Теорія ймовір. та матем. статист.
Вип. 93, 2015, стор. 115–126

Teor. Ĭmovir. ta Matem. Statyst.
No. 93, 2015, pp. 115–126

АДАПТИВНИЙ ТЕСТ ДЛЯ ПЕРЕВІРКИ РІВНОСТІ МАТЕМАТИЧНИХ СПОДІВАНЬ ЗА СПОСТЕРЕЖЕННЯМИ ІЗ СУМІШІ

УДК 519.21

Р. Є. МАЙБОРОДА І О. В. СУГАКОВА

Анотація. Розглянуто задачу перевірки рівності математичних сподівань двох компонентів суміші зі змінними концентраціями. Побудовано адаптивний тест, який мінімізує асимптотичну імовірність помилки другого роду на локальних альтернативах.

Abstract. Testing of means homogeneity for two components of mixture with varying mixing probabilities is considered. An adaptive test is constructed which minimizes the asymptotic second type error for local alternatives.

Аннотация. Рассматривается задача проверки равенства математических ожиданий двух компонентов смеси с переменными концентрациями. Построен адаптивный критерий минимизирующий асимптотическую вероятность ошибки второго рода на локальных альтернативах.

1. Вступ

Моделі сумішей кількох імовірнісних розподілів часто використовують для опису статистичних даних біологічних та соціологічних досліджень [11, 9]. При аналізі таких даних природно виникає задача перевірки гіпотез про рівність певних імовірнісних характеристик (математичних сподівань, дисперсій, функцій розподілу) різних компонентів суміші. У даній роботі розглядається задача перевірки рівності певних функціональних моментів двох компонентів суміші. При цьому вважається, що концентрації компонентів (імовірності змішування) у суміші є відомими і можуть бути різними для різних спостережень. Така задача розглядалась у роботі [10], де запропонований тест, що використовує техніку мінімаксного оцінювання відповідних моментів.

Природним узагальненням цього підходу є використання на роль статистики тесту лінійної комбінації спостережень з деякими ваговими коефіцієнтами $\mathbf{b}$. Різні варіанти $\mathbf{b}$ породжують тести різної потужності (при заданому рівні значущості). Тому для вибору оптимальних $\mathbf{b}$ ми проводимо дослідження потужності тестів на локальних альтернативах і обираємо $\mathbf{b}^{\mathrm{opt}}$, при яких асимптотична потужність найбільша. Оскільки $\mathbf{b}^{\mathrm{opt}}$ залежить від невідомих розподілів компонентів, остаточний адаптивний тест використовує їх оцінки $\hat{\mathbf{b}}$. Основний результат роботи полягає в тому, що тест на основі $\hat{\mathbf{b}}$ має таку ж асимптотичну потужність, як і гіпотетичний тест на основі теоретично оптимальних $\mathbf{b}^{\mathrm{opt}}$.

Результати роботи [10] узагальнюються у [8] на випадок перевірки рівності моментів більше ніж двох компонентів. (У цій роботі при побудові тестів використовуються мінімаксні оцінки моментів). Тест для рівності розподілів двох компонентів на основі вейвлет-оцінок щільності розподілу розглянуто у [7]. Тести однорідності

Ключові слова і фрази. адаптивні алгоритми, локальні альтернативи, моделі сумішей зі змінними концентраціями, тест однорідності середніх.

Статтю підготовлено за матеріалами доповіді на міжнародній конференції “Probability, Reliability and Stochastic Optimization (PRESTO-2015)”, 7–10 квітня 2015, Київ, Україна.

сумішей для аналізу цензурованих даних запропоновано у [5]. Сучасним задачам аналізу сумішей зі змінними концентраціями присвячені також роботи [3, 6].

Далі у роботі постановці задачі присвячено п. 2. У п. 3 вміщені допоміжні відомості про оцінювання функціональних моментів за спостереженнями із суміші зі змінними концентраціями. Загальний опис тестів з довільними невипадковими ваговими коефіцієнтами $\mathbf{b}$ міститься у п. 4. Дослідження асимптотичної потужності таких тестів на локальних альтернативах проведено у п 5. У п. 6 побудовано адаптивний тест з оціненими оптимальними коефіцієнтами і сформульована основна теорема про його оптимальність. П. 7 присвячений результатам імітаційного моделювання, у яких потужність адаптивного та неадаптивного тестів порівнюється на вибірках фіксованого обсягу. Технічні доведення винесені в додаток.

2. Постановка задачі

У моделі суміші зі змінними концентраціями статистичні дані являють собою набір $X_n = (\xi_{1;n}, \dots, \xi_{n;n})$ спостережень $\xi_{j;n}$, що є незалежними випадковими елементами деякого вимірного простору $\mathcal{X}$ з розподілом

$$\mathsf{P}\{\xi_{j;n} \in A\} = \Psi_{j;n}(A) = \sum_{i=1}^{M} p^i_{j;n} F_i(A), \tag{1}$$

де F_i — розподіл i-того компоненту суміші, $p^i_{j;n}$ — концентрація i-того компоненту під час j-того спостереження (змішуюча імовірність, див. [4]). У даній роботі вважається, що концентрації компонентів є відомими, а розподіли — повністю невідомими.

Нехай $g\colon \mathcal{X} \to \mathbb{R}$ — деяка вимірна функція, така, що відповідні функціональні моменти компонентів

$$\bar{g}_i = \int_{\mathcal{X}} g(x)\, F_i(dx), \qquad i = 1, \dots, M,$$

є скінченними. В роботі [10] розглянуто задачу перевірки гіпотези про рівність функціональних моментів двох компонентів суміші. Не обмежуючи загальності можна вважати, що перевіряється гіпотеза $H_0\colon \bar{g}_1 = \bar{g}_2$. Запропонований у [10] тест базується на статистиці, що являє собою навантажене середнє вигляду

$$S_n = S_n(\mathbf{b}) = \frac{1}{n} \sum_{j=1}^{n} b_{j;n} g(\xi_{j;n}), \tag{2}$$

де $b_{j;n} = b^{\mathrm{naive}}_{j;n}$ — наївні вагові коефіцієнти, що визначаються як різниця мінімаксних вагових коефіцієнтів (3) для оцінки $\bar{g}_1$ та $\bar{g}_2$. У тесті стьюдентизована статистика S порівнюється з порогом, що забезпечує заданий асимптотичний рівень значущості тесту α.

Крім $b^{\mathrm{naive}}_{j;n}$ у (2) можна використовувати також інші вагові коефіцієнти, які задовольняють природну умову $\mathsf{E}\, S_n(\mathbf{b}) = 0$ при виконанні H_0. Природно виникає задача вибору оптимальних таких коефіцієнтів. Цю задачу і буде розглянуто далі.

3. Оцінювання моментів у моделі суміші зі змінними концентраціями

Перш ніж описувати побудову тестів, нагадаємо деякі основні підходи до оцінювання моментів за спостереженнями з суміші зі змінними концентраціями.

Надалі для трикутних масивів $\mathbf{b}$, $\mathbf{p}^i = (p^i_{j:n}, j = 1, \dots, n; n = 1, 2, \dots)$ та подібних їм кутовими дужками $\langle \cdot \rangle_n$ позначається усереднення по всіх спостереженнях. При цьому операції піднесення до ступеня, додавання та множення всередині дужок слід розуміти поелементно:

$$\langle \mathbf{b} \rangle_n = \frac{1}{n}\sum_{j=1}^{n} b_{j:n}, \qquad \langle \mathbf{p}^i\mathbf{p}^k \rangle_n = \frac{1}{n}\sum_{j=1}^{n} p^i_{j:n}p^k_{j:n},$$

тощо. Позначимо $\langle \mathbf{b} \rangle = \lim_{n\to\infty}\langle \mathbf{b} \rangle_n$ якщо ця границя існує. Введемо у розгляд матрицю $\Gamma_n = (\langle \mathbf{p}^i\mathbf{p}^k \rangle_n)_{i,k=1}^M$, котру можна трактувати як матрицю Грама для системи масивів $(\mathbf{p}^1,\dots,\mathbf{p}^M)$ у скалярному добутку $\langle \mathbf{p}^i,\mathbf{p}^k \rangle_n = \langle \mathbf{p}^i\mathbf{p}^k \rangle_n$. Відповідно $\Gamma = \lim_{n\to\infty}\Gamma_n$.

Нехай $\det\Gamma_n \neq 0$. Мінімаксні вагові коефіцієнти для оцінювання розподілу i-того компоненту суміші визначаються як

$$a^i_{j;n} = \frac{1}{\det\Gamma_n}\sum_{m=1}^{M}(-1)^{m+i}\gamma_{mi;n}p^m_{j;n}. \tag{3}$$

де $\gamma_{mi;n}$ — (m,i)-тий мінор матриці Γ_n. (Про властивість мінімаксності цих коефіцієнтів див. [4], п. 2.1.) Відповідна оцінка для $\bar{g}_i$ визначається як

$$\hat{g}_{i;n} = \frac{1}{n}\sum_{j=1}^{n} a^i_{j;n}g(\xi_{j;n}). \tag{4}$$

Асимптотичні властивості таких оцінок описані у розділі 2 [4]. Зокрема, там показано, що за досить широких умов ці оцінки є консистентними.

4. Побудова неадаптивного тесту.

Розглянемо статистику $S_n(\mathbf{b})$, визначену рівністю (2). Нехай виконується H_0 : $\bar{g}_1 = \bar{g}_2$. Тоді $\mathsf{E}\,S_n(\mathbf{b}) = 0$, якщо

$$\left\langle \mathbf{b}(\mathbf{p}^1+\mathbf{p}^2)\right\rangle_n = 0, \qquad \langle \mathbf{b}\mathbf{p}^i \rangle_n = 0 \quad \text{при } i = 2,\dots,M. \tag{5}$$

Якщо при цьому $\langle \mathbf{b}\mathbf{p}^1 \rangle_n \neq 0$, то при $\bar{g}_1 \neq \bar{g}_2$, $\mathsf{E}\,S_n(\mathbf{b}) = \langle \mathbf{b}\mathbf{p}^1 \rangle_n(\bar{g}_1 - \bar{g}_2) \neq 0$. Таким чином, близькі до 0 значення $S_n(\mathbf{b})$ свідчать на користь основної гіпотези, а її відхилення від 0 — на користь альтернативи. Щоб вибрати поріг тесту, який забезпечить заданий асимптотичний рівень значущості, розглянемо стьюдентизовану версію статистики $S_n(\mathbf{b})$. Для цього помітимо, що її дисперсія дорівнює

$$\operatorname{Var} S_n(\mathbf{b}) = \frac{1}{n}\left[\sum_{m=1}^{M}\left\langle (\mathbf{b})^2\mathbf{p}^m\right\rangle_n \overline{g^2}_m - \sum_{m_1,m_2=1}^{M}\left\langle (\mathbf{b})^2\mathbf{p}^{m_1}\mathbf{p}^{m_2}\right\rangle_n \bar{g}_{m_1}\bar{g}_{m_2}\right],$$

де

$$\overline{g^2}_m = \int (g(x))^2\,F_m(dx).$$

Оцінка $\hat{g}_{i;n}$ для $\bar{g}_i$ задається (4). Природною оцінкою для $\overline{g^2}_m$ є

$$\widehat{g^2}_{m;n} = \frac{1}{n}\sum_{j=1}^{n} a^m_{j;n}(g(\xi_{j;n}))^2. \tag{6}$$

Підставляючи ці оцінки замість справжніх значень відповідних моментів у формулу для $\operatorname{Var} S(\mathbf{b})$, отримуємо оцінку

$$D_n(\mathbf{b}) = D_n = \frac{1}{n}\left[\sum_{m=1}^{M}\left\langle (\mathbf{b})^2\mathbf{p}^m\right\rangle_n \widehat{g^2}_{m;n} - \sum_{m_1,m_2=1}^{M}\left\langle (\mathbf{b})^2\mathbf{p}^{m_1}\mathbf{p}^{m_2}\right\rangle_n \widehat{g}_{m_1;n}\widehat{g}_{m_2;n}\right]$$

Позначимо

$$D_\infty(\mathbf{b}) = \lim_{n\to\infty} n\operatorname{Var} S_n(\mathbf{b}) = \sum_{m=1}^{M}\left\langle (\mathbf{b})^2\mathbf{p}^m\right\rangle \overline{g^2}_m - \sum_{m_1,m_2=1}^{M}\left\langle (\mathbf{b})^2\mathbf{p}^{m_1}\mathbf{p}^{m_2}\right\rangle \bar{g}_{m_1}\bar{g}_{m_2}$$

(якщо ці границі існують).

Лема 4.1. *Нехай*

1. $\det\Gamma \neq 0$.
2. *Існують границі* $\langle(\mathbf{b})^2\mathbf{p}^{m_1}\mathbf{p}^{m_2}\rangle$, $m_1, m_2 = 1,\dots,M$.
3. $\overline{g^2}_m < \infty$, $m = 1,\dots,M$.
4. $D_\infty(\mathbf{b}) \neq 0$

Тоді $D_n(\mathbf{b})/\operatorname{Var} S_n(\mathbf{b}) \to 1$ *за імовірністю при* $n\to\infty$.

Визначимо стьюдентизовану статистику як

$$T_n(\mathbf{b}) = T_n = \frac{S_n(\mathbf{b})}{\sqrt{D_n(\mathbf{b})}}. \tag{7}$$

Тест на основі $T_n(\mathbf{b})$ для перевірки H_0 має вигляд

$$\pi_{\mathbf{b}}(X_n) = \mathbb{1}\{|T_n(\mathbf{b})| > \lambda_{\alpha/2}\},$$

де $\lambda_{\alpha/2}$ — квантиль рівня $1-\alpha/2$ стандартного нормального розподілу. Тобто тест $\pi_{\mathbf{b}}(X_n)$ приймає гіпотезу H_0 якщо $|T_n(\mathbf{b})| \le \lambda_{\alpha/2}$, і відхиляє її, якщо $|T_n(\mathbf{b})| > \lambda_{\alpha/2}$.

Використовуючи лему 4.1 та асимптотичну нормальність $S_n(\mathbf{b})$ ([10]), отримуємо наступну теорему.

Теорема 4.1. *Нехай виконані умови 1–4 леми 4.1 і крім того,*

$$\sup_{j,n}|b_{j;n}| < \infty.$$

Тоді справедливі наступні твердження:

1. *При виконанні* H_0,

$$T_n(\mathbf{b}) \Rightarrow N(0,1) \quad \textit{при } n\to\infty.$$

2. *Асимптотичний рівень значущості тесту* $\pi_{\mathbf{b}}(X_n)$ *дорівнює:*

$$\lim_{n\to\infty} \mathsf{P}_{H_0}\{\pi_{\mathbf{b}}(X_n) = 1\} = \alpha.$$

5. Асимптотичне дослідження неадаптивних тестів при альтернативах, що зближуються.

Для дослідження асимптотичної поведінки тестів на альтернативах, що зближуються, обмежимось випадком, коли $\mathcal{X} = \mathbb{R}$, $g(x) = x$, розподіли F_i, $i = 2,\dots,M$ всіх компонентів, крім першого, не залежать від обсягу вибірки. Розподіл першого компоненту фіксований з точністю до зсуву

$$F_{1;n}(x) = F_{1;\infty}(x - v_n), \tag{8}$$

причому зсув $v_n = v/\sqrt{n}$. Тут $F_{1;\infty}$ — деякий фіксований розподіл, такий, що

$$\int x\, F_{1;\infty}(dx) = \int x\, F_2(dx) \tag{9}$$

(тобто H_0 виконується для суміші, у якій $F_1 = F_{1;\infty}$).

Відмітимо, що з практичної точки зору таке обмеження не є принциповим. Дійсно, якщо розподіл спостережуваної вибірки $X_n = (\xi_{1;n},\dots,\xi_{n;n})$ фіксованого обсягу n описується рівністю (1), то набір $(\eta_1,\dots,\eta_n)$ де $\eta_{j;n} = g(\xi_{j;n})$ також являє собою вибірку із суміші зі змінними концентраціями, причому розподіл першого компоненту можна задати у вигляді (8) з $F_{1;\infty}$, що задовольняє (9) і $v = \sqrt{n}\int g(x)(F_1(dx) - F_2(dx))$.

Таким чином, гіпотеза $H_{1;n}$ полягає в тому, що у рамках моделі (1) $F_1(x) = F_{1;n}$ задається (8) і виконано (9).

Будемо позначати $\bar g_1 = \int g(x) x\, F_{1;\infty}(dx)$.

Наступна теорема описує асимптотику імовірності помилки другого роду для тесту $\pi_{\mathbf{b}}$ на альтернативах $H_{1;n}$.

Теорема 5.1. *Нехай*

1. $\det \Gamma \neq 0$.
2. *Для всіх* $m_1, m_2 = 1, \ldots, M$ *існують* $\langle (\mathbf{b})^2 \mathbf{p}^{m_1} \mathbf{p}^{m_2} \rangle$.
3. $\overline{g^2}_i < \infty$ *для всіх* $i = 1, \ldots, M$.
4. $D_\infty(\mathbf{b}) \neq 0$.
5. $\sup_{j,n} |b_{j;n}| < \infty$.

Тоді

$$\beta(\mathbf{b}) \stackrel{\text{def}}{=} \lim_{n\to\infty} \mathsf{P}_{H_{1;n}}\{\pi_{\mathbf{b}}(X_n) = 0\} = \mathsf{P}\{|R(\mathbf{b}) + \zeta| < \lambda_{\alpha/2}\},$$

де

$$R(\mathbf{b}) = \frac{v\langle \mathbf{b}\mathbf{p}^1 \rangle}{\sqrt{D_\infty(\mathbf{b})}}, \tag{10}$$

ζ — *стандартна нормальна випадкова величина.*

6. Адаптивний тест

При заданому рівні значущості найкращим є тест з найменшою імовірністю помилки другого роду. З нашого асимптотичного дослідження випливає, що на $H_{1;n}$ найменша асимптотична імовірність помилки другого роду для тестів $\pi_{\mathbf{b}}$ досягається тоді, коли $R(\mathbf{b})$ буде максимальним. При цьому $\mathbf{b}$ повинно задовольняти умови (5). Оскільки множення $\mathbf{b}$ на константу не змінює $R(\mathbf{b})$, можна накласти додаткову умову нормування $\langle \mathbf{b}\mathbf{p}^1 \rangle = 1$. Тоді задача максимізації $R(\mathbf{b})$ зводиться до мінімізації $D_\infty(\mathbf{b})$ за умов:

$$\left\langle \mathbf{b}\mathbf{p}^1 \right\rangle = 1, \qquad \left\langle \mathbf{b}\mathbf{p}^2 \right\rangle = -1, \qquad \left\langle \mathbf{b}\mathbf{p}^i \right\rangle = 0, \qquad i = 3, \ldots, M,$$

які еквівалентні умовам (5) при виконанні умови нормування. Замінимо цю задачу “допграничним” варіантом: будемо мінімізувати $D_n(\mathbf{b})$ за умов

$$\left\langle \mathbf{b}\mathbf{p}^1 \right\rangle_n = 1, \qquad \left\langle \mathbf{b}\mathbf{p}^2 \right\rangle_n = -1, \qquad \left\langle \mathbf{b}\mathbf{p}^i \right\rangle_n = 0, \qquad i = 3, \ldots, M.$$

Використовуючи метод множників Лагранжа, отримуємо розв'язок цієї задачі умовної оптимізації:

$$\mathbf{b} = \mathbf{b}^{\text{opt}} = (b_{j;n}^{\text{opt}}(\mathbf{G}))_{j=1}^n,$$

де

$$b_{j;n}^{\text{opt}}(\mathbf{G}) = \frac{1}{d_{j;n}(\mathbf{G})} \sum_{m=1}^M \lambda_m(\mathbf{G}) p_{j;n}^m,$$

$$\mathbf{G} = (\bar{g}_1, \ldots, \bar{g}_M, \overline{g^2}_1, \ldots, \overline{g^2}_M),$$

$$d_{j;n}(\mathbf{G}, \mathbf{p}) = d_{j;n}(\mathbf{G}) = \operatorname{Var}(\xi_{j;n}) = \sum_{m=1}^M p_{j;n}^m \overline{g^2}_m - \sum_{m_1,m_2=1}^M p_{j;n}^{m_1} p_{j;n}^{m_2} \bar{g}_{m_1} \bar{g}_{m_2},$$

$$\lambda(\mathbf{G}) = (\lambda_1(\mathbf{G}), \ldots, \lambda_M(\mathbf{G}))^T = \Gamma_n^{-1}(\mathbf{G})(1, -1, 0, \ldots, 0)^T,$$

$$\Gamma_n(\mathbf{G}) = (\Gamma_{k,l;n}(\mathbf{G}))_{k,l=1}^M, \ \Gamma_{k,l;n}(\mathbf{G}) = \frac{1}{n} \sum_{j=1}^n \frac{p_{j;n}^k p_{j;n}^l}{d_{j;n}(\mathbf{G})},$$

$$\Gamma(\mathbf{G}) = \lim_{n\to\infty} \Gamma_n(\mathbf{G}).$$

(Зрозуміло, що для існування цього розв'язку повинно виконуватись $\det \Gamma_n(\mathbf{G}) \neq 0$. Якщо виконана умова $\det \Gamma(\mathbf{G}) \neq 0$, то $\det \Gamma_n(\mathbf{G}) \neq 0$ при достатньо великих n.)

Теорема 6.1. *Нехай*

1. $\overline{g^2}_i < \infty$ *для всіх* $i = 1, \dots, M$.
2. *Границя в означенні* $\Gamma(\mathbf{G})$ *існує і є скінченною,* $\det\Gamma(\mathbf{G}) \neq 0$.
3. *Для всіх* $m_1, m_2 = 1, \dots, M$ *існують* $\langle(\mathbf{b}^{\mathrm{opt}})^2\mathbf{p}^{m_1}\mathbf{p}^{m_2}\rangle$.
4. $\sup_{j,n}|b_{j;n}^{\mathrm{opt}}| < \infty$.

Тоді найменше можливе значення $\beta(\mathbf{b})$ *досягається при* $\mathbf{b} = \mathbf{b}^{\mathrm{opt}}$.

Оскільки $\mathbf{b}^{\mathrm{opt}}$ залежить від невідомих параметрів моделі $\mathbf{G}$, безпосередньо використати ці вагові коефіцієнти для побудови тесту неможливо. Адаптивний підхід полягає в тому, щоб оцінити $\mathbf{G}$ деякою оцінкою, наприклад,

$$\hat{\mathbf{G}} = \left(\hat{g}_{1;n}, \dots, \hat{g}_{M;n}, \widehat{g^2}_{1;n}, \dots, \widehat{g^2}_{M;n}\right)$$

і підставити цю оцінку замість справжніх значень у формулу для $\mathbf{G}^{\mathrm{opt}}$. В результаті отримуємо оцінені оптимальні вагові коефіцієнти $\hat{\mathbf{b}} = \mathbf{b}^{\mathrm{opt}}(\hat{\mathbf{G}})$. Адаптивний тест $\hat{\pi}(X_n)$ використовує статистику

$$\hat{T}_n = \frac{S(\hat{\mathbf{b}})}{\sqrt{D_n(\hat{\mathbf{b}})}}.$$

Тест $\hat{\pi}(X_n)$ приймає H_0 якщо $|\hat{T}_n| \le \lambda_{\alpha/2}$ і відхиляє, якщо $|\hat{T}_n| > \lambda_{\alpha/2}$.

Теорема 6.2. *Нехай виконіні умови теореми 6.1 і* $\sigma_m^2 = \overline{g_m^2} - (\overline{g_m})^2 > 0$ *для всіх* $m = 1, \dots, M$. *Тоді*

$$\beta(\hat{\mathbf{b}}) \stackrel{\mathrm{def}}{=} \lim_{n\to\infty} \mathsf{P}_{H_{1;n}}\{\hat{\pi}(X_n) = 0\} = \beta(\mathbf{b}^{\mathrm{opt}}(\mathbf{G})).$$

Таким чином, асимптотична імовірність помилки другого роду адаптивного тесту така ж, як у найкращого можливого неадаптивного.

Зауваження. Для оцінки $\mathbf{G}$ можна використовувати не тільки навантажені середні з мінімаксними коефіцієнтами, а і будь-які інші $\sqrt{n}$-консистентні оцінки. Оскільки мінімаксні вагові коефіцієнти для деяких j приймають від'ємні значення, оцінки на їх основі можуть мати незручні властивості при невеликих обсягах вибірки, зокрема оцінка для $d_{j;n} = \operatorname{Var}\xi_{j;n}$ може виявитись від'ємною. Тому при побудові оцінених оптимальних вагових коефіцієнтів доцільно використовувати виправлені оцінки для моментів, описані у [2].

7. Результати імітаційного моделювання

Порівняння якості роботи тестових алгоритмів на вибірках скінченного обсягу було проведено за допомогою імітаційного моделювання. В усіх експериментах моделювались двокомпонентні суміші зі змінними концентраціями, причому концентрації першого компоненту $p_{j;n}^i = p_j$ генерувались як незалежні рівномірно розподілені на $[0,1]$ випадкові величини. Відповідно, $p_{j;n}^2 = 1 - p_j$. Для розподілів компонентів суміші використана наступна схема. Вибирається базова функція розподілу F_0 з нульовим математичним сподіванням та одиничною дисперсією. Функція розподілу m-того компоненту визначається як $F_m(x) = F_0((x-\mu_m)/\sigma_m)$, де μ_m — математичне сподівання, σ_m^2 — дисперсія m-того компоненту. Для перевірки імовірності помилки першого роду α моделювались дані з розподілом, що відповідає основній гіпотезі, тобто $\mu_1 = \mu_2 = 0$. Для перевірки імовірності помилки другого роду β дані моделювались у відповідності до альтернативи $H_{1;n}\colon \mu_1 = v/\sqrt{n}$, $\mu_2 = 0$ (n — кількість спостережень). Таким чином, розподіл даних у різних експериментах визначається параметрами F_0, σ_1^2, σ_2^2 та v.

Табл. 1. Результати імітаційного моделювання у експерименті 1

n	β^{naive}	β^{adapt}	α^{naive}	α^{adapt}
100	0.1715	0.1376	0.0412	0.0344
250	0.2047	0.1501	0.0458	0.039
500	0.1388	0.0776	0.0449	0.0414
1000	0.1325	0.0686	0.0461	0.0422
5000	0.1173	0.0492	0.0486	0.0422
10000	0.1052	0.0411	0.0477	0.047
∞	0.104869	0.037512	0.05	0.05

У кожному експерименті для кожного обсягу вибірки n генерувались 10 000 модельованих вибірок із заданим розподілом, що відповідає H_0 і 10 000 вибірок з розподілом, що відповідає $H_{1;n}$. Для кожної вибірки гіпотеза H_0 про однорідність середніх $\mu_1 = \mu_2$ перевірялась двома тестами: наївним тестом з ваговими коефіцієнтами $\mathbf{b}^{\text{naive}} = \mathbf{a}^1 - \mathbf{a}^2$ та адаптивним тестом з коефіцієнтами $\hat{\mathbf{b}}$, запропонованим у попередньому розділі. Підраховувались відповідні частоти помилок першого (α^{naive} і α^{adapt}) та другого роду (β^{naive} і β^{adapt}) наївного (naive) та адаптивного (adapt) тестів для різних обсягів вибірки n. Отримані частоти наводяться у таблицях. У нижньому рядочку кожної таблиці наведено асимптотичні значення відповідних імовірностей $\alpha = 0.05$ (для помилок першого роду) та $\beta(\mathbf{b})$ (для помилок другого роду).

Для оцінювання невідомих параметрів $\mathbf{G}$ використовувались виправлені моментні оцінки з [2].

Експеримент 1. F_0 — стандартний нормальний розподіл, $\sigma_1^2 = 0.1$, $\sigma_2^2 = 1$ $v = 10$. При виборі значень параметрів враховувалось, що імовірність помилки другого роду адаптивного тесту помітно менша, ніж відповідна імовірність наївного, тоді, коли дисперсії компонентів сильно відрізняються одна від одної. Таким чином, у цьому експерименті можна побачити, як реалізуються теоретичні переваги адаптивного тесту над наївним при скінченних обсягах вибірки. Результати експерименту наведено у таблиці 1.

Ці результати показують, що точність наближення імовірності помилки першого роду асимптотичними формулами є невисокою, але достатньою для практичних потреб при обсягах вибірки більше 500 для обох тестів. Асимптотичні формули дають точніші наближення імовірностей помилок обох родів для наївного тесту, ніж для адаптивного. Однак виявлені в експерименті відмінності точності наближення не принципові для застосування тестів. У той же час, імовірність помилки другого роду для адаптивного тесту при всіх обсягах вибірки менша, ніж для наївного тесту. Приблизно двократна перевага адаптивного тесту за цим параметром досягається вже при $n = 500$. Асимптотичне співвідношення $\beta^{\text{naive}}/\beta^{\text{adapt}} = 2.795$.

Експеримент 2. F_0 — стандартний нормальний розподіл, $\sigma_1^2 = 0.5$, $\sigma_2^2 = 0.5$ $v = 10$. У цьому експерименті дисперсії обох компонентів обрано однаковими, тому адаптивний тест не має переваг над наївним. Оскільки адаптивний тест вимагає використання більш складної техніки оцінювання, можна очікувати, що за точністю відповідності асимптотичній теорії він буде програвати наївному. Метою експерименту була перевірка цієї гіпотези.

Результати експерименту наведено у таблиці 2.

Ці результати підтверджують меншу точність асимптотичного наближення для адаптивного тесту порівняно з наївним, однак, з точки зору прикладного застосування, це зменшення точності не є суттєвим.

Табл. 2. Результати імітаційного моделювання у експерименті 2

n	β^{naive}	β^{adapt}	α^{naive}	α^{adapt}
100	0.1838	0.1878	0.0556	0.0562
250	0.1057	0.1088	0.055	0.0561
500	0.1283	0.1287	0.0474	0.0478
1000	0.0799	0.0809	0.0513	0.0512
5000	0.0807	0.0804	0.0495	0.0493
10000	0.0903	0.0912	0.0478	0.0479
∞	0.0790173	0.0790173	0.05	0.05

Табл. 3. Результати імітаційного моделювання у експерименті 3

n	β^{naive}	β^{adapt}	α^{naive}	α^{adapt}
100	0.1872	0.1413	0.0398	0.0369
250	0.1632	0.0931	0.0412	0.0386
500	0.1807	0.0939	0.0457	0.040944
1000	0.1315	0.0586	0.051	0.0435
5000	0.1375	0.0616	0.047	0.044
10000	0.1114	0.0402	0.0452	0.0414
∞	0.104869	0.037512	0.05	0.05

Експеримент 3. F_0 — розподіл χ^2 з шістьма ступенями вільності, центрований математичним сподіванням та нормований своїм середньоквадратичним відхиленням, $\sigma_1^2 = 0.1$, $\sigma_2^2 = 1$ $v = 10$. Асимптотичні результати для цього випадку такі ж, як і у першому експерименті. Однак тепер розподіли компонентів не є симетричними і мають правий хвіст більш важкий, ніж у нормального розподілу. Мета експерименту — перевірити вплив цих особливостей на точність асимптотичних формул.

Результати експерименту наведено у таблиці 3.

Як показують ці результати, певне погіршення точності асимптотичних формул відбувається лише при малих обсягах вибірки. За помилкою другого роду адаптивний тест переважає наївний при всіх обсягах вибірки.

Варто відмітити, що хоча імовірність помилки першого роду для розглянутих тестів при малих обсягах вибірки помітно відрізнялась від номінальної 0.05, але в усіх випадках реальна частота помилок була меншою ніж номінальна імовірність, тобто тести виявили консервативну поведінку.

8. Висновки

Розглянута техніка адаптивного тестування дозволяє суттєво поліпшити імовірність помилки другого роду при перевірці однорідності середніх за спостереженнями із суміші зі змінними концентраціями. Адаптивні тести можна рекомендувати для застосування при обсягах вибірки $n \geq 500$, особливо у випадках, коли дисперсії компонентів суміші суттєво відрізняються. Але і при менших обсягах вибірки та за умови однорідності дисперсій використання адаптивного тесту не приводить до помітного збільшення імовірностей похибок порівняно з наївним тестом. Асиметрія розподілу та помірно важкі хвости не приводять до помітного погіршення роботи адаптивного тесту.

Додаток

Доведення леми 4.1. За теоремою 3.1.1 з [4] з умов 1 і 3 леми випливає, що $\widehat{g^2}_{m;n} \to \overline{g^2}_m$, $\hat{g}_{m;n} \to \hat{g}_m$ за імовірністю. З умови 2 леми отримуємо $\langle(\mathbf{b})^2\mathbf{p}^m\rangle_n \to \langle(\mathbf{b})^2\mathbf{p}^m\rangle$ і $\langle(\mathbf{b})^2\mathbf{p}^{m_1}\mathbf{p}^{m_2}\rangle_n \to \langle(\mathbf{b})^2\mathbf{p}^{m_1}\mathbf{p}^{m_2}\rangle$. Тому

$$\frac{D_n(\mathbf{b})}{\operatorname{Var} S_n(\mathbf{b})} = \frac{nD_n(\mathbf{b})}{n \operatorname{Var} S_n(\mathbf{b})} \to \frac{D_\infty(\mathbf{b})}{D_\infty(\mathbf{b})} = 1. \qquad \square$$

Для подальших доведень нам знадобиться наступна спеціальна конструкція аналізованих даних.

Розглянемо набір випадкових величин η_j^m, $\kappa_{j;n}$ $j = 1, 2, \ldots, m = 1, \ldots, M$, незалежних в сукупності з розподілами $\eta_j^1 \sim F_{1,\infty}$, $\eta_j^m \sim F_m$, $m = 2, \ldots, M$, $\mathsf{P}\{\kappa_{j;n} = m\} = p_{j;n}^m$. Позначимо $\delta_{j;n}^m = \mathbb{1}\{\kappa_{j;n} = m\}$.

Випадкові величини

$$\tilde{\xi}_{j;n} = \sum_{m=1}^{M} \delta_{j;n}^m \eta_j^m \tag{11}$$

мають розподіл (1), що відповідає H_0, а розподіл даних $X_n = (\xi_{1;1}, \ldots, \xi_{n;n})$, де

$$\xi_{j;n} = \tilde{\xi}_{j;n} + \delta_{j;n}^1 v_n, \tag{12}$$

відповідає локальній альтернативі $H_{1;n}$.

Таким чином, для підрахунку імовірностей, пов'язаних з тестами, що розглядаються, можна вважати, що при виконанні альтернативи дані задаються формулою (12). Основній гіпотезі відповідає (12) з $v_n = 0$. Надалі ми будемо використовувати саме таку структуру даних.

Доведення теореми 5.1. Нагадаємо, що тепер, на відміну від ситуації попередньої леми, ми припускаємо, що виконана альтернатива.

Позначимо $\tilde{S}_n(\mathbf{b}) = \frac{1}{n}\sum_{j=1}^n b_{j;n}\tilde{\xi}_{j;n}$. За теоремою 3.1.2 з [4],

$$\sqrt{n}\tilde{S}_n(\mathbf{b}) \Rightarrow N(0, D_\infty(\mathbf{b})).$$

Розглянемо

$$\sqrt{n}(S_n(\mathbf{b}) - \tilde{S}_n(\mathbf{b})) = \frac{1}{\sqrt{n}}\sum_{j=1}^n \delta_{j;n}^1 b_{j;n}\frac{v}{\sqrt{n}} = v\langle\delta^1\mathbf{b}\rangle_n \to v\langle\mathbf{b}\mathbf{p}^1\rangle$$

за законом великих чисел (теорема 3 з п.3 розділу 8 [1]).

Отже,

$$\sqrt{n}S_n(\mathbf{b}) \Rightarrow N\left(\langle\mathbf{b}\mathbf{p}^1\rangle, D_\infty(\mathbf{b})\right). \tag{13}$$

Оскільки $\hat{g}_{m;n} - \frac{1}{n}\sum_{j=1}^n b_{j;n}\tilde{\xi}_{j;n} = v_n\frac{1}{n}\sum_{j=1}^n b_{j;n}\delta_{j;n}^1 \to 0$ за імовірністю, то $\hat{g}_{m;n} \to \bar{g}_m$. Аналогічно, $\widehat{g^2}_{m;n} \to \overline{g^2}_m$ за імовірністю. Тому $D_n(\mathbf{b})/D_\infty(\mathbf{b}) \to 1$ при $n \to \infty$. Звідси та з (13) випливає, що

$$T_n(\mathbf{b}) \Rightarrow N\left(\frac{\langle\mathbf{b}\mathbf{p}^1\rangle}{\sqrt{D_\infty(\mathbf{b})}}, 1\right). \tag{14}$$

Оскільки $\{\pi_{\mathbf{b}}(X_n) = 0\} = \{|T_n(\mathbf{b})| < \lambda_{\alpha/2}\}$, то з (14) випливає твердження теореми. $\square$

Доведення теореми 6.2. Покажемо, що

$$n(D_n(\hat{\mathbf{b}}) - D_n(\mathbf{b}^{\text{opt}})) \to 0 \tag{15}$$

і
$$\Delta_n(\hat{\mathbf{G}}) \stackrel{\text{def}}{=} \sqrt{n}\big(S_n(\mathbf{b}) - S_n(\hat{\mathbf{b}})\big) = \frac{1}{n}\sum_{j=1}^{n}(b_{j;n}^{\text{opt}}(\mathbf{G}) - b_{j;n}^{\text{opt}}(\hat{\mathbf{G}}))\xi_{j;n} \to 0 \tag{16}$$
за імовірністю при $n \to \infty$. Звідси випливатиме твердження теореми.

Доведемо (15). Позначимо $\sigma_{min}^2 = \min_m \sigma_m^2$. Тоді
$$\begin{aligned} d_{j;n}(\mathbf{G}) &= \operatorname{Var}(\xi_{j;n}) = \mathsf{E}\operatorname{Var}(\xi_{j;n} \mid \kappa_{j;n}) + \operatorname{Var}\mathsf{E}(\xi_{j;n} \mid \kappa_{j;n}) \\ &\geq \mathsf{E}\operatorname{Var}(\xi_{j;n}|\kappa_{j;n}) = \sum_{m=1}^{M} p_{j;n}^{m}\sigma_m^2 \geq \sigma_{min}^2. \end{aligned} \tag{17}$$

Крім того,
$$\begin{aligned} &\sup_{\mathbf{p}}\big|d_{j;n}(\mathbf{G},\mathbf{p}) - d_{j;n}(\hat{\mathbf{G}},\mathbf{p})\big| \\ &\quad\leq \sup_{\mathbf{p}}\left|\sum_{m=1}^{M} p_{j;n}^{m}(\overline{g_m^2} - \widehat{g^2}_{m;n})\right| + \sup_{\mathbf{p}}\left|\sum_{m_1,m_2=1}^{M} p_{j;n}^{m_1}p_{j;n}^{m_2}(\bar{g}_{m_1}\bar{g}_{m_2} - \hat{g}_{m_1;n}\hat{g}_{m_2;n})\right| \\ &\quad\leq C\left(|\mathbf{G}|\cdot|\mathbf{G}-\hat{\mathbf{G}}| + |\mathbf{G}-\hat{\mathbf{G}}|^2\right), \end{aligned}$$
де C — константа, що не залежить від $\mathbf{G}$, $\hat{\mathbf{G}}$ та $\mathbf{p}$.

В умовах теореми $\varepsilon_n \stackrel{\text{def}}{=} |\hat{\mathbf{G}} - \mathbf{G}| \to 0$ при $n \to \infty$ за імовірністю.

Оцінимо
$$\begin{aligned} \Big|\Gamma_{kl;n}(\mathbf{G}) - \Gamma_{kl;n}(\hat{\mathbf{G}})\Big| &\leq \frac{1}{n}\sum_{j=1}^{n}\frac{p_{j;n}^{k}p_{j;n}^{l}|d_{j;n}(\mathbf{G}) - d_{j;n}(\hat{\mathbf{G}})|}{d_{j;n}(\mathbf{G})d_{j;n}(\hat{\mathbf{G}})} \\ &\leq \frac{1}{n}\sum_{j=1}^{n}\frac{C(|\mathbf{G}|\varepsilon_n + \varepsilon_n^2)}{\sigma_{min}^2(\sigma_{min}^2 - C(|\mathbf{G}|\varepsilon_n + \varepsilon_n^2))} \to 0 \end{aligned}$$
при $n \to \infty$. Отже $\Gamma(\hat{\mathbf{G}}) \to \Gamma(\mathbf{G})$ за імовірністю. Враховуючи $\det\Gamma(\mathbf{G}) \neq 0$, отримуємо $\Gamma^{-1}(\hat{\mathbf{G}}) \to \Gamma^{-1}(\mathbf{G})$ і $\lambda(\hat{\Gamma}) \to \lambda(\Gamma)$ при $n \to \infty$ за імовірністю. Звідси випливає (15).

Доведемо (16). За побудовою $\mathbf{b}^{\text{opt}}$ для будь-якого $\tilde{\mathbf{G}} = (\tilde{G}_1, \dots, \tilde{G}_{2M})$
$$\left\langle \mathbf{b}^{\text{opt}}(\tilde{\mathbf{G}})\mathbf{p}^1 \right\rangle_n = -\left\langle \mathbf{b}^{\text{opt}}(\tilde{\mathbf{G}})\mathbf{p}^2 \right\rangle_n, \qquad \left\langle \mathbf{b}^{\text{opt}}(\tilde{\mathbf{G}})\mathbf{p}^m \right\rangle_n = 0, \qquad m = 3, \dots, M.$$
Тому
$$\mu(\tilde{\mathbf{G}}) = \mathsf{E}\sqrt{n}\Delta_n(\tilde{\mathbf{G}}) = v\big(\langle \mathbf{b}^{\text{opt}}(\mathbf{G})\mathbf{p}^1\rangle_n - \langle \mathbf{b}^{\text{opt}}(\tilde{\mathbf{G}})\mathbf{p}^1\rangle_n\big).$$
Враховуючи, що $\hat{\mathbf{G}} \to \mathbf{G}$ за імовірністю при $n \to \infty$, так само, як при доведенні (15), отримуємо
$$\mu(\tilde{\mathbf{G}}) \to 0 \quad \text{за імовірністю при } n \to \infty. \tag{18}$$
Позначимо $\xi'_{j;n} = \xi_{j;n} - \mathsf{E}\,\xi_{j;n}$,
$$\tilde{\Delta}_n(\tilde{\mathbf{G}}) \stackrel{\text{def}}{=} \Delta_n(\tilde{\mathbf{G}}) - \mu(\tilde{\mathbf{G}}) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\Big(b_{j;n}^{\text{opt}}(\mathbf{G}) - b_{j;n}^{\text{opt}}(\tilde{\mathbf{G}})\Big)\,\xi'_{j;n}.$$
Нехай $\varepsilon > 0$. Позначимо $G^k = (G_1^k, \dots, G_{2M}^k)$, $k = 0, 1$, де $G_i^0 = G_i - \varepsilon$, $G_i^1 = G_i + \varepsilon$,
$$K^{\varepsilon} = \bigotimes_{i=1}^{2M}[G_i^0, G_i^1].$$
Надалі ε обирається достатньо малим а n_0 — достатньо великим, щоб $\det\Gamma_n(\tilde{\mathbf{G}}) \neq 0$ для всіх $\tilde{\mathbf{G}} \in K^{\varepsilon}$ при всіх $n > n_0$.

Ми покажемо, що для всіх $\lambda > 0$,

$$P_\varepsilon(\lambda) = \sup_{n>n_0} \mathsf{P}\left\{ \sup_{\tilde{\mathbf{G}}\in K^\varepsilon} |\tilde{\Delta}_n(\tilde{\mathbf{G}})| > \lambda \right\} \to 0 \quad \text{при } \varepsilon \to 0. \tag{19}$$

Для цього скористаємось нерівністю соболевського типу, що обмежує рівномірну норму функції L_2-нормами її похідних.

Для будь-якого $\alpha = (\alpha_1, \dots, \alpha_{2M}) \in \{0,1\}^{2M}$ і будь-якого $u = (u_1, \dots, u_{2M}) \in K^\varepsilon$, позначимо $K^\varepsilon(\alpha) = \bigotimes_{i:\alpha_i=1}[G_i^0, G_i^1]$, $u|_\alpha$ — вектор у $\mathbb{R}^{2M}$, у якого i-та координата дорівнює G_i^0, якщо $\alpha_i = 0$ і дорівнює u_i, якщо $\alpha_i = 1$, $D^\alpha = \prod_{i:\ \alpha_i=1} \frac{\partial}{\partial u_i}$, $(du)^\alpha = \prod_{i:\ \alpha_i=1} du_i$. За лемою 7.1.1 з [4], для будь-якої $2M$ разів диференційовної функції $f\colon K^\varepsilon \to \mathbb{R}$

$$\sup_{u\in K^\varepsilon} |f(u)| \le f(\mathbf{G}^0) + \|f\|_H \bar{V}(\varepsilon) \tag{20}$$

де

$$\bar{V}(\varepsilon) = \left(\sum_{\alpha\in\{0,1\}^{2M},\alpha\neq 0} V_\alpha(\varepsilon) \right)^{1/2},$$

$V_\alpha(\varepsilon) = \prod_{i:\alpha_i=1}(G_i^1 - G_i^0)$,

$$\|f\|_H = \left(\sum_{\alpha\in\{0,1\},\alpha\neq 0} \int_{K^\varepsilon(\alpha)} (D^\alpha f(u|_\alpha))^2 (du)^\alpha \right)^{1/2}.$$

Для доведення (19) скористаємось (20) з $f(u) = \tilde{\Delta}(u)$.

Зрозуміло, що

$$\begin{aligned} J_1(\varepsilon) &\stackrel{\text{def}}{=} \mathsf{E}(\tilde{\Delta}_n(G^0))^2 = \operatorname{Var}\tilde{\Delta}_n(G^0) \\ &= \frac{1}{n}\sum_{j=1}^n (b_{j;n}^{\text{opt}}(\mathbf{G}) - b_{j;n}^{\text{opt}}(\mathbf{G})^0)^2 \operatorname{Var}\xi_{j;n} \\ &\le \sigma_{max}^2 \sup_{n>n_0} (b_{j;n}^{\text{opt}}(\mathbf{G}) - b_{j;n}^{\text{opt}}(\mathbf{G}^0))^2 \to 0 \end{aligned} \tag{21}$$

при $\varepsilon \to 0$.

Далі, $\mathsf{E}\,\|\tilde{\Delta}_n\|_H^2 \le \sum_{\alpha\in 0,1^{2M},\alpha\neq 0} J_\alpha(\varepsilon)$, де

$$\begin{aligned} J_\alpha(\varepsilon) &= \mathsf{E}\int_{K^\varepsilon(\alpha)} (D^\alpha \tilde{\Delta}_n(u|_\alpha))^2 (du)^\alpha \\ &= \frac{1}{n^2}\int_{K^\varepsilon(\alpha)} \mathsf{E}\left(\sum_{j=1}^n D^\alpha(b_{j;n}^{\text{opt}}(u|_\alpha)\xi'_{j;n} \right)^2 (du)^\alpha \\ &\le \max_{m=1,\dots,M} \overline{g_m^2} \sup_{j=1\dots,n;n>n_0;u\in K^\varepsilon(\alpha)} \left(D^\alpha(b_{j;n}^{\text{opt}}(u))\right)^2. \end{aligned} \tag{22}$$

Легко бачити, що для достатньо малого ε_0 і $0 < \varepsilon < \varepsilon_0$,

$$\sup_{j=1,\dots,n;n>n_0;u\in K^\varepsilon} (D^\alpha(b_{j;n}^{\text{opt}}(u))^2 < C, \tag{23}$$

де C — константа, що не залежить від ε.

Враховуючи (20-23), за нерівністю Чебишова отримуємо

$$P_\varepsilon(\lambda) \le \frac{2(J_1(\varepsilon) + C\bar{V}_\varepsilon)}{\lambda^2} \to 0$$

при $\varepsilon \to 0$.

Ця нерівність забезпечує виконання (19).

Тепер помітимо, що

$$\mathsf{P}\{|\tilde{\Delta}_n(\hat{\mathbf{G}}_n)| > \lambda\} \le \mathsf{P}\{\hat{\mathbf{G}} \notin K^\varepsilon\} + \mathsf{P}\left\{\sup_{\tilde{\mathbf{G}}\in K^\varepsilon} |\tilde{\Delta}_n(\tilde{\mathbf{G}})| > \lambda\right\}.$$

Внаслідок (19) другий доданок можна зробити як завгодно малим, обравши достатньо мале ε. Зафіксувавши ε, отримуємо, що перший доданок прямує до 0 при $n \to \infty$ внаслідок консистентності $\hat{\mathbf{G}}$ як оцінки $\mathbf{G}$. Звідси, з урахуванням (18), випливає (16).

З (15) та (16) випливає, що для будь-якого $\varepsilon > 0$,

$$J_n(\varepsilon) \stackrel{\text{def}}{=} \mathsf{P}\left\{\left|\frac{S_n(\mathbf{b}^{\text{opt}})}{\sqrt{nD_n(\mathbf{b}^{\text{opt}})}} - \frac{S_n(\hat{\mathbf{b}})}{\sqrt{nD_n(\hat{\mathbf{b}})}}\right| \ge \varepsilon\right\} \to 0, \tag{24}$$

при $n \to \infty$.

Оскільки $\beta(\mathbf{b}^{\text{opt}}) = \lim_{n\to\infty} \mathsf{P}\{|S_n(\mathbf{b}^{\text{opt}})|/\sqrt{nD_n(\mathbf{b}^{\text{opt}})} < \lambda_{\alpha/2})$ є неперервною функцією від $\lambda_{\alpha/2}$, то з (24) випливає, що $\beta(\hat{\mathbf{b}}) = \lim_{n\to\infty} \mathsf{P}\{|S_n(\hat{\mathbf{b}})|/\sqrt{nD_n(\hat{\mathbf{b}})} < \lambda_{\alpha/2}) = \beta(\mathbf{b}^{\text{opt}})$.

Теорему доведено. □

Література

1. А. А. Боровков, *Теория вероятностей*, “Наука”, Москва, 1986.
2. О. О. Кубайчук, Р. Є. Майборода, *Виправлені оцінки моментів по спостереженнях із суміші*, Теор. ймовірн. та математ. статист. **70** (2004), 74–81.
3. О. Доронін, *Адаптивне оцінювання у семіпараметричній моделі суміші*, Теор. ймовірн. та математ. статист. **91** (2014), 26–37.
4. Р. Є. Майборода, О. В. Сугакова, *Оцінювання та класифікація за спостереженнями із суміші*, ВПЦ “Київський ун-т”, Київ, 2008.
5. А. Ю. Рижов, *Перевірка гіпотези про однорідність компонент суміші зі змінними концентраціями за цензурованими даними*, Теорія ймовірностей та математична статистика **72** (2005), 129–139.
6. А. Щербіна, *Оцінювання середнього у моделі суміші зі змінними кнцентраціями*, Теорія ймовірностей та математична статистика **84** (2011), 142–154.
7. F. Autin and Ch. Pouet, *Test on the components of mixture densities*, Statistics & Risk Modelling **28** (2011), no. 4, 389–410.
8. A. Doronin, R. Maiboroda, *Testing hypotheses on moments by observations from a mixture with varying concentrations*, Modern Stochastics: Theory and Applications **1** (2014), no. 2, 195–209.
9. G. J. McLachlan and D. Peel, *Finite mixture models*, Wiley-Interscience, 2000.
10. Rostyslav Maiboroda and Olena Sugakova, *Statistics of mixtures with varying concentrations with application to DNA microarray data analysis*, Journal of Nonparametric Statistics **24** (2012), no. 1, 201–215.
11. D. M. Titterington, A. F. M. Smith, and U. E. Makov, *Statistical Analysis of Finite Mixture Distribution*, Wiley, New York, 1985.

Кафедра теорії ймовірностей, математичної статистики і актуарної математики механіко-математичний факультет, Національний університет ім. Тараса Шевченка, проспект Глушкова, 6, Київ 03127, Україна

Адреса електронної пошти: `mre@univ.kiev.ua`

Кафедра вищої математики та теоретичної радіофізики, радіофізичний факультет Національний університет ім. Тараса Шевченка, проспект Глушкова, 6, Київ 03127, Україна

Адреса електронної пошти: `sugak@univ.kiev.ua`