

Ш. ЛУВСАНЦЕРЕН

ОЦЕНКИ НАИБОЛЬШЕГО ПРАВДОПОДОБИЯ И ДОВЕРИТЕЛЬНЫЕ МНОЖЕСТВА ДЛЯ НЕИЗВЕСТНЫХ ПАРАМЕТРОВ СТАЦИОНАРНОГО ГАУССОВСКОГО ПРОЦЕССА МАРКОВСКОГО ТИПА

(Представлено академиком А. Н. Колмогоровым 7 VII 1954)

Предварительно сформулируем две общих теоремы. Первая из них является перенесением на более общий случай известной теоремы об асимптотическом поведении оценок наибольшего правдоподобия, построенных по независимым наблюдениям (см. (1), стр. 544, и (5)).

Теорема 1. Рассмотрим последовательность пространств элементарных событий $\Omega^{(n)}$ ($n = 1, 2, \dots$). Пусть в $\Omega^{(n)}$ задано семейство распределений вероятностей $P_\theta^{(n)}(A)^*$, зависящее от параметра θ из некоторого числового интервала I . Предположим, что $P_\theta^{(n)}(A)$ задается плотностью $p_n(\omega; \theta)$ ($\omega \in \Omega^{(n)}$, $\theta \in I$) относительно некоторой меры $\mu^{(n)}$ в $\Omega^{(n)}$. Предположим далее, что для последовательности функций $L_n(\omega; \theta) = \partial \ln p_n(\omega; \theta) / \partial \theta$ и $C_n(\theta) = D L_n(\omega; \theta)$ для всякого $\varepsilon > 0$ при $n \rightarrow \infty$ равномерно по $\theta \in I$ выполнены условия:

$$а) P_\theta^{(n)} \left\{ \left| \frac{L_n(\omega; \theta)}{C_n(\theta)} \right| > \varepsilon \right\} \rightarrow 0; \quad б) P_\theta^{(n)} \left\{ \left| \frac{\partial L_n(\omega; \theta)}{\partial \theta} C_n^{-1}(\theta) + 1 \right| > \varepsilon \right\} \rightarrow 0;$$

в) для любого $\delta > 0$ существует постоянная K , не зависящая от θ и n такая, что

$$P_\theta^{(n)} \left\{ \left| \left[\frac{\partial^2 L_n(\omega; \theta)}{\partial \theta^2} \right]_{\theta=\tilde{\theta}} C_n^{-1}(\tilde{\theta}) \right| > K \right\} < \delta \quad \text{для } \theta, \tilde{\theta} \in I, n = 1, 2, \dots$$

Тогда:

I. С вероятностью, равномерно по $\theta \in I$ стремящейся к 1, уравнение наибольшего правдоподобия

$$L_n(\omega; \theta) = 0 \tag{1}$$

имеет корень $\theta_n(\omega)$ такой, что для любого $\varepsilon > 0$ при $n \rightarrow \infty$ равномерно относительно $\theta \in I$

$$P_\theta^{(n)} \{ |\theta_n(\omega) - \theta| > \varepsilon \} \rightarrow 0^{**}. \tag{2}$$

II. Если $\theta_n(\omega)$ — корень, удовлетворяющий соотношению (2), то

$$\lim_{n \rightarrow \infty} P_\theta^{(n)} \{ p_n(\omega; \theta_n(\omega)) = \max \} = 1.$$

III. Решение $\theta_n(\omega)$, удовлетворяющее (2), единственно в следующем смысле: если $\theta_n^{(1)}(\omega)$ и $\theta_n^{(2)}(\omega)$ — два решения уравнения (1), удов-

* A — элемент некоторого борелевского тела подмножеств пространства $\Omega^{(n)}$.

** Корень $\theta_n(\omega)$, удовлетворяющий условию (2), называется равномерно состоятельной оценкой для θ .

летворяющие условию (2), то при $n \rightarrow \infty$ равномерно по $\theta \in I$ $P_0^{(n)}\{\theta_n^{(1)}(\omega) \neq \theta_n^{(2)}(\omega)\} \rightarrow 0$.

IV. Если при $n \rightarrow \infty$ равномерно относительно $\theta \in I$ имеем $P_0^{(n)}\{C_n^{-1/2}(\theta)L_n(\omega; \theta) < x\} \rightarrow F(x)$ * ($F(x)$ — некоторая функция распределения), то при $n \rightarrow \infty$ равномерно по $\theta \in I$ $P_0^{(n)}\{C_n^{1/2}(\theta)(\theta_n(\omega) - \theta) < x\} \rightarrow F(x)$.

Во всех известных до настоящего времени случаях функция распределения $F(x)$, фигурирующая в теореме 1, является нормальной функцией распределения, и поэтому для больших n по заданному α ($0 < \alpha < 1$) можно найти число x_α такое, что имеет место приближенное равенство

$$P_0^{(n)}\{C_n^{-1/2}(\theta)|L_n(\omega; \theta)| < x_\alpha\} \sim 1 - 2\alpha^{**}, \text{ где } \alpha = \frac{1}{V2\pi} \int_{x_\alpha}^{\infty} e^{-t^2/2} dt.$$

Обозначим через $M_\alpha(\omega)$ и $\tilde{M}_\alpha(\omega)$ доверительные множества для θ , определяемые, соответственно, неравенствами:

$$C_n^{1/2}(\theta)|\theta_n - \theta| \leq x_\alpha, \quad |\tilde{L}_n(\omega; \theta)| = C_n^{-1/2}(\theta)|L_n(\omega; \theta)| \leq x_\alpha.$$

Предположим, что выполнено следующее условие: для всякого $\varepsilon > 0$ при $n \rightarrow \infty$ равномерно относительно $\theta \in I$

$$P_0^{(n)}\left\{\left|\sqrt{\frac{C_n(\theta_n)}{C_n(\theta)}} - 1\right| > \varepsilon\right\} \rightarrow 0. \quad (B)$$

Тогда $P_0^{(n)}\{C_n^{1/2}(\theta_n)|\theta_n - \theta| < x_\alpha\} \sim 1 - 2\alpha$, и мы можем рассматривать еще одно доверительное множество $\tilde{\tilde{M}}_\alpha(\omega)$, определяемое неравенством $C_n^{1/2}(\theta_n)|\theta_n - \theta| \leq x_\alpha$.

Известно, что доверительное множество $\tilde{M}_\alpha(\omega)$ в случае независимых наблюдений обладает определенными оптимальными свойствами⁽²⁾. Не такие доверительные множества часто бывают очень сложны. Напротив, доверительное множество $\tilde{\tilde{M}}_\alpha(\omega)$ является наиболее простым, ибо оно представляет собой просто интервал. Теорема 2 показывает, что при широких условиях все три типа доверительных множеств асимптотически эквивалентны.

Теорема 2. Пусть $C_n(\theta) < \infty$ всюду, кроме, быть может, конечного числа точек и пусть функция распределения оценки θ_n сходится к нормальной равномерно по $\theta \in I$ при $n \rightarrow \infty$. Предположим, что выполнено условие (B). Тогда доверительные множества $M_\alpha(\omega)$, $\tilde{M}_\alpha(\omega)$ и $\tilde{\tilde{M}}_\alpha(\omega)$ асимптотически эквивалентны в том смысле, что попарные отношения их лебеговских мер при $n \rightarrow \infty$ сходятся по вероятности к 1 равномерно по $\theta \in I$.

Отметим, что теоремы 1 и 2 распространяются и на случай нескольких неизвестных параметров.

Переходим к изучению марковского гауссовского стационарного процесса

$$\xi_1, \xi_2, \dots, \xi_n, \dots \quad (3)$$

Такой процесс определяется следующими тремя параметрами: $a = M\xi_k = M\xi_{k+1}$, $\sigma^2 = D\xi_k = D\xi_{k+1}$, $\lambda = R(\xi_k, \xi_{k+1})$, где символы M , D

* Здесь и всюду в дальнейшем, говоря о сходимости функций распределения, мы подразумеваем так называемую слабую сходимость, т. е. сходимость в каждой точке непрерывности предельной функции распределения.

** Символ $\sim$ означает, что отношение левой части к правой при $n \rightarrow \infty$ сходится к 1 равномерно относительно $\theta \in I$.

и R обозначают, соответственно, математическое ожидание, дисперсию и коэффициент корреляции соответствующих величин, а k — любое целое положительное число. Вопрос об оценке параметров процесса (3)* решается нами для случая, когда σ^2 известна. Не ограничивая общности, можно принять $\sigma = 1$.

Плотность вероятности величин $\xi_1, \xi_2, \dots, \xi_n$ имеет вид

$$p_n(x_1, \dots, x_n; a, \sigma, \lambda) = \\ = (2\pi)^{-\frac{n}{2}} (1 - \lambda^2)^{-\frac{n-1}{2}} \sigma^{-n} \exp \left\{ -\frac{1}{2} \frac{1}{1 - \lambda^2} \sum_{i,j=1}^n q_{ij} \frac{(x_i - a)}{\sigma} \frac{(x_j - a)}{\sigma} \right\},$$

где $\|q_{ij}\| = \begin{vmatrix} 1 & -\lambda & 0 & \dots & 0 \\ -\lambda & 1 + \lambda^2 & & & \\ 0 & & & & \\ \vdots & & & & \\ 0 & \dots & 0 & 1 + \lambda^2 & -\lambda \\ & & & -\lambda & 1 \end{vmatrix}$

Лемма. Пусть $a = 0$ и $\sigma = 1$. Тогда уравнение наибольшего правдоподобия для λ .

$$\frac{\partial \ln p_n(x_1, \dots, x_n; \lambda)}{\partial \lambda} = 0$$

с вероятностью, стремящейся к 1, когда $n \rightarrow \infty$, имеет корень $\lambda_n = \lambda_n(x_1, \dots, x_n)$, который является равномерно состоятельной оценкой для λ .

Оценка λ_n выражается через достаточные статистики нашей задачи

$$T_1 = \frac{x_1 x_2 + \dots + x_{n-1} x_n}{n-1}, \quad T_2 = \frac{x_1^2 + 2x_2^2 + \dots + 2x_{n-1}^2 + x_n^2}{n-1},$$

а именно является корнем уравнения $\frac{\lambda - T_1}{T_2 - 2} = -\frac{\lambda}{1 + \lambda^2}$, для решения которого можно дать весьма простой графический метод.

В случае $a = 0$ и $\sigma = 1$ имеет место следующая теорема о предельном распределении оценки λ_n .

Теорема 3. а) Функция распределения $\frac{V(n-1)(1+\lambda^2)}{1-\lambda^2} (\lambda_n - \lambda)$ при $n \rightarrow \infty$ стремится к нормальной функции распределения $\Phi(x)$ с параметрами 0 и 1 равномерно относительно $\lambda \in (-1, 1)$.

б) Преобразованная оценка

$$\varphi(\lambda_n) = \int_0^{\lambda_n} \frac{V(1+\lambda^2)}{1-\lambda^2} d\lambda$$

асимптотически распределена по нормальному закону с параметрами $\varphi(\lambda)$ и $1/V(n-1)$, причем стремление к предельному распределению будет происходить равномерно по $\lambda \in (-1, 1)$.

Основываясь на теореме 3, можно построить доверительные интервалы для $\varphi(\lambda)$ и λ следующим образом:

$$P_\lambda^{(n)} \left\{ \varphi(\lambda_n) - x_\alpha \frac{1}{V(n-1)} < \varphi(\lambda) < \varphi(\lambda_n) + x_\alpha \frac{1}{V(n-1)} \right\} \sim 1 - 2\alpha;$$

$$P_\theta^{(n)} \left\{ \lambda_n - x_\alpha \frac{1 - \lambda_n^2}{V(n-1)(1 + \lambda_n^2)} < \lambda < \lambda_n + x_\alpha \frac{1 - \lambda_n^2}{V(n-1)(1 + \lambda_n^2)} \right\} \sim 1 - 2\alpha.$$

* Этот вопрос был поставлен А. Н. Колмогоровым в 1948 г. В работе (3) Ю. В. Линника построены доверительные множества для λ при четырех разных предположениях относительно известности a и σ . Доверительные множества Ю. В. Линника действуют для любых, в том числе малых, значений n . Но, в отличие от рассматриваемых нами доверительных множеств, они не являются оптимальными в каком-нибудь смысле.

Доверительный интервал, определяемый последним соотношением, является оптимальным в смысле Уилкса.

В случае, когда известны λ и σ^2 ($\sigma = 1$), уравнение наибольшего правдоподобия для a дает оценку:

$$a_n = a_n(x_1, \dots, x_n) = \frac{x_1 + x_n + (1-\lambda) \sum_{i=2}^{n-1} x_i}{2 + (1-\lambda)(n-2)}.$$

Оценка a_n распределена по нормальному закону с параметрами a и $\frac{V1+\lambda}{V2+(1-\lambda)(n-2)}$. Отметим, что a_n является состоятельной оценкой для a , но скорость ее сходимости к a сильно ухудшается при значениях λ , близких к 1.

В случае неизвестных a и λ (при известной $\sigma^2 = 1$) имеет место следующая теорема о совместном предельном распределении оценок наибольшего правдоподобия для a и λ .

Теорема 4. Система уравнений наибольшего правдоподобия

$$\frac{\partial \ln p_n(x_1, \dots, x_n; a, \lambda)}{\partial \lambda} = 0, \quad \frac{\partial \ln p_n(x_1, \dots, x_n; a, \lambda)}{\partial a} = 0$$

с вероятностью, стремящейся к 1, когда $n \rightarrow \infty$, имеет решение $\lambda_n(x_1, \dots, x_n)$, $a_n(x_1, \dots, x_n)$, такое, что при $n \rightarrow \infty$ равномерно в области $-1 < \lambda < 1$, $-\infty < a < \infty$

$$P_{a, \lambda}^{(n)} \left\{ \frac{V(n-1)(1+\lambda^2)}{1-\lambda^2} (\lambda_n - \lambda) < x, \frac{V(n-2)(1-\lambda)+2}{V1+\lambda} (a_n - a) < y \right\} \rightarrow \\ \rightarrow \frac{1}{2\pi} \int_{-\infty}^x \int_{-\infty}^y e^{-1/2(a^2+v^2)} dudv.$$

Итак, оценки a_n и λ_n оказываются асимптотически независимыми. Из теоремы 4 можно вывести, что

$$P_{a, \lambda}^{(n)} \left\{ \frac{(n-1)(1+\lambda_n^2)}{(1-\lambda_n^2)^2} (\lambda_n - \lambda)^2 + \frac{(n-2)(1-\lambda)+2}{1+\lambda_n} (a_n - a)^2 < \chi_\gamma^2 \right\} \sim 1 - \gamma,$$

где χ_γ^2 — корень уравнения $\gamma = e^{-\chi_\gamma^2/2}$ и $0 < \gamma < 1$.

В силу аналога теоремы 2 для случая нескольких неизвестных параметров можно утверждать, что задаваемое этой формулой доверительное множество для одновременной оценки параметров a и λ является оптимальным в смысле Уилкса.

В заключение пользуюсь случаем принести глубокую благодарность доктору физико-математических наук Е. Б. Дынкину за руководство при выполнении настоящей работы.

Поступило
7 VII 1954

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

- ¹ Г. Крамер, Математические методы статистики, М., ИЛ, 1948.
² S. S. Wilks, Ann. Math. Statist., 9, 166 (1938). ³ Ю. В. Линник, Изв. АН СССР, сер. матем., 14, № 6 (1950). ⁴ U. Grenander, Ark. f. Mat., 1, Н. 3 (1950).
⁵ J. Huzurbazar, Ann. Eugenics, 14, part 3 (1948).

* Эта формула является тривиальным видоизменением формулы Гренандера (⁴) для случая непрерывного марковского гауссовского стационарного процесса с известной корреляционной функцией.