УДК 519.21

О.О. Кубайчук

АСИМПТОТИЧНА ПОВЕДІНКА ОЦІНКИ МЕТОДУ МІНІМІЗАЦІЇ ЕМПІРИЧНОГО РИЗИКУ ДЛЯ БАЄСОВОГО ПОРОГА

Вступ

Задача класифікації об'єкта за спостереженням його числової характеристики має велике наукове та практичне значення. Елементарний приклад такої класифікації — визначення людини як хворої, якщо її температура перевищує 37°. Для розв'язання цієї задачі розглядаються порогові класифікатори вигляду

$$g_t(\xi)=\begin{cases}1 \text{ при } \xi\le t,\\ 2 \text{ при } \xi>t,\end{cases}$$

тобто об'єкт належить до першого класу, якщо його характеристика (ξ) не перевищує поріг t, і до другого класу — в іншому випадку. В наведеному прикладі таким порогом є температура 37°.

Зазвичай найкращим (баєсовим) вважається такий поріг $t=t^B$, при якому g_t має найменшу ймовірність помилки. Найбільш поширеними методами оцінювання t^B є емпірично баєсова класифікація (ЕБК) [1, 2] та метод мінімізації емпіричного ризику (МЕР) [3, 4]. Випадок, коли навчаюча вибірка отримана із суміші із змінними концентраціями, розглядається в [5]. Там же досліджується асимптотична поведінка цих двох методів для вибірки із суміші із змінними концентраціями.

Але часто трапляються ситуації, що призводять до розв'язання задач класифікації об'єкта за спостереженням його числової характеристики при наявності більше як одного порога (наприклад, два). Прикладом може бути визначення людини як хворої, якщо її температура перевищує 37° або є меншою за 36°. Інший випадок — людина хвора, якщо рівень гемоглобіну перевищує 84 одиниці або є меншим за 72 одиниці. Частково це питання розглядається в [6].

Постановка задачі

Метою даної статті є дослідження асимптотичної поведінки оцінок методу МЕР у випадку двох порогів для вибірки із суміші із змінними концентраціями.

Методика побудови оцінки для баєсового порога

Розглядається задача класифікації деякого об'єкта O за спостереженням його числової характеристики $\xi=\xi(O)$. Вважаємо, що об'єкт може належати лише одному з двох класів. Розглядаємо порогові класифікатори вигляду

$$g^1_{t_1,t_2}(\xi)=\begin{cases}1 \text{ при } \xi\in[t_1,t_2],\\ 2 \text{ при } \xi\notin[t_1,t_2],\end{cases} \quad (1)$$

$$g^2_{t_1,t_2}(\xi)=\begin{cases}1 \text{ при } \xi\notin[t_1,t_2],\\ 2 \text{ при } \xi\in[t_1,t_2].\end{cases} \quad (2)$$

В одному з наведених вище прикладів — визначення людини (об'єкта) як хворої — характеристикою ξ є її температура, порогом t_1 — 36°, порогом t_2 — 37° для порогового класифікатора вигляду (1).

Найкращим (баєсовим) вважають поріг $\bar t^B=(t_1^B,t_2^B)$, при якому $g^i_{t_1,t_2}$, $i=1,2$, має найменшу ймовірність помилки. Об'єкт O, в якого спостерігається деяка числова характеристика $\xi=\xi(O)$, може належати одному з двох класів; невідомий номер класу, якому належить O, позначимо $\mathrm{ind}(O)$. Вважаються відомими апріорні ймовірності $p_i=P(\mathrm{ind}(O)=i)$, $i=1,2$. Характеристика ξ — випадкова, її розподіл залежить від $\mathrm{ind}(O): P(\xi(O)<x \mid \mathrm{ind}(O)=i)=H_i(x)$. Розподіли H_i невідомі, але вважатимемо, що вони мають неперервні щільності відносно міри Лебега — h_i.

Множину всіх класифікаторів позначимо G. Ймовірність помилки класифікатора визначається формулою

$$L(g^1_{\bar t})=L^1(\bar t)=P\{g^1_{\bar t}(\xi(O))\ne \mathrm{ind}(O)\}=$$

$$=p_1(H_1(t_1)+1-H_1(t_2))+p_2(H_2(t_2)-H_2(t_1)),$$

$$L(g^2_{\bar t})=L^2(\bar t)=P\{g^2_{\bar t}(\xi(O))\ne \mathrm{ind}(O)\}=$$

$$=p_1(H_1(t_2)-H_1(t_1))+p_2(H_2(t_1)+1-H_2(t_2)).$$

Відповідно до позначень, введених у [6], маємо

$$L_1^1(t_1) = p_1 H_1(t_1) - p_2 H_2(t_1),$$

$$L_2^1(t_2) = p_1(1 - H_1(t_2)) + p_2 H_2(t_2),$$

$$L_1^2(t_1) = -p_1 H_1(t_1) + p_2 H_2(t_1),$$

$$L_2^2(t_2) = p_1 H_1(t_2) + p_2(1 - H_2(t_2)).$$

Розглянемо перший випадок. Функції H_i (а значить, і h_i) вважаються невідомими. Їх можна оцінити за даними, що являють собою вибірку із суміші із змінними концентраціями: $\{\xi_{j:N}\}_{j=1}^N$, $\xi_{j:N}$ – не залежні між собою при фіксованому N і $P\{\xi_{j:N} < x\} = w_{j:N} H_1(x) + (1 - w_{j:N}) H_2(x)$, де $w_{j:N}$ – відома концентрація об'єктів першого класу в суміші в момент j-го спостереження [7]. Для оцінки функції розподілу H_i використовують зважені емпіричні функції розподілу $\widehat{H}_i^N(x) = \frac{1}{N}\sum_{j=1}^N a_{j:N}^i 1\{\xi_j < x\}$, де $1\{A\}$ – індикатор події A; $a_{j:N}^i$ – вагові коефіцієнти [7]:

$$a_{j:N}^1 = \frac{1}{\Delta_N}((1 - S_N^1) w_{j:N} + (S_N^2 - S_N^1)),$$

$$a_{j:N}^2 = \frac{1}{\Delta_N}(S_N^2 - S_N^1 w_{j:N}), \tag{3}$$

$$S_N^k = \frac{1}{N}\sum_{j=1}^N (w_{j:N})^k, \ \Delta_N = S_N^2 - (S_N^1)^2.$$

Для оцінки щільностей розподілів h_i використовують ядерні оцінки

$$\widehat{h}_i^N(x) = \frac{1}{Nk_N}\sum_{j=1}^N a_{j:N}^i K\left(\frac{x - \xi_{j:N}}{k_N}\right),$$

де K – ядро; k_N – параметр згладжування [8, 9].

Побудуємо оцінку порога $\bar{t}^B$ методом МЕР [6]. Нехай $\widehat{\bar{t}}_N$ оцінка для $\bar{t}^B$. Крім того, нехай $L_N^1(t_1, t_2)$ є оцінкою для $L^1(t_1, t_2)$. Тоді маємо

$$L_N^1(t_1, t_2) =$$
$$= p_1(\widehat{H}_1^N(t_1) + 1 - \widehat{H}_1^N(t_2)) + p_2(\widehat{H}_2^N(t_2) - \widehat{H}_2^N(t_1)) =$$
$$= L_{N_1}^1(t_1) + L_{N_2}^1(t_2),$$

$$L_{N_1}^1(t_1) = p_1 \widehat{H}_1^N(t_1) - p_2 \widehat{H}_2^N(t_1),$$

$$L_{N_2}^1(t_2) = p_1(1 - \widehat{H}_1^N(t_2)) + p_2 \widehat{H}_2^N(t_2).$$

І аналогічно отримуємо

$$L_{N_1}^2(t_1) = -p_1 \widehat{H}_1^N(t_1) + p_2 \widehat{H}_2^N(t_1),$$

$$L_{N_2}^2(t_2) = p_1 \widehat{H}_1^N(t_2) + p_2(1 - \widehat{H}_2^N(t_2)).$$

Нехай щільності h_i існують і є неперервно диференційовними в деяких околах точок t_1^B і t_2^B.

Позначимо

$$r_{N_i} = \left[\frac{1}{N}\sum_{j=1}^N (p_2 a_{j:N}^2 - p_1 a_{j:N}^1)^2 \times \right.$$
$$\left. \times (w_{j:N} h_1(t_i^B) + (1 - w_{j:N}) h_2(t_i^B))\right]^{1/2}, \ i = 1, 2.$$

Нехай існує $\lim_{N\to\infty} r_{N_i} = r_i$, $i = 1, 2$ (умови існування границі розглянемо далі). Також позначимо

$$W_{N_i}^1(\tau) = N^{2/3}(L_{N_i}^1(t_i^B + N^{-1/3}\tau) -$$
$$- L_{N_i}^1(t_i^B) - L_i^1(t_i^B + N^{-1/3}\tau) + L_i^1(t_i^B)),$$

$$W_{N_i}^2(\tau) = N^{2/3}(L_{N_i}^2(t_i^B + N^{-1/3}\tau) -$$
$$- L_{N_i}^2(t_i^B) - L_i^2(t_i^B + N^{-1/3}\tau) + L_i^2(t_i^B)), \ i = 1, 2.$$

Основні результати

Будемо вважати, що виконуються такі умови:

A) $\bar{t}^B$ існує і є єдиною точкою глобального мінімуму $L^1(\bar{t})$ (t_i^B є точкою глобального мінімуму $L_i^1(t_i)$, $i = 1, 2$);

B_k) існують границі $S^i = \lim_{N\to\infty} S_N^i$, $i = 1, 2, \ldots, k$, і $\Delta = S^2 - (S^i)^2 > 0$. (Для випадку $L^2(\bar{t})$ формулювання умов аналогічне.)

Перед тим як перейти до вивчення асимптотичної поведінки процесів $W_{N_i}^1$ і $W_{N_i}^2$, $i = 1, 2$, доведемо кілька допоміжних тверджень.

Зауваження. Для існування $\lim_{N\to\infty} r_{N_i} = r_i$, $i = 1,2$, достатньо виконання умови B_3.

Лема 1. Мають місце рівності

$$\sum_{j=1}^{N} a_{j:N}^1 w_{j:N} = N,\ \sum_{j=1}^{N} a_{j:N}^1 = N,\ \sum_{j=1}^{N} a_{j:N}^2 = N,$$

$$\sum_{j=1}^{N} a_{j:N}^2 w_{j:N} = 0 .$$

Доведення. З (3) маємо

$$a_{j:N}^1 = \frac{1}{\Delta_N}((1-S_N^1)w_{j:N} + (S_N^2 - S_N^1)) .$$

Тоді

$$\sum_{j=1}^{N} a_{j:N}^1 w_{j:N} =$$

$$= \sum_{j=1}^{N} \frac{1}{\Delta_N}((1-S_N^1)w_{j:N} + (S_N^2 - S_N^1))w_{j:N} =$$

$$= \sum_{j=1}^{N} \frac{1}{\Delta_N}(w_{j:N}^2 - w_{j:N}^2 \frac{1}{N}\sum_{j=1}^{N} w_{j:N} +$$

$$+ w_{j:N}\frac{1}{N}\sum_{j=1}^{N} w_{j:N}^2 - w_{j:N}\frac{1}{N}\sum_{j=1}^{N} w_{j:N}) =$$

$$= \frac{1}{\Delta_N}\left[\sum_{j=1}^{N} w_{j:N}^2 - \frac{1}{N}\sum_{j=1}^{N} w_{j:N}\sum_{j=1}^{N} w_{j:N}\right].$$

Крім того, отримаємо

$$\sum_{j=1}^{N} a_{j:N}^2 = \sum_{j=1}^{N} \frac{1}{\Delta_N}(S_N^2 - S_N^1 w_{j:N}) =$$

$$= \frac{\frac{1}{N} N\sum_{j=1}^{N} w_{j:N}^2 - \frac{1}{N}\left(\sum_{j=1}^{N} w_{j:N}\right)^2}{\frac{1}{N}\sum_{j=1}^{N} w_{j:N}^2 - \frac{1}{N^2}\left(\sum_{j=1}^{N} w_{j:N}\right)^2} = N .$$

Аналогічно, маємо

$$\sum_{j=1}^{N} a_{j:N}^1 = \frac{1}{\Delta_N}(NS_N^1(1-S_N^1) + (S_N^2 - S_N^1)N) =$$

$$= \frac{N(S_N^2 - (S_N^1)^2)}{S_N^2 - (S_N^1)^2} = N,$$

$$\sum_{j=1}^{N} a_{j:N}^2 w_{j:N} = \frac{1}{\Delta_N}(NS_N^2 S_N^1 - S_N^1 NS_N^2) = 0 .$$

Лему доведено.

Зауважимо, що з леми 1 випливає

$$\sum_{j=1}^{N} a_{j:N}^1 (1 - w_{j:N}) = 0$$

і

$$\sum_{j=1}^{N} a_{j:N}^2 (1 - w_{j:N}) = N .$$

Лема 2. Нехай виконується умова *A*. Тоді матимемо

$$W_{N_1}^1(\tau_2) - W_{N_1}^1(\tau_1) =$$

$$= N^{-1/3}\sum_{j=1}^{N}(-b_{j:N})(I\{\xi_{j:N} \in A_N\} - P\{\xi_{j:N} \in A_N\}), \quad (4)$$

$$W_{N_2}^1(\tau_2) - W_{N_2}^1(\tau_1) =$$

$$= N^{-1/3}\sum_{j=1}^{N} b_{j:N}(I\{\xi_{j:N} \in A_N\} - P\{\xi_{j:N} \in A_N\}), \quad (5)$$

$$W_{N_1}^2(\tau_2) - W_{N_1}^2(\tau_1) =$$

$$= N^{-1/3}\sum_{j=1}^{N} b_{j:N}(I\{\xi_{j:N} \in A_N\} - P\{\xi_{j:N} \in A_N\}), \quad (6)$$

$$W_{N_2}^2(\tau_2) - W_{N_2}^2(\tau_1) =$$

$$= N^{-1/3}\sum_{j=1}^{N}(-b_{j:N})(I\{\xi_{j:N} \in A_N\} - P\{\xi_{j:N} \in A_N\}), \quad (7)$$

де $\tau_1 < \tau_2$; $A_N = A_N(\tau_1, \tau_2) = [N^{-1/3}\tau_1, N^{-1/3}\tau_2)$; $b_{j:N} = p_2 a_{j:N}^2 - p_1 a_{j:N}^1$.

Доведення. Проведемо повне доведення формул (6) і (7). Формули (4) і (5) отримані аналогічно (через обмеження обсягу даної статті викладки не наводяться). Враховуючи лему 1, маємо

$$W_{N_1}^2(\tau_2) - W_{N_1}^2(\tau_1) =$$

$$= N^{2/3}[(L_{N_1}^2(t_1^B + N^{-1/3}\tau_2) - L_{N_1}^2(t_1^B) -$$

$$- L_1^2(t_1^B + N^{-1/3}\tau_2) + L_1^2(t_1^B)) - (L_{N_1}^2(t_1^B + N^{-1/3}\tau_1) -$$

$$- L_{N_1}^2(t_1^B) - L_1^2(t_1^B + N^{-1/3}\tau_1) + L_1^2(t_1^B))] =$$

$$= N^{2/3}[-p_1\widehat{H}_1^N(t_1^B + N^{-1/3}\tau_2) +$$

$$+ p_2\widehat{H}_2^N(t_1^B + N^{-1/3}\tau_2) + p_1\widehat{H}_1^N(t_1^B + N^{-1/3}\tau_1) -$$

$$- p_2\widehat{H}_2^N(t_1^B + N^{-1/3}\tau_1) + p_1H_1(t_1^B + N^{-1/3}\tau_2) -$$

$$- p_2H_2(t_1^B + N^{-1/3}\tau_1) - p_1H_1(t_1^B + N^{-1/3}\tau_1) +$$

$$+ p_2H_2(t_1^B + N^{-1/3}\tau_1)] = N^{-1/3} \times$$

$$\times \sum_{j=1}^{N}[(p_2a_{j:N}^2 I\{\xi_{j:N} \in A_N\} - p_1a_{j:N}^1 I\{\xi_{j:N} \in A_N\}) -$$

$$- (p_2a_{j:N}^2 P\{\xi_{j:N} \in A_N\} - p_1a_{j:N}^1 P\{\xi_{j:N} \in A_N\})] =$$

$$= N^{-1/3}\sum_{j=1}^{N}(p_2a_{j:N}^2 - p_1a_{j:N}^1)(I\{\xi_{j:N} \in A_N\} -$$

$$- P\{\xi_{j:N} \in A_N\}).$$

Далі маємо

$$W_{N_2}^2(\tau_2) - W_{N_2}^2(\tau_1) =$$

$$= N^{2/3}[(L_{N_2}^2(t_2^B + N^{-1/3}\tau_2) - L_{N_2}^2(t_2^B) -$$

$$- L_2^2(t_2^B + N^{-1/3}\tau_2) + L_2^2(t_2^B)) -$$

$$- (L_{N_2}^2(t_2^B + N^{-1/3}\tau_1) - L_{N_2}^2(t_2^B) -$$

$$- L_2^2(t_2^B + N^{-1/3}\tau_1) + L_2^2(t_2^B))] = N^{2/3} \times$$

$$\times[p_1\widehat{H}_1^N(t_2^B + N^{-1/3}\tau_2) + p_2(1 - \widehat{H}_2^N(t_2^B + N^{-1/3}\tau_2)) -$$

$$- p_1\widehat{H}_1^N(t_2^B + N^{-1/3}\tau_1) - p_2(1 - \widehat{H}_2^N(t_2^B + N^{-1/3}\tau_2)) -$$

$$- p_1H_1(t_2^B + N^{-1/3}\tau_2) - p_2(1 - H_2(t_2^B + N^{-1/3}\tau_1)) +$$

$$+ p_1H_1(t_2^B + N^{-1/3}\tau_1) + p_2(1 - H_2(t_2^B + N^{-1/3}\tau_1))] =$$

$$= N^{2/3}[p_1\widehat{H}_1^N(t_2^B + N^{-1/3}\tau_2) -$$

$$- p_2\widehat{H}_2^N(t_2^B + N^{-1/3}\tau_2) - p_1\widehat{H}_1^N(t_2^B + N^{-1/3}\tau_1) +$$

$$+ p_2\widehat{H}_2^N(t_2^B + N^{-1/3}\tau_2) - p_1H_1(t_2^B + N^{-1/3}\tau_2) +$$

$$+ p_2H_2(t_2^B + N^{-1/3}\tau_1) + p_1H_1(t_2^B + N^{-1/3}\tau_1) -$$

$$- p_2H_2(t_2^B + N^{-1/3}\tau_1)] = -N^{-1/3} \times$$

$$\times \sum_{j=1}^{N}(p_2a_{j:N}^2 - p_1a_{j:N}^1)(I\{\xi_{j:N} \in A_N\} -$$

$$- P\{\xi_{j:N} \in A_N\}) =$$

$$= N^{-1/3}\sum_{j=1}^{N}(-b_{j:N})(I\{\xi_{j:N} \in A_N\} - P\{\xi_{j:N} \in A_N\}).$$

Лему доведено.

Справедливим є твердження про асимптотичну поведінку процесів $W_{N_i}^1$.

Теорема 1. Нехай виконується умова B_3. Тоді на будь-якому скінченному інтервалі $U_i = [\tau_{i-}, \tau_{i+}]$ випадковий процес $W_{N_i}^1$ слабко збігається при $N \to \infty$ до процесу r_iW у просторі $D(U_i)$ функцій без розривів другого роду з рівномірною метрикою ($i = 1,2$), де W – двосторонній стандартний вінерівський процес.

Аналогічне твердження справедливе і для $W_{N_i}^2$, $i = 1,2$.

Д о в е д е н н я. З [10], враховуючи неперервність траєкторій $W_{N_i}^1$, $W_{N_i}^2$, досить довести асимптотичну нормальність скінченновимірних розподілів $W_{N_i}^1$, $W_{N_i}^2$, збіжність других моментів приростів і щільність розподілів $W_{N_i}^1$, $W_{N_i}^2$ у $D(U_i)$.

Спочатку підрахуємо $E(W_{N_i}^1(\tau_2) - W_{N_i}^1(\tau_1))^2$, $E(W_{N_i}^2(\tau_2) - W_{N_i}^2(\tau_1))^2$, $i = 1,2$. За лемою 2 виконуються рівності (4)–(7). Тоді отримаємо

$$E(W_{N_2}^1(\tau_2) - W_{N_2}^1(\tau_1))^2 = N^{-2/3} \times$$

$$\times \sum_{j=1}^{N}(b_{j:N})^2 E(I\{\xi_{j:N} \in A_N\} - P\{\xi_{j:N} \in A_N\})^2 =$$

$$= N^{-2/3}\sum_{j=1}^{N}(b_{j:N})^2 E(I^2\{\xi_{j:N} \in A_N\} - 2I\{\xi_{j:N} \in A_N\} \times$$

$$\times P\{\xi_{j:N} \in A_N\} + P^2\{\xi_{j:N} \in A_N\}) =$$

$$= N^{-2/3}\sum_{j=1}^{N}(b_{j:N})^2[(w_{j:N}H_1(A_N) + (1 - w_{j:N}) \times$$

$$\times H_2(A_N)) - (w_{j:N}H_1(A_N) + (1 - w_{j:N})H_2(A_N))^2],$$

$$E(W_{N_1}^1(\tau_2) - W_{N_1}^1(\tau_1))^2 = N^{-2/3} \times$$

$$\times \sum_{j=1}^{N}(-b_{j:N})^2 E(I\{\xi_{j:N} \in A_N\} - P\{\xi_{j:N} \in A_N\})^2 =$$

$$= N^{-2/3}\sum_{j=1}^{N}(b_{j:N})^2[(w_{j:N}H_1(A_N) + (1 - w_{j:N}) \times$$

$$\times H_2(A_N)) - (w_{j:N} H_1(A_N) + (1-w_{j:N}) H_2(A_N))^2].$$

Нехай спочатку $H_i(A_N) \sim h_i(t_2^B) N^{-1/3}(\tau_2 - \tau_1)$. Тоді маємо

$$E(W_{N_2}^1(\tau_2) - W_{N_2}^1(\tau_1))^2 =$$

$$= N^{-2/3} \sum_{j=1}^{N} (b_{j:N})^2 [(w_{j:N} h_1(t_2^B) N^{-1/3}(\tau_2 - \tau_1) +$$

$$+ (1-w_{j:N}) h_2(t_2^B) N^{-1/3}(\tau_2 - \tau_1)) -$$

$$- (w_{j:N} h_1(t_2^B) N^{-1/3}(\tau_2 - \tau_1) +$$

$$+ (1-w_{j:N}) h_2(t_2^B) N^{-1/3}(\tau_2 - \tau_1))^2] =$$

$$= N^{-2/3} \sum_{j=1}^{N} N^{-1/3} (b_{j:N})^2 \times$$

$$\times [(w_{j:N} h_1(t_2^B) + (1-w_{j:N}) h_2(t_2^B))(\tau_2 - \tau_1)] \times$$

$$\times [1 - w_{j:N} h_1(t_2^B) N^{-1/3}(\tau_2 - \tau_1) -$$

$$- (1-w_{j:N}) h_2(t_2^B) N^{-1/3}(\tau_2 - \tau_1)] \to r_2^2(\tau_2 - \tau_1),$$

$$N \to \infty,$$

де

$$r_{N_2} = [N^{-1} \sum_{j=1}^{N} (p_2 a_{j:N}^2 - p_1 a_{j:N}^1)^2 \times$$

$$\times [w_{j:N} h_1(t_2^B) + (1-w_{j:N}) h_2(t_2^B)]]^{1/2}; \; r_2 = \lim_{N\to\infty} r_{N_2}.$$

Далі, нехай $H_i(A_N) \sim h_i(t_1^B) N^{-1/3}(\tau_2 - \tau_1)$. Аналогічно до попереднього отримуємо

$$E(W_{N_1}^1(\tau_2) - W_{N_1}^1(\tau_1))^2 =$$

$$= N^{-1} \sum_{j=1}^{N} (b_{j:N})^2 [(w_{j:N} h_1(t_1^B) +$$

$$+ (1-w_{j:N}) h_2(t_1^B))(\tau_2 - \tau_1)] \times$$

$$\times [1 - w_{j:N} h_1(t_1^B) N^{-1/3}(\tau_2 - \tau_1) -$$

$$- (1-w_{j:N}) h_2(t_1^B) N^{-1/3}(\tau_2 - \tau_1)] \to r_1^2(\tau_2 - \tau_1),$$

$$N \to \infty,$$

$$r_{N_1} = [N^{-1} \sum_{j=1}^{N} (p_2 a_{j:N}^2 - p_1 a_{j:N}^1)^2 \times$$

$$\times [w_{j:N} h_1(t_1^B) + (1-w_{j:N}) h_2(t_1^B)]]^{1/2}, \; r_1 = \lim_{N\to\infty} r_{N_1},$$

$$E(W_{N_1}^2(\tau_2) - W_{N_1}^2(\tau_1))^2 \to r_1^2(\tau_2 - \tau_1) =$$

$$= E(r_1 W(\tau_2) - r_1 W(\tau_1))^2, \; N \to \infty,$$

$$E(W_{N_2}^2(\tau_2) - W_{N_2}^2(\tau_1))^2 \to r_2^2(\tau_2 - \tau_1) =$$

$$= E(r_2 W(\tau_2) - r_2 W(\tau_1))^2, \; N \to \infty.$$

Асимптотична нормальність скінченновимірних розподілів $W_{N_i}^1, W_{N_i}^2$, $i = 1,2$, випливає з центральної граничної теореми з умовою Ліндеберга із врахуванням рівномірної обмеженості всіх доданків у сумах (4)–(7), яка має місце, оскільки вимагається виконання умови B_3. Залишилось перевірити щільність сім'ї розподілів $W_{N_i}^1, W_{N_i}^2$. Доведення проведемо для $W_{N_2}^1$. Для того щоб використати відповідний критерій (теорема 15.6 з [10]), покажемо, що для всіх $\tau_1 < \tau < \tau_2$

$$J \doteq E(W_{N_2}^1(\tau) - W_{N_2}^1(\tau_1))^2 (W_{N_2}^1(\tau_2) - W_{N_2}^1(\tau))^2 \le$$

$$\le c_1(\tau_2 - \tau_1)^2, \tag{8}$$

де c_1 не залежить від N, τ_1, τ, τ_2. Покладемо

$$\eta_j(\tau_1, \tau_2) =$$

$$= b_{j:N}(I\{\xi_{j:N} \in A_N(\tau_1,\tau_2)\} - P\{\xi_{j:N} \in A_N(\tau_1,\tau_2)\}).$$

Тоді маємо

$$J \le \frac{c_2}{N^{4/3}} \left(\sum_{j \ne k} \{E(\eta_j(\tau,\tau_2))^2 (\eta_k(\tau_1,\tau))^2 + \right.$$

$$+ |E\eta_j(\tau,\tau_2)\eta_j(\tau_1,\tau)\eta_k(\tau,\tau_2)\eta_k(\tau_1,\tau)|\} +$$

$$\left. + \sum_{j=1}^{N} E\,(\eta_j(\tau_1,\tau_2))^2 (\eta_j(\tau_1,\tau))^2 \right) \le$$

$$\le \frac{c_2}{N^{4/3}} (N^2 c_3 (h^*)^2 N^{-1/3}(\tau - \tau_1) N^{-1/3}(\tau_2 - \tau) +$$

$$+ N^2 c_4 (h^*)^4 N^{-2/3}(\tau - \tau_1)(\tau_2 - \tau)) \le c_1(\tau_2 - \tau_1)^2.$$

Тут c_2 – абсолютна константа; c_2, c_3, c_4 залежать лише від $\sup_{j,N} |b_{j:N}| < \infty$ (нерівність $\sup_{j,N} |b_{j:N}| < \infty$ має місце, оскільки вимагається виконання умови B_3); $h^* = \sup_{t \in U}(h_1(t) + h_2(t))$. Далі, використовуючи твердження 15.4 та 15.6 з

[10], отримаємо твердження леми. Щільність сім'ї розподілів $W^1_{N_1}, W^2_{N_1}, W^2_{N_2}$ доводиться аналогічно до розподілу $W^1_{N_2}$.

Лема 3. Для будь-якого $\varepsilon > 0$ $\exists D > 0$, таке, що для всіх N

$$P\{\exists\tau\in\mathbb{R}: |W^1_{N_i}(\tau)| > D(1+|\tau|)\} < \varepsilon,$$

$$P\{\exists\tau\in\mathbb{R}: |W^2_{N_i}(\tau)| > D(1+|\tau|)\} < \varepsilon,\ i=1,2.$$

Доведення. Доведення проведемо для $W^1_{N_2}$. Позначимо

$$\mathrm{M} \doteq \sup_{\tau\in[\tau_-,\tau_+]} \min(|W^1_{N_2}(\tau) - W^1_{N_2}(\tau_1)|, \\ |W^1_{N_2}(\tau_2) - W^1_{N_2}(\tau)|).$$

З нерівності 15.30 з [10] і (8) випливає

$$J_1 \doteq P\{\mathrm{M} > \varepsilon/2\} \le \frac{16C_1}{\varepsilon^4}(\tau_- - \tau_+)^2.$$

Так само, як і при доведенні (8), отримаємо $EW^1_{N_2}(\tau) \le C_5\tau^2$, тому за нерівністю Чебишова маємо

$$J_2(\tau) \doteq P\left\{|W^1_{N_2}(\tau)| > \frac{\varepsilon}{2}\right\} \le \frac{16C_5\tau^2}{\varepsilon^4}.$$

Отже, отримуємо

$$P\{\sup_{\tau\in[\tau_-,\tau_+]} |W^1_{N_2}(\tau)| > \varepsilon\} \le$$

$$\le J_1(\tau_-) + J_2(\tau_+) \le \frac{C_6}{\varepsilon^4}((\tau_+ - \tau_-)^2 + \tau_-^2 + \tau_+^2).$$

Таким чином, маємо співвідношення

$$P\{\exists\tau\in\mathbb{R}: |W^1_{N_2}(\tau)| > D(1+|\tau|)\} \le$$

$$\le \sum_{j=0}^{\infty}[P\{\exists r\in[i,i+1]: |W^1_{N_2}(\tau)| > D(1+i)\} +$$

$$+ P\{\exists r\in[-i-1,-i]: |W^1_{N_2}(\tau)| > D(1+i)\}] \le$$

$$\le C_7\sum_{i=0}^{\infty}\left(\frac{(i+1-i)^2}{D^4(i+1)^4} + \frac{i^2}{D^4(i+1)^4} + \frac{(i+1)^2}{D^4(i+1)^4}\right) \le \frac{C_8}{D^4},$$

звідки й випливає твердження леми.

Введемо позначення $f(x) \doteq p_2h_2'(x) - p_1h_1'(x)$.

Теорема 2. Якщо виконані умови A, B_3, h_i існують і є неперервно диференційовними в деякому околі точок t_1^B, t_2^B і $f(t_i^B) \ne 0$, то

$$N^{1/3}(\hat{t}^{\mathrm{MER}}_{N_1} - t_1^B) \Rightarrow \left(\frac{2r_1}{f(t_1^B)}\right)^{2/3} z,$$

$$N^{1/3}(\hat{t}^{\mathrm{MER}}_{N_2} - t_2^B) \Rightarrow \left(\frac{2r_2}{f(t_2^B)}\right)^{2/3} z,$$

де $r_i = \lim_{N\to\infty} r_{N_i}$; $z = \arg\min_{t\in\mathbb{R}}(W(t) + t^2)$; $W(t)$ — двосторонній вінерівський процес, $i = 1,2$.

Доведення. Розглянемо випадок $W^1_{N_i}$. Нехай $\bar{U}$ — окіл точки $\bar{t}^B$, на якому h_i є неперервно диференційовними. Згідно з теоремою 2 в [6] $\bar{\hat{t}}^{\mathrm{MER}}_N \to \bar{t}^B$ за ймовірністю. Отже, для довільного $\varepsilon > 0$ і достатньо великих N для $A_{N_i} = \{\hat{t}^{\mathrm{MER}}_{N_i} \in U_i\}$ виконується нерівність $P(A_{N_i}) > 1 - \frac{\varepsilon}{2}$. Позначимо $\bar{\tau}_N = N^{1/3}(\bar{\hat{t}}^{\mathrm{MER}}_N - \bar{t}^B)$. Покажемо, що знайдеться таке $S = S_{\varepsilon i}$, що для достатньо великих N справедливо $P\{|\tau_{N_i}| < S_i\} > 1 - \varepsilon$.

Оскільки

$$L^1_{N_1}(t_i^B + N^{-1/3}\tau) - L^1_{N_1}(t_i^B) =$$

$$= N^{-2/3}W^1_{N_1}(\tau) + L^1_1(t_i^B + N^{-1/3}\tau) - L^1_1(t_i^B),$$

то

$$\tau_{N_i} = \arg\min_{\tau} v_i(\tau),$$

де

$$v_i(\tau) = W^1_{N_i}(\tau) + N^{2/3}(L^1_i(t_i^B + N^{-1/3}\tau) - L^1_i(t_i^B)).$$

Будемо вважати, що подія $\bar{\hat{t}}^{\mathrm{MER}}_N \in \bar{U}$ відбулася. Оскільки H_i — двічі диференційовні на $\bar{U}$ функції і $\bar{t}^B$ — точка мінімуму $\bar{L}^1$, то

$$L_1^1(t_1^B + N^{-1/3}\tau) - L_1^1(t_1^B) = \frac{1}{2} L_1^{1''}(\xi) N^{-2/3}\tau^2 ,$$

$$L_2^1(t_2^B + N^{-1/3}\tau) - L_2^1(t_2^B) = \frac{1}{2} L_2^{1''}(\eta) N^{-2/3}\tau^2 ,$$

де $(\xi,\eta)\in \bar{U}$ — проміжна точка; $L_1^{1''}(\xi) = f_1(\xi)$; $L_2^{1''}(\eta) = f_2(\eta)$.

Оскільки $f_1(t_1^B) > 0$, $f_2(t_2^B) > 0$, то можна вибрати таке $\bar{U}$, щоб $f_1(\xi) > c_1 > 0$ на U_1, $f_2(\eta) > c_2 > 0$ на U_2. Тоді $v_i(\tau) \ge W_{N_i}^1(\tau) + \frac{c_i}{2}\tau^2$ при $N^{-1/3}\tau + t_i^B \in U_i$, $i = 1,2$.

Позначимо $B_{N_i} = \{|W_{N_i}^1(\tau)| < D(1+|\tau|)\ \forall\tau\in \mathbb{R}\}$. За лемою 3, для достатньо великих D $P(B_{N_i}) > 1-\varepsilon/2$, $i = 1,2$. Але, якщо виконана рівність B_{N_i}, то при $c_i\tau^2/2 > D(1+|\tau|)$ маємо $v_i(\tau) > 0 = v_i(0)$. Тому в таких точках мінімум $v_i(\tau)$ досягатись не може. Отже, при виконанні умови $A_{N_i} \cap B_{N_i}$ отримаємо $c_i\tau_{N_i}^2/2 < D(1+|\tau_{N_i}|)$, тобто $|\tau_{N_i}| \le \max(1, 2D/c_i) = S_i$. Оскільки $P(A_{N_i} \cap B_{N_i}) > 1-\varepsilon$, нерівність $P\{|\tau_{N_i}| < S_i\} > > 1-\varepsilon$ доведена.

Аналогічно (або з використанням закону повторного логарифма для W) доводиться, що для достатньо великих S_i матимемо

$$P\left\{ |\arg\min_{\tau\in\mathbb{R}} \tilde{W}_i(\tau)| < S_i \right\} > 1-\varepsilon ,$$

де $\tilde{W}_i(\tau) = r_i W(\tau) + \frac{f_i(t)}{2}\tau^2$. За теоремою 1, на інтервалі $[-S,S]$ $W_{N_i}^1$ слабко збігається до $r_i W$, а внаслідок неперервної диференційовності h_i матимемо, що $N^{2/3}(L_i^1(t_i^B + N^{-1/3}\tau) - - L_i^1(t_i^B))$ рівномірно збігається до $\frac{f_i(0)}{2}\tau^2$. Враховуючи, що функціонал $\arg\min_{\tau\in[-S,S]}$ є неперервним відносно рівномірної норми на $[-S,S]$, отримуємо, що для всіх x

$$P\left\{ \arg\min_{\tau\in[-S_i,S_i]} v_i(\tau) < x \right\} \to P\left\{ \arg\min_{\tau\in[-S_i,S_i]} \tilde{W}_i(\tau) < x \right\}.$$

Це разом з нерівностями $P\{|\tau_{N_i}| < S_i\} > 1-\varepsilon$ і $P\left\{ \left| \arg\min_{\tau\in\mathbb{R}} \tilde{W}_i(\tau) \right| < S_i \right\} > 1-\varepsilon$ забезпечує збіжність τ_{N_i} до $\arg\min_{\tau\in\mathbb{R}} \tilde{W}_i(\tau) = \arg\min_{\tau\in\mathbb{R}}(W(\tau) + \alpha_i\tau^2)$, де $\alpha_i = f_i(t_i^B)/2r_i$. Враховуючи, що $W(\alpha_i\tau) \doteq \doteq \sqrt{\alpha_i}W(\tau)$ ("$\doteq$" позначає рівність за розподілом), отримуємо

$$\arg\min_{\tau\in\mathbb{R}}(W(\tau) + \alpha_i\tau^2) \doteq$$

$$\doteq \arg\min_{\tau\in\mathbb{R}}(\alpha_i^{-1/3}W(\alpha_i^{2/3}\tau) + \alpha_i^{-1/3}(\alpha_i^{2/3}\tau)^2) = \alpha_i^{-2/3}z,$$

$$\tau^0 = \alpha_i^{2/3}\tau,\ \tau = \tau^0\alpha_i^{-2/3},$$

$$W(\tau) + \alpha_i\tau^2 = W(\tau) + (\sqrt{\alpha_i}\tau)^2 = W(\tau^0\alpha_i^{-2/3}) +$$

$$+ (\sqrt{\alpha_i}\tau^0\alpha_i^{-2/3})^2 = \alpha_i^{-1/3}W(\tau^0) + \alpha_i^{-1/3}(\tau^0)^2 =$$

$$= \alpha_i^{-1/3}(W(\alpha_i^{2/3}\tau) + (\alpha_i^{2/3}\tau)^2). \quad .$$

Крім того, відзначимо, що $f_1 = -f_2$, і позначивши $f_2 = f$, маємо $f_1 = -f$. Теорему доведено.

Висновки

Дослідження асимптотики оцінки методу МЕР у випадку, коли об’єкт належить одному з двох класів за наявності двох порогів, дає можливість зробити ще один важливий крок у розв’язанні задач класифікації. Надалі буде розглянуто асимптотичну поведінку оцінки ЕБК для цього ж випадку і проведено порівняння отриманих оцінок.

О.А. Кубайчук

АССИМПТОТИЧЕСКОЕ ПОВЕДЕНИЕ ОЦЕНКИ МЕТОДА МИНИМИЗАЦИИ ЭМПИРИЧЕСКОГО РИСКА ДЛЯ БАЕССОВОГО ПОРОГА

Рассматривается ассимптотическое поведение оценки для баессового порога, построенной методом минимизации эмпирического риска для выборки из смеси со сменными концентрациями.

O.O. Kubaychuk

THE ESTIMATOR ASYMPTOTIC BEHAVIOR OF THE EMPIRICAL RISK MINIMIZATION METHOD FOR BAYESIAN BORDER

This paper highlights the estimator asymptotic behavior for the Bayesian border. Specifically, it is constructed for a sample from the mixture with varying concentrations by the Empirical Risk Minimization method.

1. *Деврой Л., Дьерфи Л.* Непараметрическое оценивание плотности. – М.: Мир, 1988. – 408 с.
2. *Іванько Ю.О., Майборода Р.Є.* Експоненціальні оцінки емпірично-баєсового ризику при класифікації суміші зі змінними концентраціями // Укр. мат. журн. – 2002. – **54**, № 10. – С. 1421–1428.
3. *Вапник В.Н.* Индуктивные принципы поиска эмпирических закономерностей // Распознавание. Классификация. Прогноз. – М., 1988. – Вып.1. – С. 17–81.
4. *Vapnik V.N.* The nature of Statistical Learning Theory. – New York: Springer-Verlag, 1996. – 314 p.
5. *Іванько Ю.О., Майборода Р.Є.* Асимптотика порогових класифікаторів, побудованих за вибіркою з суміші зі змінними концентраціями // Теорія ймовірностей та мат. статистика. – 2006. – № 74. – С. 34–43.
6. *Кубайчук О.О.* Асимптотика оцінки для баєсового порогу // Вісн. КНУ. Сер. Математика. Механіка. – 2008. – № 19. – С. 47–50.
7. *Майборода Р.Є.* Статистичний аналіз сумішей. – К.: ВПЦ "Київський університет", 2003. – 176 с.
8. *Сугакова О.В.* Асимптотика ядерної оцінки щільності порогу по спостереженнях суміші зі змінними концентраціями // Теорія ймовірностей та мат. статистика. – 1998. – **59**, № 10. – С. 156–166.
9. *Іванько Ю.О.* Асимптотика ядерних оцінок щільностей та їх похідних, побудованих за спостереженнями із суміші зі змінними концентраціями // Вісн. КНУ. Сер. Математика. Механіка. – 2003. – № 9. – С. 29–35.
10. *Биллингсли П.* Сходимость вероятностных мер. – М.: Наука, 1977. – 352 с.

Рекомендована Радою
фізико-математичного факультету
НТУУ "КПІ"

Надійшла до редакції
21 вересня 2009 року