Теорія ймовір. та матем. статист.
Вип. 91, 2014, стор. 26–37

Teor. Ĭmovir. ta Matem. Statyst.
No. 91, 2014, pp. 26–37

АДАПТИВНЕ ОЦІНЮВАННЯ У СЕМІПАРАМЕТРИЧНІЙ МОДЕЛІ СУМІШІ

УДК 519.21

О. В. ДОРОНІН

Анотація. Обговорюється модель суміші зі змінними концентраціями. Розглядається параметризація перших K із M компонент суміші, $1 \leq K \leq M$. Розвивається техніка адаптивного семіпараметричного оцінювання на основі узагальнених оцінюючих рівнянь. Доводиться консистентність та асимптотична нормальність введених оцінок.

1. Вступ

Семіпараметрична модель суміші може бути застосована при аналізі соціологічних даних. Розглянемо подібне дослідження на прикладі парламентських виборів, у яких беруть участь дві політичні партії. Виборці можуть проголосувати за одну з двох партій, або зробити інший вибір. Відповідно до цього вибору, кожна людина O буде віднесена до однієї з підпопуляцій Ξ_m, $m = 1, \dots, M$ (у даному випадку $M = 3$) загальної популяції Ξ. Який саме вибір зробив конкретний виборець, точно невідомо. Проте ймовірність $p^m(O)$, $m = 1, \dots, M$, того, що людина зробила певний вибір, може бути оцінена, якщо відомо, до якої виборчої дільниці приписана людина, і скільки всього на цій дільниці було віддано голосів у підтримку того чи іншого вибору.

Дослідників може цікавити, який зв'язок існує між політичним вибором людини O, і деякими іншими показниками $\xi(O)$ (задоволення рівнем доходу людини, соціальне положення, сімейний стан, релігійні погляди, тощо). Такі показники можуть бути зібрані, наприклад, під час перепису населення.

Нехай досліджується задоволеність економічним достатком $\xi(O)$ за десятибальною шкалою. Перші два компоненти суміші сформовані людьми, які віддали свій голос за першу та другу партію відповідно. Тому можливо припустити, що перші два компоненти мають нормальний розподіл з різними середніми та однаковою дисперсією. У загальному випадку припускається, що для перших K підпопуляцій $(1 \leq K \leq M)$ розподіли H_m показників $\xi(O)$ у m-ій підпопуляції належать до деякого параметричного сімейства розподілів з евклідовим параметром $t \in \Theta \subset \mathbb{R}^d$. Істинне значення параметра позначимо як ϑ. Воно вважається невідомим. Задача дослідника в тому, щоб якомога точніше оцінити ϑ. Розподіли характеристик для інших підпопуляцій H_m, $m = (K+1), \dots, M$, вважаються повністю невідомими.

У даній роботі досліджується техніка адаптивних оцінок у розглянутій вище семіпараметричній моделі суміші. Ця техніка полягає в тому, що спочатку параметр ϑ оцінюється за допомогою деякої пілотної оцінки $\tilde{\vartheta}_N$, де N — кількість спостережень $\xi_{j;N}$ з популяції Ξ_N. Потім оцінка $\tilde{\vartheta}_N$ покращується оптимальною у певному сенсі оцінкою $\hat{\vartheta}_N$.

2010 *Mathematics Subject Classification.* Primary 62G05, 62G20, 62F12; Secondary 62P25, 62G30.
Ключові слова і фрази. Адаптивне оцінювання, модель сумішей, узагальнені оціночні рівняння.

Дана робота є продовженням роботи [2]. В останній була знайдена нижня межа для матриці розсіяння оцінок методу узагальнених оціночних рівнянь (Generalized Estimating Equations, GEE).

Тут розглядається модель суміші зі змінними концентраціями. Подібні моделі також можна знайти у роботах [3, 4, 5, 6, 9, 12, 14, 15]. Класичну модель сумішей, де ймовірності p^m є однаковими для всіх спостережень $\xi_{j;N}$, $j = 1, \ldots, N$, та приклади її застосування розглянуто в [16, 20]. Деякі спеціальні класи моделі сумішей досліджено в [10, 11, 13].

У розділі 2 розглянуто попередні результати до даної роботи. Техніка адаптивного оцінювання обговорюється у розділі 3. Результати імітаційного моделювання поміщено у розділ 4.

2. Огляд попередніх результатів

Нехай спостереження $\xi_{j;N} := \xi(O_{j;N})$, $j = 1, \ldots, N$, належать деякому метричному простору $\mathfrak{X}$ із σ-скінченною мірою μ, заданій на борельовій σ-алгебрі $\mathfrak{B}(\mathfrak{X})$. Через $F_m(A) := \mathsf{P}[\xi(O) \in A \mid \operatorname{ind}(O) = m]$ позначимо розподіл m-го компонента (розподіл $\xi(O)$ за умови, що об'єкт O належить m-ій підпопуляції), $m = 1, \ldots, M$, $A \in \mathfrak{B}(\mathfrak{X})$. Набір ймовірностей $(p^m_{j;N})_{j=1,\ldots,N,m=1,\ldots,M}$ вважаемо відомим.

Надалі для зручності нульовий вектор з простору $\mathbb{R}^d$ позначатимемо як $\mathbb{O}_d$. Для дійсної $m \times n$ матриці A будемо писати $A \in \mathbb{R}^{m\times n}$.

Для матриці $p_{\cdot;N} := (p^i_{j;N})_{j=1,\ldots,N,i=1,\ldots,M} \in \mathbb{R}^{N\times M}$, побудованої з набору концентрацій, матриці Грамма (за умови, що вони існують) позначимо як

$$\Gamma_N := \frac{1}{N} p^T_{\cdot;N} p_{\cdot;N}, \qquad \Gamma := \lim_{N\to\infty} \Gamma_N. \tag{1}$$

У [15] розглянута задача мінімізації максимальної дисперсії навантажених емпіричних розподілів $\hat{F}_{i;N}(A) := \frac{1}{N}\sum_{j=1}^N a^i_{j;N}\mathbb{I}_{\{\xi_{j;N}\in A\}}$, що береться серед усіх $A \in \mathfrak{B}(\mathfrak{X})$ та розподілів $F_i(A)$ за умови незсуненості $\hat{F}_{i;N}(A)$.

Отримані мінімаксні вагові коефіцієнти визначаються як

$$a^i_{\cdot;N} := p_{\cdot;N}\Gamma^+_N e_i, \tag{2}$$

де Γ^+_N позначає псевдо-обернену матрицю Мура–Пенроуза для матриці Γ_N, а $e_i := (\mathbb{I}_{\{k=i\}})_{k=1,\ldots,M} \in \mathbb{R}^M$.

Зауважимо, що вагові коефіцієнти $a^i_{j;N}$ можуть приймати від'ємні значення, що призводить до погіршення оцінок. Тому у [5] розглядаються виправлені емпіричні функції розподілу $\hat{F}^+_{i;N}$, яким відповідають деякі невід'ємні вагові коефіцієнти $\tilde{a}^i_{j;N}$:

$$\hat{F}^+_{i;N}(x) := \min\left[1, \sup_{y<x}\hat{F}_{i;N}(y)\right] = \frac{1}{N}\sum_{j=1}^N \tilde{a}^i_{j;N}\mathbb{I}_{\{\xi_{j;N}<x\}}. \tag{3}$$

Навантажені моменти для вимірних функцій $g^i(x;t)\colon \mathfrak{X}\times\Theta \to \mathbb{R}^d$ позначимо як

$$\hat{g}^i_N(t) := \frac{1}{N}\sum_{j=1}^N a^i_{j;N} g^i(\xi_{j;N};t), \qquad i = 1, \ldots, K. \tag{4}$$

Означення 2.1. Будемо казати, що випадкова послідовність ϕ_N зрештою співпадає з випадковою послідовністю ψ_N, якщо $\phi_N = \psi_N$ м.н. починаючи з деякого випадкового номера N.

Означення 2.2. GEE-оцінка $\hat{\vartheta}_N$ для ϑ визначається в [2] як вимірна функція від вибірки $\xi_{1;N}, \ldots, \xi_{N;N}$ така, що зрештою

$$\sum_{k=1}^{K} \hat{g}_N^k(\hat{\vartheta}_N) = \mathbb{O}_d. \tag{5}$$

Означення 2.3. Нехай послідовність оцінок ζ_N, $N \geq 1$, є асимптотично нормальною, тобто $\sqrt{N} \cdot \zeta_N$ збігається за розподілом до деякої випадкової величини з розподілом $\mathcal{N}(a, \Sigma)$. Величину Σ будемо назвати коефіцієнтом розсіяння, якщо ζ_N — випадкове число, і матрицею розсіяння, якщо ζ_N — випадковий вектор.

Приклад 2.4. Розглянутий у вступі приклад соціологічного дослідження може бути змодельований трьохкомпонентною сумішшю зі змінними концентраціями. Нехай всі три компоненти нормально розподілені (мають розподіл $\mathcal{N}(m_1, \sigma^2)$, $\mathcal{N}(m_2, \sigma^2)$, $\mathcal{N}(m_3, \sigma_3^2)$ відповідно). Оберемо істинні параметри всіх компонент як $m_1 = -3$, $m_2 = 2$, $\sigma = 2$, $m_3 = 0$, $\sigma_3 = 2$. Концентрації компонентів візьмемо як $p^{\cdot}_{j;N} := \frac{1}{S_{j;N}}(u^1_{j;N}, u^2_{j;N}, 2u^3_{j;N})^T$, де $u^i_{j;N}$ — незалежні рівномірно розподілені на $[0, 1]$ випадкові величини, $i = 1, 2, 3$, $S_{j;N} := u^1_{j;N} + u^2_{j;N} + 2u^3_{j;N}$, $j = 1, \ldots, N$. Розподіл третього компонента вважаємо повністю невідомим. Істинне значення параметра виглядає як $t = (m_1, m_2, \sigma)^T$. Моментну оцінку $\tilde{\vartheta}_N$ для ϑ визначимо як $\tilde{m}_{i;N} := \frac{1}{N}\sum_{j=1}^{N} a^i_{j;N}\xi_{j;N}$, $i = 1, 2$, $\tilde{\sigma}_N := \sqrt{\frac{1}{2}\sum_{k=1}^{2}\frac{1}{N}\sum_{j=1}^{N} a^k_{j;N}(\xi_{j;N} - \tilde{m}_{k;N})^2}$ (за умови, що вираз під знаком кореня є невід'ємним). Аналогічним чином визначається покращена моментна оцінка $\tilde{\vartheta}_N^{impr}$, використовуючи виправлені вагові коефіцієнти $\tilde{a}^i_{j;N}$ (див. (3)). У [2] були обчислені асимптотичні коефіцієнти розсіяння моментних оцінок: $d^m(\tilde{m}_{1;N}) = 281.204$, $d^m(\tilde{m}_{2;N}) = 244.921$, $d^m(\tilde{\sigma}_N) = 217.29$. В той час нижня межа була знайдена як $d^*(\hat{m}_1) = 57.24$, $d^*(\hat{m}_2) = 60.52$, $d^*(\hat{\sigma}) = 19.21$.

У [2] було показано, що завдяки вдалому підбору оціночних функцій $g^k(x; t)$, $k = 1, \ldots, K$ можливо досягти значного зменшення коефіцієнтів розсіяння оцінок, порівняно з використанням деяких простих фукнцій.

3. Адаптивне оцінювання

У [2] знайдено оптимальні функції $g^{*;k}(x; t)$, $k = 1, \ldots, K$, на яких досягається нижня межа матриці розсіяння оцінок. Але вони залежать від щільностей невідомих компонентів суміші. Отже, їх застосування на практиці є неможливим. Тому пропонується використання адаптивної техніки оцінювання, аналогічної розглянутій у [15] на випадок задання параметричної моделі тільки для першого компонента суміші.

Нехай задані деякі базисні функції $u^k(x; t)\colon \mathfrak{X} \times \Theta \to \mathbb{R}^{L_k}$, $k = 1, \ldots, K$. Розглянемо оціночні функції $g^k(x; t)$, $k = 1, \ldots, K$, вигляду

$$g^k(x; t; B^k) := B^k \cdot u^k(x; t), \tag{6}$$

де $B^k \in \mathbb{R}^{d \times L_k}$, $k = 1, \ldots, K$, — матриці коефіцієнтів. Через $B^{*;k}$, $k = 1, \ldots, K$, позначимо матриці B^k, які мінімізують матрицю розсіяння оцінки у класі всіх GEE-оцінок з оціночними функціями вигляду (6) (див. зауваження до леми 3.3). Порівняння матриць розуміється в сенсі Льовнера, тобто для матриць A і B пишемо, що $A \geq B$, якщо матриця $(A - B)$ — невід'ємно визначена. Матриці $B^{*;k}$ залежать від ϑ і щільностей невідомих розподілів. Тому їх варто замінити деякою оцінкою $\hat{B}^k_N$.

Отже, оцінку $\hat{\vartheta}_N$, визначену в (5), можливо побудувати як розв'язок рівняння

$$\sum_{k=1}^{K} \hat{g}_N^k\left(t; \hat{B}^k_N\right) = \mathbb{O}_d \tag{7}$$

відносно $t \in \Theta$. Взагалі кажучи, вказане рівняння розв'язати доволі складно. Так само як перевірити, чи має воно рівно один корінь. Тому не можна бути впевненим, що оцінка $\hat{\vartheta}_N$ існує і є консистентною.

Натомість пропонується почати з деякої $\sqrt{N}$-консистентної "пілотної" оцінки $\tilde{\vartheta}_N$ для ϑ (це може бути оцінка методу моментів). Позначимо

$$V\left(t; \{B^k\}\right) := \sum_{k=1}^{K} B^k \int \frac{\partial}{\partial t^T} u^k(x;t)\, F_k(dx;t) \in \mathbb{R}^{d\times d}. \tag{8}$$

Тоді наближення рівняння (7) матиме форму

$$\mathbb{O}_d = \sum_{k=1}^{K} \hat{g}_N^k(t; \hat{B}_N^k) \approx \sum_{k=1}^{K} \hat{B}_N^k \cdot \hat{u}_N^k(\tilde{\vartheta}_N) + V(\tilde{\vartheta}_N; \{\hat{B}_N^k\}) \cdot (t - \tilde{\vartheta}_N). \tag{9}$$

Якщо матриці $\hat{B}_N^k$ обрати так, щоб $V\big(\tilde{\vartheta}_N; \{\hat{B}_N^k\}\big) \approx \mathbb{I}_{d\times d}$, то наближений розв'язок рівняння (7) матиме вигляд

$$\check{\vartheta}_N := \tilde{\vartheta}_N - \sum_{k=1}^{K} \hat{B}_N^k \cdot \hat{u}_N^k(\tilde{\vartheta}_N). \tag{10}$$

Це і є запропонована адаптивна оцінка для ϑ. Далі буде показано, що її матриця розсіяння є такою ж, як і матриця розсіяння $D[\{B^{*;k}\}]$ із леми 3.3.

Зауваження 3.1. Довільна вимірна фукнція $g^k(x;t)$ може бути наближена із заданою точністю за допомогою функцій вигляду $B^k u^k(x;t)$ за умови вибору достатньо широкого базису $u^k(x;t) \in \mathbb{R}^{L_k}$, $k = 1, \ldots, K$. Отже, можна сподіватись, що при правильному виборі $u^k(x;t)$ матриця розсіяння оцінки $\check{\vartheta}_N$, визначеної в (7), буде близькою до нижньої межі, знайденої в [2].

Зауваження 3.2. На матрицю $V(\tilde{\vartheta}_N; \{\hat{B}_N^k\})$ обов'язково слід накладати умову невиродженості в деякому околі $\mathfrak{U} \subset \Theta$ для ϑ. Дійсно, нехай $V(\tilde{\vartheta}_N; \{\hat{B}_N^k\})$ — вироджена (вона також буде виродженою і в деякому околі $\tilde{\vartheta}_N$, якщо брати $u^k(x;t)$ неперервними). Тоді існує вектор $v \in \mathbb{R}^d$, $v \neq 0$ такий, що $V(\tilde{\vartheta}_N; \{\hat{B}_N^k\}) \cdot v = \mathbb{O}_d$. І якщо існує $\hat{\vartheta}_N$ — розв'язок наближення (9), то для довільного $c \in \mathbb{R}$ значення $\hat{\vartheta}_N + cv$ також буде розв'язком цього наближення. Отже, зміщення оцінки $\hat{\vartheta}_N$ у даному випадку можливо зробити як завгодно великим.

3.1. **Асимптотика адаптивної оцінки.** Припустимо, що функції розподілу $F_1, \ldots, F_M$ є абсолютно неперервними відносно міри μ. Щільності розподілу кожного компонента суміші позначимо через

$$\begin{gathered} f_k(\cdot;t) := \frac{dF_k(\cdot;t)}{d\mu(\cdot)}, \qquad f_k(\cdot) := f_k(\cdot,\vartheta), \qquad k = 1, \ldots, K, \\ f_k(\cdot) := \frac{dF_k(\cdot)}{d\mu(\cdot)}, \qquad k = K+1, \ldots, M. \end{gathered} \tag{11}$$

Для зручності визначимо матрицю з усіх базисних функцій $u^k(x;t)$, $k = 1, \ldots, K$:

$$U(x;t) := \mathrm{Diag}(u^{\cdot}(x;t)) := \begin{pmatrix} u^1(x;t) & \mathbb{O}_{L_1} & \cdots & \mathbb{O}_{L_1} \\ \mathbb{O}_{L_2} & u^2(x;t) & \cdots & \mathbb{O}_{L_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{O}_{L_K} & \mathbb{O}_{L_K} & \cdots & u^K(x;t) \end{pmatrix} \in \mathbb{R}^{L\times K}.$$

Її математичне сподівання від m-го компонента позначимо як

$$\bar{U}^m(t) := \begin{cases} \int U(x;t) f_m(x;t)\,\mu(dx), & m = 1,\dots,K, \\ \int U(x;t) f_m(x)\,\mu(dx), & m = K+1,\dots,M, \end{cases} \quad \in \mathbb{R}^{L\times K}.$$

Нехай $\{b_{j;N}\}_{j=1,\dots,N}$, $\{c_{j;N}\}_{j=1,\dots,N}$ — деякі набори коефіцієнтів. Визначимо оператор осереднення $\langle\cdot\rangle_N$ та відповідні арифметичні операції як

$$\langle b_{\cdot;N}\rangle_N := \frac{1}{N}\sum_{j=1}^{N} b_{j;N}, \qquad \langle b_{\cdot;N}\cdot c_{\cdot;N}\rangle_N := \frac{1}{N}\sum_{j=1}^{N} b_{j;N}\cdot c_{j;N}. \tag{12}$$

Введемо формально випадкові величини η_m з функцією розподілу F_m, $m = 1,\dots,M$. Наступні позначення зручні при обчисленні виразів виду $\operatorname{Cov}[\hat{u}_N^k, \hat{u}_N^l]$. При цьому матриці $\alpha_{r,s;N}$ з'являються при доданках виду $\mathsf{E}[u^k(\eta_r)]\,\mathsf{E}[u^l(\eta_s)]^T$, а $\beta_{m;N}$ — при доданках $\mathsf{E}[u^k(\eta_m)u^l(\eta_m)^T]$.

$$\begin{aligned} \alpha_{r,s;N} &:= \left(\alpha_{r,s;N}^{k,l}\right)_{k,l=1,\dots,K} \\ &:= \left(\langle a_{\cdot;N}^k a_{\cdot;N}^l p_{\cdot;N}^r p_{\cdot;N}^s\rangle_N\right)_{k,l=1,\dots,K} \in \mathbb{R}^{K\times K}, \qquad r,s = 1,\dots,M; \\ \beta_{m;N} &:= \left(\beta_{m;N}^{k,l}\right)_{k,l=1,\dots,K} \\ &:= \left(\langle a_{\cdot;N}^k a_{\cdot;N}^l p_{\cdot;N}^m\rangle_N\right)_{k,l=1,\dots,K} \in \mathbb{R}^{K\times K}, \qquad m = 1,\dots,M. \end{aligned}$$

$$\begin{aligned} \alpha_{r,s} &:= \lim_{N\to\infty} \alpha_{r,s;N} \in \mathbb{R}^{K\times K}, \qquad r,s = 1,\dots,M; \\ \beta_m &:= \lim_{N\to\infty} \beta_{m;N} \in \mathbb{R}^{K\times K}, \qquad m = 1,\dots,M. \end{aligned} \tag{13}$$

Завдяки наступному позначенню вираз $\sum_{m=1}^{M}\mathsf{E}[U(\eta_m)\beta_m U(\eta_m)^T]$ переписується у вигляді $\int U(x)R(x)U(x)^T\,\mu(dx)$.

$$R(x;t) := \sum_{m=1}^{K}\beta_m f_m(x;t) + \sum_{m=K+1}^{M}\beta_m f_m(x); \tag{14}$$

Сумарну розмірність простору з $u^k(x;t)$, $k = 1,\dots,K$, позначимо як

$$L := \sum_{k=1}^{K} L_k.$$

А сукупну матрицю з невідомих коефіцієнтів — як

$$B := (B^1 \cdots B^K) \in \mathbb{R}^{d\times L}.$$

Через $\mathbb{Z}(\vartheta)$ позначимо матрицю розсіяння вектора $(\hat{u}_N^1(\vartheta),\dots,\hat{u}_N^K(\vartheta))^T$. Зазначимо, що тоді $B\mathbb{Z}(\vartheta)B^T$ — матриця розсіяння виразу $\sum_{k=1}^{K} B^k\hat{u}_N^k(\vartheta)$.

$$\begin{aligned} \mathbb{Z}_1(\vartheta) &:= \sum_{r,s=1}^{M}\bar{U}(\vartheta)^r\alpha_{r,s}\bar{U}^s(\vartheta)^T \in \mathbb{R}^{L\times L}; \\ \mathbb{Z}_2(\vartheta) &:= \int U(x;\vartheta)R(x;\vartheta)U(x;\vartheta)^T\,\mu(dx) \in \mathbb{R}^{L\times L}; \\ \mathbb{Z}(\vartheta) &:= \mathbb{Z}_2(\vartheta) - \mathbb{Z}_1(\vartheta) \in \mathbb{R}^{L\times L}. \end{aligned}$$

Аналогічно до [2], припустимо, що для функцій $u^k(x;t)$, $k = 1,\dots,K$, ненульових матриць B^k, і деякого відкритого околу $\mathfrak{U} \subset \Theta$ для ϑ виконуються наступні умови.

(a1) $u^1(x;t),\dots,u^K(x;t)$ — диференційовані за t для майже всіх $x \pmod{\mu}$, $t \in \mathfrak{U}$.

(a2) Існує $\delta > 0$ таке, що $\int \sup_{t\in\mathfrak{U}} \left\| \frac{\partial}{\partial t^T} u^k(x;t) \right\|^{1+\delta} F_i(dx) < \infty$, $i = 1,\dots,M$, $k = 1,\dots,K$.
(a3) $\int \left\| u^k(x;\vartheta) \right\|^2 F_i(dx) < \infty$, $i = 1,\dots,M$, $k = 1,\dots,K$.
(b1) $B^k \int u^k(x;\vartheta) F_k(dx;\vartheta) = \mathbb{O}_d$, $k = 1,\dots,K$ (незсуненість).
(b2) Існує скінченна невироджена матриця $V(t;\{B^k\})$ (див. (8)), $t \in \mathfrak{U}$.

Лема 3.3. *Нехай*
(i) $\tilde\vartheta_N$ — $\sqrt{N}$*-консистентна оцінка для* ϑ.
(ii) *Для функцій* $u^k(x;t)$, $k = 1,\dots,K$, *виконуються умови (a1)–(a3).*
(iii) *Для матриць* B^k, $k = 1,\dots,K$, *виконуються умови (b1), (b2).*
(iv) *Існують граничні матриці* $\alpha_{r,s}$, β_m, *визначені в* (13), $m,r,s = 1,\dots,M$.
(v) *Адаптивна оцінка визначена як* $\hat\vartheta_N := \tilde\vartheta_N - V(\tilde\vartheta_N;\{B^k\})^{-1} \sum_{k=1}^K B^k \hat u_N^k(\tilde\vartheta_N)$.

Тоді $\sqrt{N}\cdot(\hat\vartheta_N - \vartheta) \xrightarrow{W_\vartheta} \mathcal{N}(\mathbb{O}_d, D[\{B^k\}])$, *де* $D[\{B^k\}] := V^{-1} B \mathbb{Z}(\vartheta) B^T (V^{-1})^T$, $V := V(\vartheta;\{B^k\})$.

Доведення. Аналогічно доведенню теореми 3.2 з [2]. □

Зауваження 3.4. У теоремі 3.2 з [2] була знайдена матриця розсіяння GEE-оцінки. Вона співпадає з матрицею розсіяння $D[\{B^k\}]$, знайденою у лемі 3.3, якщо в якості оціночних функцій взяти $g^k(x;t) = B^k u^k(x;t)$, $k = 1,\dots,K$.

3.2. **Оптимальні коефіцієнти** B^***.** Позначимо похідну від щільності компонента по параметру як

$$f_k'(x;t) := \frac{\partial f_k(x;t)}{\partial t^T} \in \mathbb{R}^d, \qquad k = 1,\dots,K. \tag{15}$$

Для запису умови $V(\vartheta;\{B^k\}) = \mathbb{I}_{d\times d}$ (див. (8)) введемо у розгляд матриці $\mathbb{V}_1^k(t)$. При доведенні теореми 3.5 буде пояснено, чому дані матриці мають саме таку форму.

$$\mathbb{V}_1^k(t) := \int u^k(x;t) f_k'(x;t)^T \,\mu(dx) \in \mathbb{R}^{L_k\times d}, \qquad k = 1,\dots,K.$$

Моменти $u^k(x;t)$ позначимо як

$$\mathbb{V}_2^k(t) := \int u^k(x;t) f_k(x;t) \,\mu(dx) \in \mathbb{R}^{L_k}, \qquad k = 1,\dots,K.$$

У наступних позначеннях умови **(b1)** та $V(\vartheta;\{B^k\}) = \mathbb{I}_{d\times d}$ записуються як $B\mathbb{V}(\vartheta) = -E_0$. У доведенні теореми 3.5 це пояснено детальніше. Отже, позначимо:

$$\mathbb{V}^k(t) := \left(\mathbb{V}_1^k(t) \mathbb{O}_{L_k\times(k-1)} \mathbb{V}_2^k(t) \mathbb{O}_{L_k\times(K-k)}\right) \in \mathbb{R}^{L_k\times(d+K)}, \qquad k = 1,\dots,K;$$

$$\mathbb{V}(t) := \begin{pmatrix} \mathbb{V}^1(t) \\ \vdots \\ \mathbb{V}^K(t) \end{pmatrix} \in \mathbb{R}^{L\times(d+K)};$$

$$E_0 := \left(\mathbb{I}_{d\times d} \mathbb{O}_{d\times K}\right) \in \mathbb{R}^{d\times(d+K)}.$$

Нарешті, оптимальні коефіцієнти при $u^k(x;\vartheta)$, $k = 1,\dots,K$ записуються у вигляді

$$B^* := -E_0 \left[\mathbb{V}(\vartheta)^T \mathbb{Z}(\vartheta)^{-1} \mathbb{V}(\vartheta)\right]^{-1} \mathbb{V}(\vartheta)^T \mathbb{Z}(\vartheta)^{-1} \in \mathbb{R}^{d\times L}.$$

Теорема 3.5. *В умовах леми 3.3 нехай виконується наступне.*
(i) *Існують неперервні похідні від щільностей, визначені в (15).*
(ii) *Існують невироджені матриці* $\mathbb{Z}(\vartheta)$ *та* $\mathbb{V}(\vartheta)^T \mathbb{Z}(\vartheta)^{-1} \mathbb{V}(\vartheta)$.

Тоді $D[\{B^k\}] \ge D[\{B^{*;k}\}]$ *у тому розумінні, що матриця* $(D[\{B^k\}] - D[\{B^{*;k}\}])$ — *невід'ємно визначена.*

Доведення. Обмеження на пробні функції.

Позначимо $g^k(x;t) := B^k u^k(x;t)$. З умови **(b1)** маємо:

$$\mathbb{O}_d = \frac{\partial}{\partial t_i} \int g^k(x;t) f_k(x;t)\,\mu(dx)$$
$$= \int \left(\frac{\partial}{\partial t_i} g^k(x;t)\right) f_k(x;t)\,\mu(dx) + \int g^k(x;t) \left(\frac{\partial}{\partial t_i} f_k(x;t)\right) \mu(dx).$$

Звідси $V(\vartheta;\{B^k\}) = -\sum_{k=1}^K \int g^k(x;t) f_k'(x;t)^T \mu(dx)$.

Зазначимо, що без втрати загальності можна вважати, що $V(\vartheta;\{B^k\}) = \mathbb{I}_{d\times d}$. Дійсно, GEE-оцінка $\hat{\vartheta}_N$ не зміниться, якщо замінити функції $g^k(x;t)$ функціями $V(\vartheta;\{B^k\})^{-1} g^k(x;t)$, $k = 1,\ldots,K$.

Отже, обмеження на функції $g^k(x;t)$, $k = 1,\ldots,K$ можуть бути записані як

$$\begin{cases} -\sum_{k=1}^K \int g^k(x;\vartheta)\cdot f_k'(x;\vartheta)^T\,\mu(dx) = \mathbb{I}_{d\times d}, \\ \int g^k(x;\vartheta) f_k(x;\vartheta)\,\mu(dx) = \mathbb{O}_d, \end{cases} \qquad k = 1,\ldots,K. \tag{16}$$

Умови (16) переписуються у формі

$$\begin{cases} -\sum_{k=1}^K B^k \mathbb{V}_1^k(\vartheta) = \mathbb{I}_{d\times d}, \\ B^k \mathbb{V}_2^k(\vartheta) = \mathbb{O}_d, \end{cases} \qquad k = 1,\ldots,K.$$

Або, рівносильно,

$$B\mathbb{V}(\vartheta) = -E_0. \tag{17}$$

Задача мінімізації. Для доведення теореми потрібно показати, що для довільного $c \in \mathbb{R}^d$ і $B \in \mathbb{R}^{d\times R}$, яке задовольняє умови (17), виконується

$$c^T(D[\{B^k\}] - D[\{B^{*;k}\}])c \geq 0.$$

Іншими словами, для всіх $c \in \mathbb{R}^d$ потрібно мінімізувати $c^T B\mathbb{Z}(\vartheta)B^T c$ за обмежень (17).

Позначимо $b := B^T c$. Задача мінімізації $c^T B\mathbb{Z}(\vartheta)B^T c$ за умов (17) приймає форму

$$\begin{cases} b^T\mathbb{Z}(\vartheta)b \to \min; \\ \mathbb{V}(\vartheta)^T b = -E_0^T c. \end{cases} \tag{18}$$

Оптимальні коефіцієнти. Введемо змінну $\lambda \in \mathbb{R}^R$. Функціонал Лагранжа, який відповідає задачі мінімізації (18), виражається як

$$\mathfrak{L}(b) := b^T\mathbb{Z}(\vartheta)b - 2\big(\mathbb{V}(\vartheta)^T b + E_0^T c\big)^T \lambda.$$

Знайдемо стаціонарну точку цього функціоналу. З рівняння $\frac{\partial \mathfrak{L}(\cdot)}{\partial b} = 0$ маємо $b^* = \mathbb{Z}(\vartheta)^{-1}\mathbb{V}(\vartheta)\lambda^*$. Із умов (18) випливає $\mathbb{V}(\vartheta)^T\mathbb{Z}(\vartheta)^{-1}\mathbb{V}(\vartheta)\lambda^* = -E_0^T c$. Звідси $\lambda^* = -\big[\mathbb{V}(\vartheta)^T\mathbb{Z}(\vartheta)^{-1}\mathbb{V}(\vartheta)\big]^{-1}E_0^T c$. Отже, $b^* = -\mathbb{Z}(\vartheta)^{-1}\mathbb{V}(\vartheta)\big[\mathbb{V}(\vartheta)^T\mathbb{Z}(\vartheta)^{-1}\mathbb{V}(\vartheta)\big]^{-1}E_0^T c$.

Оскільки $c \in \mathbb{R}^d$ вибрано довільним чином, а $b^* = {B^*}^T c$, маємо

$$B^* = -E_0\big[\mathbb{V}(\vartheta)^T\mathbb{Z}(\vartheta)^{-1}\mathbb{Z}(\vartheta)\big]^{-1}\mathbb{V}(\vartheta)^T\mathbb{Z}(\vartheta)^{-1}.$$

Коректність розв'язку. Функціонал $b^T\mathbb{Z}(\vartheta)b$, який мінімізується у задачі (18), є квадратичною формою на афінному просторі. Оскільки матриця $\mathbb{Z}(\vartheta)$ — невід'ємно визначена (це є коваріаційна матриця певного випадкового вектора), то стаціонарна точка цього функціоналу є точкою його глобального мінімума. □

Зауваження 3.6. Умови **(b1)**, **(b2)** виконуються для матриць $B^{*;k}$, $k = 1,\ldots,K$, за їх побудовою (див. доведення теореми 3.5). Крім того, $V(\vartheta;\{B^{*;k}\}) = \mathbb{I}_{d\times d}$.

Приклад 3.7. Розглянемо модель суміші з прикладу 2.4. Для адаптивної оцінки візьмемо набір із 7 рівномірних B-сплайнів $u(x; m, \sigma) \in \mathbb{R}^7$ з вузлами в точках

$$\{m + k \cdot \sigma\}_{k=-5,\dots,5}.$$

Визначимо $u^1(x;t) := u(x;t_1,t_3)$, $u^2(x;t) := u(x;t_2,t_3)$. Для так визначеної адаптивної оцінки при оптимальному виборі B^* коефіцієнти розсіяння дорівнюють $d^a(\hat{m}^1) = 62.77$, $d^a(\hat{m}^2) = 66.40$, $d^a(\hat{\sigma}) = 22.16$. Таким чином, коефіцієнти розсіяння адаптивної оцінки є доволі близькими до нижньої межі коефіцієнтів розсіяння GEE-оцінок.

3.3. **Оцінювання B^*.** Оптимальні коефіцієнти B^* на практиці можливо оцінити наступним чином:

$$\hat{U}_N^m := \frac{1}{N}\sum_{j=1}^{N} a_{j;N}^m U(\xi_{j;N};\tilde{\vartheta}_N) \in \mathbb{R}^{L\times K}, \qquad m = 1,\dots,M;$$

$$\hat{\mathbb{Z}}_{1;N} := \sum_{r,s=1}^{M} \hat{U}_N^r \alpha_{r,s;N} (\hat{U}_N^s)^T \in \mathbb{R}^{L\times L};$$

$$\hat{\mathbb{Z}}_{2;N} := \frac{1}{N}\sum_{j=1}^{N} U(\xi_{j;N};\tilde{\vartheta}_N)\Big[\sum_{m=1}^{M} a_{j;N}^m \beta_{m;N}\Big] U(\xi_{j;N};\tilde{\vartheta}_N)^T \in \mathbb{R}^{L\times L};$$

$$\hat{\mathbb{Z}}_N := \hat{\mathbb{Z}}_{2;N} - \hat{\mathbb{Z}}_{1;N}.$$

$$\hat{B}_N^* := -E_0\big[\mathbb{V}(\tilde{\vartheta}_N)^T \hat{\mathbb{Z}}_N^{-1} \mathbb{V}(\tilde{\vartheta}_N)\big]^{-1} \mathbb{V}(\tilde{\vartheta}_N)^T \hat{\mathbb{Z}}_N^{-1} \in \mathbb{R}^{d\times L}.$$

Лема 3.8. *Припустимо, що*

(i) *Виконуються умови (a1), (a3).*

(ii) *Для деякої відкритої околу $\mathfrak{U}$, $\vartheta \in \mathfrak{U} \subset \Theta$ виконується*

$$\int \sup_{t\in\mathfrak{U}} \Big\|\frac{\partial}{\partial t} u^k(x;t)\Big\|^2 F_i(dx) < \infty, \qquad i = 1,\dots,M,\ k = 1,\dots,K$$

(посилена умова (a2)).

(iii) $\det\Gamma \neq 0$.

(iv) *Існують граничні матриці $\alpha_{r,s} \in \mathbb{R}^{K\times K}$, $\beta_m \in \mathbb{R}^{K\times K}$, $r,s,m = 1,\dots,M$.*

(v) *$\tilde{\vartheta}_N \to \vartheta$ за ймовірністю.*

Тоді $\hat{\mathbb{Z}}_N \to \mathbb{Z}(\vartheta)$ за ймовірністю.

Доведення. Аналогічно доведенню леми 2 з [15].

Позначимо як $\kappa(\cdot)$ — номер компонента суміші, а $\lambda(\cdot)$ — координату вектора $u^k(x;t)$:

$$\begin{aligned} &\kappa(1) := \cdots := \kappa(L_1) := 1; && \lambda(1) := 1,\dots,\lambda(L_1) := L_1;\\ &\qquad\vdots && \qquad\vdots \\ &\kappa(L-L_K+1) := \cdots := \kappa(L) := K; && \lambda(L-L_K+1) := 1,\dots,\lambda(L) := L_K. \end{aligned}$$

Нехай $k,l \in \{1,\dots,K\}$. Виразимо $(\mathbb{Z}_1)_{k,l}$ і $(\mathbb{Z}_2)_{k,l}$.

$$(\mathbb{Z}_1)_{k,l} = \sum_{r,s=1}^{M} \Big\langle a^{\kappa(k)} a^{\kappa(l)} p^r p^s \Big\rangle \int u_{\lambda(k)}^{\kappa(k)}(x;\vartheta) f_r(x)\,\mu(dx) \int u_{\lambda(l)}^{\kappa(l)}(x;\vartheta) f_s(x)\,\mu(dx).$$

За лемами 4 та 1 з [15], де

$$b_{j;N} := a_{j;N}^{\kappa(k)} a_{j;N}^{\kappa(l)}, \qquad u(x) := u_{\lambda(k)}^{\kappa(k)}(x;\vartheta), \qquad v(x) := u_{\lambda(l)}^{\kappa(l)}(x;\vartheta),$$

маємо, що $(\hat{\mathbb{Z}}_{1;N})_{k,l} \to (\mathbb{Z}_1)_{k,l}$ за ймовірністю.

$$(\mathbb{Z}_2)_{k,l} = \sum_{m=1}^{M} \left\langle a^{\kappa(k)} a^{\kappa(l)} p^m \right\rangle \int u^{\kappa(k)}_{\lambda(k)}(x;\vartheta) u^{\kappa(l)}_{\lambda(l)}(x;\vartheta) f_m(x)\, \mu(dx).$$

За лемою 3 з [15], де $b_{j;N} := a^{\kappa(k)}_{j;N} a^{\kappa(l)}_{j;N}$, $v(x;t) := u^{\kappa(k)}_{\lambda(k)}(x;t) u^{\kappa(l)}_{\lambda(l)}(x;t)$, маємо, що $(\hat{\mathbb{Z}}_{2;N})_{k,l} \to (\mathbb{Z}_2)_{k,l}$ за ймовірністю. □

3.4. **Адаптивна оцінка.** Визначимо адаптивну оцінку як

$$\check{\vartheta}_N := \tilde{\vartheta}_N - \sum_{k=1}^{K} \hat{B}^{*;k}_N \cdot \hat{u}^k_N(\tilde{\vartheta}_N).$$

Теорема 3.9. *Нехай виконуються умови (i)–(iv) леми 3.8, а також:*

(i) *Матриця $\mathbb{V}(\vartheta)$ — повного рангу.*

(ii) *Оцінка $\tilde{\vartheta}_N$ — $\sqrt{N}$-консистентна для ϑ.*

Тоді $\sqrt{N} \cdot (\check{\vartheta}_N - \vartheta) \xrightarrow{W} \mathcal{N}(\mathbb{O}_d, D[\{B^{*;k}\}])$.

Доведення. Аналогічно доведенню теореми 4 з [15].

Розіб'ємо $\sqrt{N}(\check{\vartheta}_N - \vartheta)$.

$$\sqrt{N}(\check{\vartheta}_N - \vartheta) = \sqrt{N}\left(\tilde{\vartheta}_N - \vartheta - \sum_{k=1}^{K} \hat{B}^{*;k}_N \hat{u}^k_N(\tilde{\vartheta}_N)\right) = -\sqrt{N} \sum_{k=1}^{K} \hat{B}^{*;k}_N \hat{u}^k_N(\vartheta) + \varepsilon_1,$$

де

$$\varepsilon_1 := \sqrt{N}\left(\mathbb{I}_{d\times d} - \sum_{k=1}^{K} \hat{B}^{*;k}_N \frac{\partial}{\partial t} \hat{u}^k_N(t)\,\bigg|_{t=\zeta}\right)(\tilde{\vartheta}_N - \vartheta),$$

ζ — проміжна точка між ϑ і $\tilde{\vartheta}_N$.

В той же час

$$\sqrt{N} \sum_{k=1}^{K} \hat{B}^{*;k}_N \hat{u}^k_N(\vartheta) = \sqrt{N} \sum_{k=1}^{K} B^{*;k} \hat{u}^k_N(\vartheta) + \varepsilon_2,$$

де $\varepsilon_2 := \sqrt{N} \sum_{k=1}^{K} (\hat{B}^{*;k}_N - B^{*;k}) \hat{u}^k_N(\vartheta)$. Отже, $\sqrt{N}(\check{\vartheta}_N - \vartheta) = -\sqrt{N} \sum_{k=1}^{K} B^{*;k} \hat{u}^k_N(\vartheta) + \varepsilon_1 + \varepsilon_2$.

Знайдемо ε_2. За побудовою, $B^* \mathbb{V}(\vartheta) = E_0$, $\hat{B}^*_N \mathbb{V}(\vartheta) = E_0$.

Звідси $B^{*;k} \bar{u}^k(\vartheta) = \mathbb{O}_d$, $\hat{B}^{*;k} \bar{u}^k(\vartheta) = \mathbb{O}_d$, де $\bar{u}(\vartheta) := \int u^k(x;\vartheta) f_k(x;\vartheta) \mu(dx)$, $k = 1, \ldots, K$.

Отже, $\varepsilon_2 = \sum_{k=1}^{K} \sqrt{N}(\hat{B}^{*;k}_N - B^{*;k})(\hat{u}^k_N(\vartheta) - \bar{u}^k(\vartheta))$. Оскільки $\hat{B}^{*;k}_N \to B^{*;k}$ за ймовірністю, а $\mathsf{E}\left[\left(\sqrt{N}(\hat{u}^k_N(\vartheta) - \bar{u}^k(\vartheta))\right)^2\right]$, маємо $\varepsilon_2 = o_p(1)$ при $N \to \infty$.

Знайдемо ε_1. За побудовою B^* маємо

$$\mathsf{E}\left[\sum_{k=1}^{K} B^{*;k} \frac{\partial}{\partial t} \hat{u}^k(t)\bigg|_{t=\vartheta}\right] = \sum_{k=1}^{K} \int B^{*;k} \frac{\partial}{\partial t} u^k(t)\Big|_{t=\vartheta} F_k(dx;\vartheta) = V(\vartheta; \{B^{*;k}\}) = \mathbb{I}_{d\times d}.$$

За лемою 3 з [15]

$$\hat{B}^{*;k}_N \frac{\partial}{\partial t} \hat{u}^k(t)\Big|_{t=\zeta} - B^{*;k} \frac{\partial}{\partial t} \hat{u}^k(t)\Big|_{t=\vartheta} \to \mathbb{O}_{d\times d}$$

за ймовірністю, $k = 1, \ldots, K$. Отже, $\varepsilon_1 = o_p(1)$ при $N \to \infty$.

Маємо: $\sqrt{N}(\check{\vartheta}_N - \vartheta) = -\sqrt{N} \sum_{k=1}^{K} B^{*;k} \hat{u}^k_N(\vartheta) + o_p(1)$.

Аналогічно лемі 3.3, $\sqrt{N} \sum_{k=1}^{K} B^{*;k} \hat{u}^k_N(\vartheta) \xrightarrow{W} \mathcal{N}(\mathbb{O}_d, D[\{B^{*;k}\}])$.

Отже, $\sqrt{N} \cdot (\check{\vartheta}_N - \vartheta) \xrightarrow{W} \mathcal{N}(\mathbb{O}_d, D[\{B^{*;k}\}])$. □

4. Моделювання

Модель суміші з прикладу 2.4 було змодельовано за допомогою пакета Mathematica. Симулювались вибірки різного розміру. Кількість спостережень була взята як 50, 100, 250, 500, 750, 1000, 2000 та 5000. Серед реалізованих оцінок були моментні оцінки, покращені моментні оцінки (див. приклад 2.4), та адаптивні оцінки (див. приклад 3.7), які будувались на основі покращених моментних. Робота оцінок була перевірена за допомогою багаторазового перегенерування вибірки (2000 разів для кожного випадку). Результати моделювання зображені на рисунку 1 в додатку.

Якість оцінок визначалась через порівняння їх середньоквадратичного відхилення (Mean Squared Error; MSE), помноженого на кількість спостережень у вибірці. Варто зазначити, що адаптивні оцінки іноді давали значні викиди у вибірках розміру до 500 спостережень включно. Тому, окрім MSE, розкид оцінок був оцінений через $\mathrm{RobVar} := N \cdot (\hat{Q}_N(3/4) - \hat{Q}_N(1/4))^2/\gamma^2$, де N — кількість спостережень у вибірці, $\hat{Q}_N(\alpha)$ — оцінка квантиля рівня α, γ — інтерквартильний розмах стандартного нормального розподілу (приблизно 1.34898).

Результати моделювання показали, що величина $N \cdot \mathrm{MSE}$ прямує до значення відповідного значення коефіцієнта розсіяння для всіх оцінок при зростанні N. Це свідчить про те, що зміщенням оцінок можна знехтувати. Адаптивні оцінки вийшли кращими за моментні при розмірах вибірки більше 500 за умови відсутності їх викиду. Тому на практиці адаптивні оцінки рекомендується застосовувати при середніх та великих розмірах вибірки. Для невеликих вибірок (близько 500 спостережень) значення адаптивної оцінки варто використовувати, якщо воно не сильно відрізняється від пілотної оцінки. Можливо, наприклад, побудувати довірчий інтервал для ϑ на основі пілотної оцінки (дисперсія для моментних оцінок знайдена у [2]), і перевірити, чи потрапляє туди адаптивна оцінка.

5. Висновки

Розглянута модель скінченної суміші зі змінними концентраціями, де частина компонентів описана параметрично. Запропонована техніка адаптивного оцінювання параметрів у даній моделі, яка фактично є реалізацією на практиці методу GEE-оцінок. Для введених адаптивних оцінок знайдені умови консистентності та асимптотичної нормальності, знайдені їх матриці розсіяння. Коефіцієнти розсіяння адаптивних оцінок виявились доволі близькими до нижньої межі GEE-оцінок, знайденої у [2]. Робота адаптивних оцінок доведена імітаційним моделюванням. Розроблена техніка оцінювання може бути застосована у соціологічних дослідженнях.

Література

1. А. А. Боровков, *Математическая статистика*, “Наука”, Москва, 1984.
2. О. В. Доронін, *Нижня межа матриці розсіяння для семіпараметричного оцінювання у моделі суміші*, Теорія ймовір. та матем. статист. (прийнято до друку)
3. О. В. Доронін, *Робастні оцінки для сумішей з гауссовою компонентою*, Вісник Київського національного університету імені Тараса Шевченка. Серія фізико-математичні науки (2012), №1, 18–23.
4. А. Лодатко, Р. Майборода, *Адаптивна моментна оцінка параметру розподілу по спостереженнях з домішкою*, Теорія ймовір. та матем. статист. **75** (2006), 61–70.
5. Р. Є. Майборода, О. В. Сугакова, *Оцінювання та класифікація за спостереженнями із суміші*, “Київський університет”, Київ, 2008.
6. Д. І. Похилько, *Вейвлет-оцінки щільності по спостереженням з суміші*, Теорія ймовір. та матем. статист. **70** (2004), 121–130.
7. А. М. Щербіна, *Оцінювання середнього у моделі суміші зі змінними концентраціями*, Теорія ймовір. та матем. статист. **84** (2011), 142–154.
8. А. М. Щербіна, *Оцінювання параметрів біноміального розподілу у моделі суміші*, Теорія ймовір. та матем. статист. **86** (2012), 182–192.

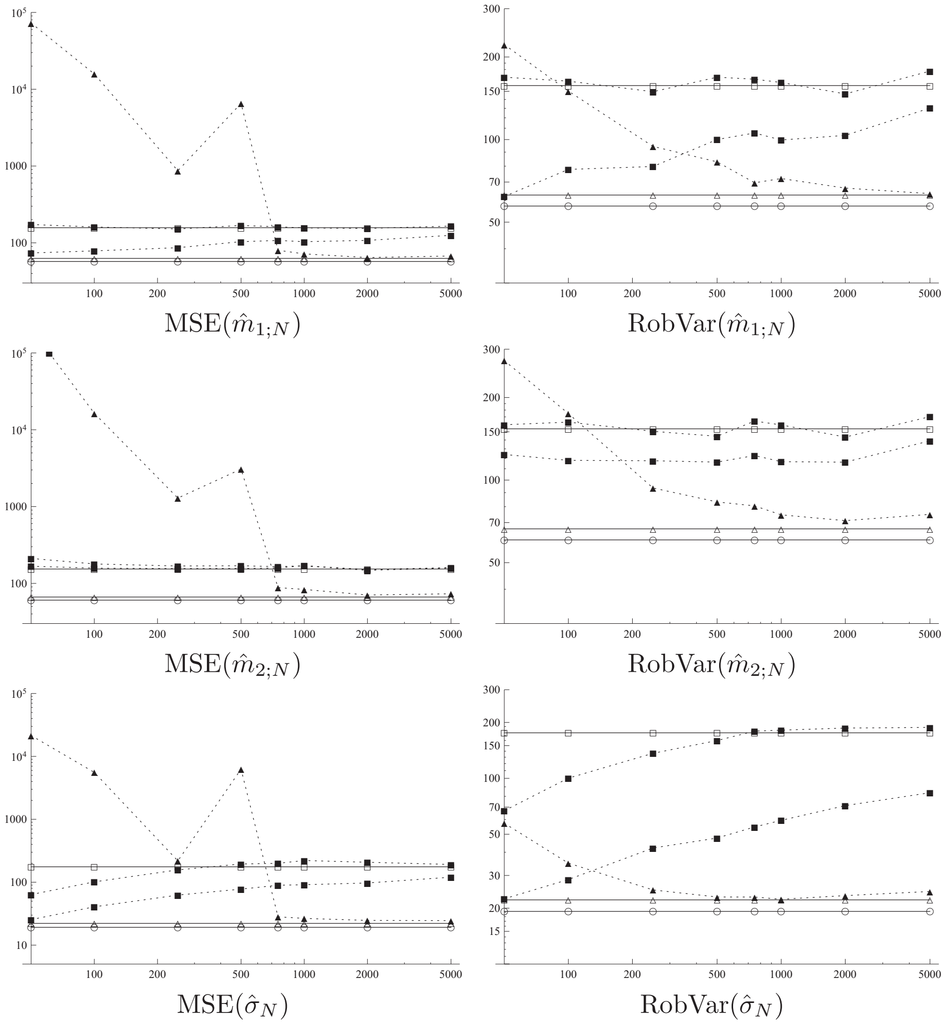

Рис. 1. Розсіяння оцінок. MSE — середньоквадратичне відхилення оцінки, помножене на кількість спостережень. RobVar — оцінка дисперсії через інтерквартильних розмах, помножена на кількість спостережень. Символ □ відповідає теоретичній дисперсії моментної оцінки, △ — адаптивної, ○ — нижній межі. Символи ■, ▲ позначають MSE чи RobVar. Верхня лінія з ■ відповідає непокращеній моментній оцінці, нижня — покращеній.

9. F. Autin and Ch. Pouet, *Test on the components of mixture densities*, Statistics & Risk Modelling **28** (2011), №4, 389–410.
10. L. Bordes, C. Delmas, and P. Vandekerkhove, *Semiparametric estimation of a two-component mixture model where one component is known*, Scandinavian Journal of Statistics **33** (2006), 733–752.
11. P. Hall and X.-H. Zhou, *Nonparametric estimation of component distributions in a multivariable mixture*, Annals of Statistics **31** (2003), №1, 201–224.
12. R. E. Maiboroda and O. O. Kubaichuk, *Improved estimators for moments constructed from observations of a mixture*, Theor. Probability and Math. Statist. **70** (2005), 83–92.
13. R. Maiboroda and O. Sugakova, *Nonparametric density estimation for symmetric distributions by contaminated data*, Metrica **75** (2012), №1, 109–126.

14. R. Maiboroda and O. Sugakova, *Statistics of mixtures with varying concentrations with application to DNA microarray data analysis*, Journal of nonparametric statistics **24** (2012), №1, 201–205.
15. R. E. Maiboroda, O. V. Sugakova, and A. V. Doronin, *Generalized estimating equations for mixtures with varying concentrations*, The Canadian Journal of Statistics **41** (2013), №2, 217–236.
16. G. J. McLachlan and D. Peel, *Finite Mixture Models*, Wiley, New York, 2000.
17. J. Shao, *Mathematical Statistics*, Springer-Verlag, New York, 1998.
18. O. Sugakova, *Adaptive estimates for the parameter of a mixture of two symmetric distributions*, Theor. Probability and Math. Statist. **82** (2011), 149–159.
19. O. Sugakova, *Empirical Bayesian classification for observations with admixture*, Theor. Probability and Math. Statist. **84** (2012), 165–172.
20. D. M. Titterington, A. F. Smith, and U. E. Makov, *Analysis of Finite Mixture Distributions*, Wiley, New York, 1985.

Кафедра теорії ймовірностей, статистики та актуарної математики, механіко-математичний факультет, Київський національний університет ім. Тараса Шевченка, проспект Глушкова, 6, Київ 03127, Україна

Адреса електронної пошти: al_doronin@ukr.net

Надійшла 04/03/2014