x) = (y_1, \ldots, y_p)$ равна

$$q(y_1, \ldots, y_p;\ \theta_1, \ldots, \theta_p) =$$

$$= (2\pi)^{-\frac{p}{2}}\left[\exp\left\{-\frac{1}{2}\sum_{j=1}^{p}\left[y_j\left(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{a^2}{|y|^2}}\right) - \theta_j\right]^2\right\}|I_1| + \right.$$

$$\left. + \exp\left\{-\frac{1}{2}\sum_{j=1}^{p}\left[y_j\left(\frac{1}{2} - \sqrt{\frac{1}{4} + \frac{a^2}{|y|^2}}\right) - \theta_j\right]^2\right\}|I_2|\right], \qquad (6)$$

где I_1, I_2 определены в (5).

Для нахождения риска, соответствующего этой оценке, полагаем в (6) $y_j = \theta_j$, $j = 1, \ldots, p$. Тогда

$$q(\theta_1, \ldots, \theta_p;\ \theta_1, \ldots, \theta_p) =$$

$$= (2\pi)^{-\frac{p}{2}}\left[\exp\left\{-\frac{1}{2}|\theta|^2\left(-\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{a^2}{|\theta|^2}}\right)^2\right\}|I_1| + \right.$$

$$\left. + \exp\left\{-\frac{1}{2}|\theta|^2\left(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{a^2}{|\theta|^2}}\right)^2\right\}|I_2|\right]. \qquad (7)$$

Пусть

$$b = b(\theta) = \sqrt{\frac{1}{4} + \frac{a^2}{|\theta|^2}}.$$

Тогда $|\theta|^2 = \dfrac{4a^2}{4b^2 - 1}$ и величина (7), если учесть (5), равна

$$\frac{1}{(\sqrt{2\pi})^p}\left[\left(\frac{1}{2} + b\right)^{p-1}\left(\frac{1}{2} + b - \frac{4b^2 - 1}{4b}\right) \times \right.$$

$$\times \exp\left\{-\frac{1}{2}\,\frac{4a^2}{4b^2 - 1}\left(b - \frac{1}{2}\right)^2\right\} +$$

$$\left. + \left(b - \frac{1}{2}\right)^{p-1}\left(\frac{1}{2} - b + \frac{4b^2 - 1}{4b}\right)\exp\left\{-\frac{1}{2}\,\frac{4a^2}{4b^2 - 1}\left(b + \frac{1}{2}\right)^2\right\}\right] =$$

$$= \frac{1}{2b(\sqrt{2\pi})^p}\left[\left(b+\frac{1}{2}\right)^p \exp\left\{-\frac{a^2}{2}\frac{b-\frac{1}{2}}{b+\frac{1}{2}}\right\} + \right.$$

$$\left. + \left(b-\frac{1}{2}\right)^p \exp\left\{-\frac{a^2}{2}\frac{b+\frac{1}{2}}{b-\frac{1}{2}}\right\}\right].$$

Таким образом, факт недопустимости оценки максимального правдоподобия следует из справедливости неравенства

$$\left(b+\frac{1}{2}\right)^p \exp\left\{-\frac{a^2}{2}\frac{b-\frac{1}{2}}{b+\frac{1}{2}}\right\} + \left(b-\frac{1}{2}\right)^p \exp\left\{-\frac{a^2}{2}\frac{b+\frac{1}{2}}{b-\frac{1}{2}}\right\} \geqslant 2b,$$

$b \geqslant \frac{1}{2}$, $a^2 = p - 1$, доказательство которого элементарно.

ЛИТЕРАТУРА

1. Farrell R. H. Estimators of a location parameter in the absolutely continuous case.— Ann. Math. Statist., 35, 3, 1964.

2. James W., Stein C. Estimation with quadratic loss function.— Pros. Fourth Berkeley Symposium on Math. Statist. and Probability, 1, 361—379, 1961.

3. Stein C. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution.— Proc. Third Berkeley Symposium on Math. Statist. and Probability, 1, 197—206, 1956.

A. L. Rukhin

ON THE ADMISSIBILITY OF MAXIMUM LIKELIHOOD ESTIMATES

Summary

The admissibility in some sense of maximum likelihood estimates of one dimensional location parameter is proved. Contradictious example for the case of multidimensional parameter is constructed.

Поступила в редколлегию 2.VI. 1969.