

У попередньому експерименті (рис. 7) фазова різниця між лазерними хвилями вібрметра була близькою до π . Це підтверджує верхній трек на рис. 7, який відображає амплітуду сигналу на частоті гетеродину. Подвоєння частоти цього сигналу і є свідченнями протифази хвиль, що інтерферують (порівняймо з рис. 5A).

Розрахунок амплітуди коливань по даним рис. 7 дає результат 30 nm, хоча в дійсності $X_1 = 6 \text{ nm}$. Тобто аномальне підсилення складає $\alpha = 5$. Точка №3 ($\alpha = 5, \eta = 1.25$), яка нанесена рис. 3, добре співпадає з результатами розрахунку.

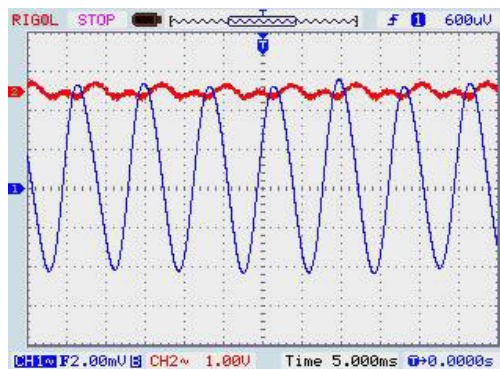


Рис. 7. Осцилограми сигналів ЛДВ. Амплітуда коливання об'єкта 6 nm, частота 120 Гц: верхня еюра – амплітуда фотоструму, нижня еюра – частота фотоструму (вихідний сигнал ЛДВ)

4. Висновки

Встановлено, що в схемі ЛДВ з двома зондувальними променями і одним опорним при умові, коли амплітуди зондувальних променів майже рівні, а початкові фази протилежні, для сигналів з амплітудою менше ніж половина довжини зондувальної хвилі спостерігається різке зростання амплітуди вихідного сигналу вібрметра. Таким чином, тільки оптичними апаратними засобами можна досягти значного "підсилення" амплітуди сигналу ЛДВ і, тим самим, підвищити чутливість до слабких акустичних сигналів.

При значному підвищенні чутливості спостерігається ріст нелінійних спотворень сигналу. Важливим результатом є те, що вразливість до нелінійних спотворень зменшується зі зменшенням амплітуди вібрації, тому запропонований метод найбільш ефективний саме для діагностики слабких акустичних сигналів.

У демонстраційному експерименті з спеціально доопрацьованим ЛДВ отримано десятикратне підвищення чутливості вібрметра до амплітуди миттєвої швидкості з рівнем нелінійних спотворень вихідного сигналу 17 dB та п'ятикратне підвищення амплітуди з рівнем нелінійних спотворень < 40 dB.

1. Яровой Л. Влияние когерентных помех на сигнал доплеровских вибрметров // Фізичні методи та засоби контролю середовищ, матеріалів та виробів "Леотест-2004". – К.; Львів. – 2004. – С. 35–39.
2. Яровой Л. К. Метод дискретного вимірювання малих переміщень за допомогою лазерного вібрметра // Вісник Київського університету. Механіка та Математика. – 2006. – Вип. 15–16. – С. 53–57.
3. Яровой Л. К. Застосування трихвильового фотозміщення для підвищення чутливості лазерного доплерівського вібрметра в субнанометровому діапазоні // Акустичний симпозиум "КОНСОНАНС-2009", Київ, 29 вересня – 1 жовтня 2009 року. – К., 2009. – С. 330–336.
4. Пат. 84937 Україна МПК G 01 B 9/02 Спосіб вимірювання вібрації / Яровой Л. К.; заявник власник патенту Київ Нац. Унів. ім. Тараса Шевченка. – № 200614074; заяв. 29.12.06; друк. 10.12.08, Бюл. № 23.
5. Dussarrat Olivier J., Clark D. F., Moir T. J. New demodulation process to reduce cochannel interference for a laser vibrometer sensing system // Proc. SPIE. Vibration Measurements by Laser Techniques, – 1998. – Vol. 3411. – P. 2–13.
6. Ombeline de La Rochefoucauld, Shyam M. Khana S., Olson E. Recording depth and signal competition in heterodyne interferometry // J. Acoust. Soc. Am. – 2005. – Vol. 117, № 3. – P. 1267–1284.
7. Yarovi L., Siegmund G. The effect of three-wave interference in laser Doppler vibrometry // Meas. Sci. Technol. – 2004. – Vol. 15. – P. 2150–2156.

Надійшла до редколегії 29.01.2011 р.

УДК 519.254.2 + 519.254

А. Щербіна, асп.

ПОРІВНЯННЯ ОЦІНОК СЕРЕДНІХ ЗНАЧЕНЬ ДЛЯ СУМІШЕЙ ЗІ ЗМІННИМИ КОНЦЕНТРАЦІЯМИ НА МОДЕЛЬОВАНИХ ДАНИХ

Розглянуто задачу оцінювання середнього значення у моделі двокомпонентної суміші із залежними спостереженнями. Описано лінійні та адаптивні непараметричні оцінки і оцінки найбільшої вірогідності для параметричної моделі. Якість цих оцінок порівнюється на модельованих вибірках.

Estimation of the mean value in the model of two-component mixture with dependent observations is considered. Linear and adaptive nonparametric estimators and maximum likelihood estimator for parametric model are described. Quality of such estimators is compared on simulated data.

1. Вступ

Задачі дослідження сумішей зі змінними концентраціями виникають в багатьох галузях природознавства. Зокрема, модель суміші природно виникає, коли об'єкти популяції, що досліджується, належать до декількох різних класів. Нехай ми досліджуємо популяцію об'єктів, вимірюючи у них певну числову характеристику X . Нехай існує M класів об'єктів, а розподіл характеристики X залежить лише від класу, до якого належить об'єкт.

Якщо при обстеженні клас, якому належить об'єкт точно не відомий, то для опису розподілу характеристики X можна використати модель суміші:

$$P(X \in A) = w^1 F_1(A) + w^2 F_2(A) + \dots + w^M F_M(A),$$

де F_l – розподіл l -ї компоненти суміші, а w^l – її концентрація у суміші (ймовірність, з якою досліджуваний об'єкт може належати до l -го класу). У класичній моделі суміші розподіли характеристик всіх об'єктів описуються наведеною вище моделлю з однаковими концентраціями компонент. Дослідження цих моделей було розпочато Ньюкомбом [7] та Пірсоном [8]. Огляд результатів у цій області можна знайти у Дж. Маклахана та Піла [6].

Останнім часом досліджуються моделі сумішей зі змінними концентраціями із незалежними спостереженнями [1-3]. У ці же статті ми будемо розглядати випадок залежних спостережень, описаний у [4]. Залежність виникає через отримання об'єктів за допомогою випадкового відбору без повернення зі скінченної популяції. Опис відповідної моделі наведено у [9]. Нехай ми хочемо визначити середню успішність учнів у школі в залежності від того, палять вони чи ні. Питання про паління є "дражливим", тому безпосереднє опитування є небажаним. Тому можна спочатку провести анонімне обстеження у класах школи, яке дозволить визначити частки курців. А потім провести вибіркове або суцільне обстеження успішності учнів.

2. Опис моделі

Ми будемо розглядати випадок двох класів $M = 2$. Нехай популяція об'єктів складається з K груп з обсягами N_1, N_2, \dots, N_K об'єктів. При цьому в i -й групі міститься N_i^1 об'єктів першого класу, та N_i^2 об'єктів другого класу. Вважаємо ці величини відомими. Отже, можна визначити концентрації об'єктів першого та другого класів за формулою

$$w_i^l = \frac{N_i^l}{N_i}, \quad i = 1, 2, \dots, K, \quad l = 1, 2.$$

Позначимо значення досліджуваної характеристики у цих об'єктів $\{X_{ij}, i = 1, 2, \dots, K, j = 1, 2, \dots, N_i\}$. Вважаємо їх незалежними випадковими величинами із невідомими розподілами, спільними для об'єктів одного класу. Нехай у цих розподілів існують скінченні другі моменти. Середні значення та дисперсії характеристики X об'єктів першого та другого класів будемо позначати μ_1, μ_2 та σ_1^2, σ_2^2 відповідно.

Далі, за допомогою простого випадкового відбору без повернення з кожної групи відбирається n_i спостережуваних об'єктів, $1 \leq n_i \leq N_i, i = 1, 2, \dots, K$. Значення характеристики X у відібраних об'єктів позначимо $\{Y_{ij}, i = 1, 2, \dots, K, j = 1, 2, \dots, n_i\}$. Іншими словами при фіксованому i множина цих характеристик є випадковою вибіркою без повернення з множини $\{X_{ij}, j = 1, 2, \dots, N_i\}$. Задача полягає у оцінюванні середніх значень характеристик X у об'єктів першого та другого класів за цими даними.

3. Опис оцінок

Ідея оцінювання полягає в тому, щоб використати відмінності концентрацій об'єктів першого та другого класів у групах. Через ці відмінності середні значення характеристик X у групах будуть відрізнятися. Для їх оцінювання будемо використовувати вибіркове середнє спостережених характеристик $T_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}$. Обчислимо математичне сподівання цієї величини:

$$E T_i = \frac{1}{n_i} \sum_{j=1}^{n_i} E Y_{ij} = E Y_{i1} = E \frac{1}{N_i} \sum_{j=1}^{N_i} X_{ij} = \frac{1}{N_i} (N_i^1 \mu_1 + N_i^2 \mu_2) = w_i^1 \mu_1 + w_i^2 \mu_2, \quad i = 1, 2, \dots, K.$$

Основними оцінками, які розглядаються у цій статті є зважені лінійні оцінки наступного вигляду:

$$\hat{\mu}_l = \frac{1}{K} \sum_{i=1}^K a_i^l T_i, \quad l = 1, 2,$$

де $a_1^l, a_2^l, \dots, a_K^l$ – певні набори вагових коефіцієнтів. Математичне сподівання цієї оцінки дорівнює

$$E \hat{\mu}_l = \frac{1}{K} \sum_{i=1}^K a_i^l E T_i = \frac{1}{K} \sum_{i=1}^K a_i^l (w_i^1 \mu_1 + w_i^2 \mu_2) = \frac{\mu_1}{K} \sum_{i=1}^K a_i^l w_i^1 + \frac{\mu_2}{K} \sum_{i=1}^K a_i^l w_i^2, \quad l = 1, 2.$$

Оскільки оцінювання для другого класу відбувається аналогічно до оцінювання для першого класу, то надалі будемо розглядати оцінки лише для першого класу, а індекс 1 у a^l опускати.

Бачимо, що оцінка $\hat{\mu}_1$ буде незміщеною при виконанні наступних умов:

$$\frac{1}{K} \sum_{i=1}^K a_i w_i^1 = 1, \quad \frac{1}{K} \sum_{i=1}^K a_i w_i^2 = 0. \tag{1}$$

Надалі будемо вважати ці умови виконаними. Проте вони не однозначно визначають вагові коефіцієнти. Отже, постає питання про вибір найкращих вагових коефіцієнтів. Природною характеристикою якості незміщеної оцінки є її дисперсія. Наведемо твердження з [4] про явний вигляд дисперсії зваженої оцінки:

Твердження 1. Дисперсія оцінки $\hat{\mu}_1$ дорівнює:

$$D \hat{\mu}_1 = \frac{1}{K^2} \sum_{i=1}^K a_i^2 d_i, \quad d_i = \frac{1}{n_i} \left(w_i^1 \sigma_1^2 + w_i^2 \sigma_2^2 + \frac{N_i - n_i}{N_i - 1} w_i^1 w_i^2 (\mu_1 - \mu_2)^2 \right), \quad i = 1, 2, \dots, K.$$

Як бачимо, дисперсія залежить від вектору невідомих параметрів $\gamma = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$. Для того, щоб побудувати оцінки, які матимуть дисперсію, близьку до мінімальної можливої, замінимо величини $\{d_i, i = 1, 2, \dots, K\}$ їх оцінками

$\{f_i, i = 1, 2, \dots, K\}$. Щоб визначити найкращі вагові коефіцієнти, будемо мінімізувати вираз $\sum_{i=1}^K a_i^2 f_i$ при умовах (1).

Введемо в розгляд наступні зважені моменти концентрацій w_i^j :

$$r_j(f) = \frac{1}{K} \sum_{i=1}^K \frac{(w_i^j)^j}{f_i}, \quad j = 0, 1, 2.$$

Застосування методу множників Лагранжа показує, що в невідродженому випадку ($r_0(f)r_2(f) - r_1^2(f) \neq 0$) найкращі ваги розраховуються за наступною формулою:

$$a_i(f) = \frac{(r_0(f) - r_1(f))w_i^1 + r_2(f) - r_1(f)}{d_i(r_0(f)r_2(f) - r_1^2(f))}, \quad i = 1, 2, \dots, K.$$

Найпростіше в якості коефіцієнтів f_i використати тотожну одиницю. Отримані коефіцієнти називаються мінімаксними і розглядаються у [2]. Будемо позначати їх $a_i(1)$.

Подальше дослідження показує, що величини d_i майже обернено пропорційно залежать від n_i . А тому природно визначити коефіцієнти f_i наступним чином $f_i = 1/n_i$, $i = 1, 2, \dots, K$. Визначені за цими коефіцієнтами ваги будемо позначати $a_i(1/n)$.

Ще однією можливістю є так звані адаптивні оцінки. Спочатку, за допомогою мінімаксних ваг ми будемо оцінку вектора параметрів $\hat{\gamma} = (\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2, \hat{\sigma}_2^2)$:

$$\hat{\mu}_l = \frac{1}{K} \sum_{i=1}^K a_i^l T_i, \quad \hat{\sigma}_l^2 = \frac{1}{K} \sum_{i=1}^K a_i^l \frac{1}{n_i} \sum_{j=1}^K Y_{ij}^2 - \hat{\mu}_l^2, \quad l = 1, 2.$$

Далі, підставляємо отримані значення у вираз для d_i , отримуючи наступну оцінку:

$$\hat{d}_i = \frac{1}{n_i} \left(w_i^1 \hat{\sigma}_1^2 + w_i^2 \hat{\sigma}_2^2 + \frac{N_i - n_i}{N_i - 1} w_i^1 w_i^2 (\hat{\mu}_1 - \hat{\mu}_2)^2 \right), \quad i = 1, 2, \dots, K.$$

Обчислені за цими величинами ваги позначимо $a_i(\hat{d}_i)$. Якщо за цими вагами знову обчислити оцінки середніх та дисперсій, і повторити процедуру адаптивного оцінювання, то отримуємо двічі адаптивні оцінки $a_i(\hat{\hat{d}}_i)$.

Також для порівняння будемо розглядати найкращі ваги, обчислені з використанням справжніх значень d_i . Будемо позначати їх $a_i(d_i)$.

4. Порівняння оцінок на модельованих вибірках

Для дослідження якості розглянутих оцінок на вибірках скінченного обсягу ми провели імітаційні експерименти, кожен з яких складався з 10000 ітерацій. Кожна ітерація виконувалась наступним чином:

1. За певними розподілами генерувалися кількості об'єктів першого та другого класів у групах N_i^1 , N_i^2 та обсяги вибірок n_i для $i = 1, 2, \dots, K$. У випадку, коли величина N_i або n_i випадали рівними нулю, то генерація повторювалася, поки не отримувалась ненульовий результат.

2. Для кожної групи визначалися кількості об'єктів першого та другого класів у вибірці за гіпергеометричним розподілом з параметрами N_i^1 , N_i^2 та n_i . Позначимо їх n_i^1 та n_i^2 відповідно.

3. Для кожної групи генерувалися n_i^1 та n_i^2 величин з розподілами, що відповідають першому та другому класам. Вони утворювали набір характеристик $\{Y_{ij}, i = 1, 2, \dots, K, j = 1, 2, \dots, n_i\}$.

4. Обчислювалися вагові коефіцієнти за різними методами, як це було описано в попередньому розділі.

5. Розраховувалися вибіркові середні (1), обчислювалися оцінки (2) для різних вагових коефіцієнтів.

На базі отриманих оцінок підраховувалося вибіркове середньоквадратичне відхилення оцінок від справжнього значення. Для порівняння середньоквадратичних відхилень при різних K ми домножали їх на K . Відмітимо, що в усіх експериментах зміщення досліджуваних оцінок виявилось малим, порівняно з дисперсією, тому графіки зміщень та дисперсій окремо не наводяться.

У наступних двох прикладах наведено конкретні розподіли випадкових величин. Для них проведено імітаційні експерименти для різних K . Назви, позначення та спосіб відображення оцінок на графіках зведений у наступній таблиці.

Коефіцієнти	Позначення	Лінії на графіках
Мінімаксні	$a_i(1)$	Пунктирна
Зважені на обсяг вибірки	$a_i(1/n)$	Штрихована
Адаптивні коефіцієнти	$a_i(\hat{d}_i)$	3 квадратами
Двічі адаптивні коефіцієнти	$a_i(\hat{\hat{d}}_i)$	3 трикутниками
Найкращі коефіцієнти	$a_i(d_i)$	Суцільна

Приклад 1. Характеристики об'єктів класів мають нормальні розподіли з параметрами (0, 1) та (1, 4). Кількості об'єктів першого та другого класів у групах генерувалися за біноміальним розподілом з 7 випробуваннями та ймові-

рністю успіху 0.5. Обсяги вибірки генерувалися за біноміальним розподілом з N_i випробуваннями та ймовірністю успіху 0.5.

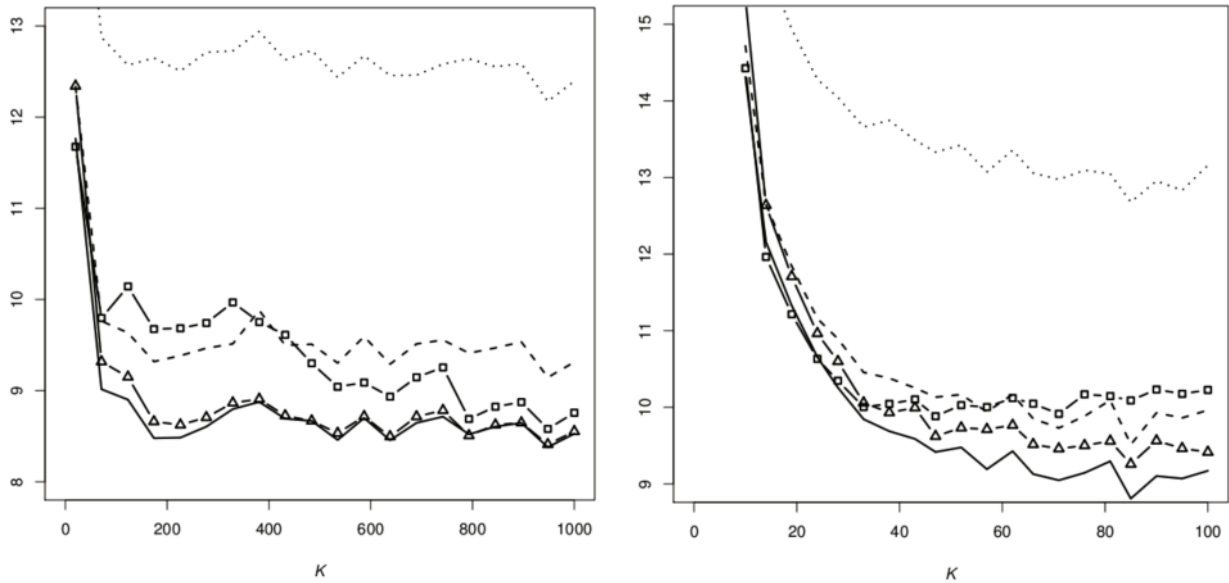


Рис. 1. Нормовані середньоквадратичні відхилення оцінок μ_1 для різних кількостей груп K

Як бачимо з рис. 1, середньоквадратичне відхилення найбільше у мінімакських оцінок. Коефіцієнти, зважені на обсяг вибірки, показують набагато кращі результати. Як і очікувалося, оцінки з найкращими коефіцієнтами мають найменше середньоквадратичне відхилення. Двічі адаптивні оцінки майже при всіх K виявилися кращими за адаптивні коефіцієнти та коефіцієнти, зважені на обсяг вибірки.

Приклад 2. Характеристики об'єктів класів мають розподіл Стюдента з трьома степенями свободи та нормальний розподіл з параметрами (1, 4). Кількості об'єктів першого та другого класів у групах генерувалися за рівномірним розподілом на множині $\{0,1,\dots,4\}$. Обсяги вибірки генерувалися за біноміальним розподілом з N_i випробуваннями та ймовірністю успіху 0.5.

Як бачимо з рис. 2, співвідношення між якістю оцінок з мінімакськими коефіцієнтами, коефіцієнтами, зважені на обсяг вибірки, та найкращими коефіцієнтами залишилось тим самим, як і в попередньому прикладі. Незвичайним є той факт, що адаптивні оцінки виявилися кращими за оцінки з найкращими коефіцієнтами. Цей ефект, можливо, виникає через наявність "важких хвостів" у розподілу Стюдента з трьома степенями свободи. Таким чином, адаптивні оцінки виявили здатність враховувати наявність викидів у вибірці при оцінюванні.

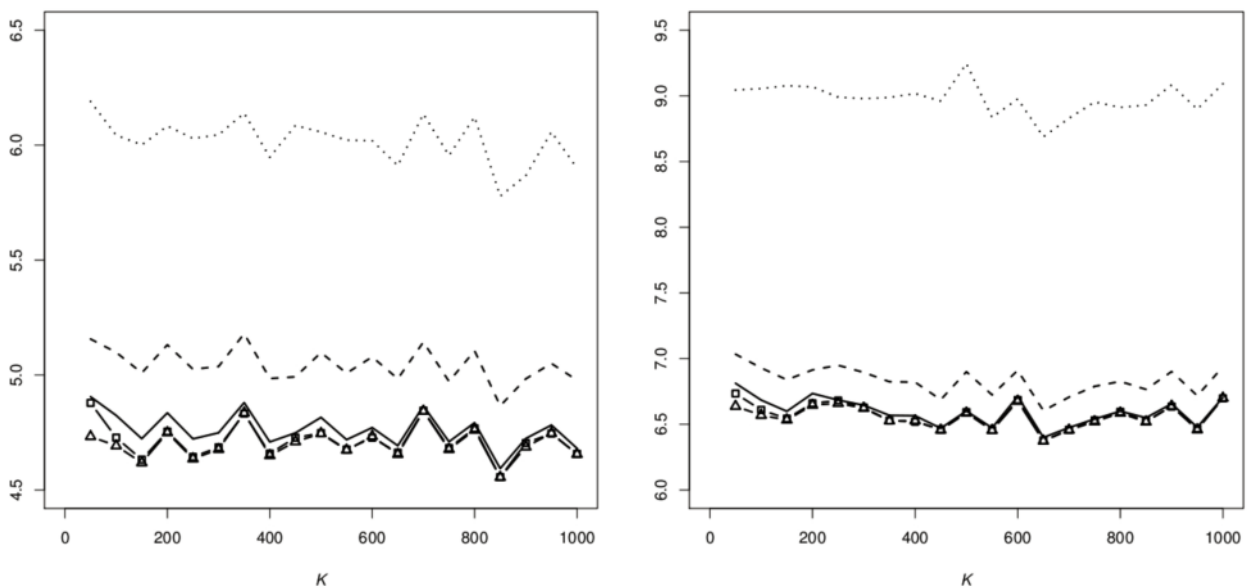


Рис. 2. Нормовані середньоквадратичні відхилення оцінок μ_1 (зліва) та μ_2 (справа) для різних кількостей груп K

Приклад 3. Розглянемо частковий випадок нашої задачі, коли характеристики об'єктів приймають тільки значення 0 та 1, причому $P(X = 1) = \mu_l$ для об'єктів, що належать l -тому класу. Спостерігаються характеристики всіх об'єктів, тобто $n_i = N_i$. До такої моделі можна застосувати оцінку максимальної вірогідності. Введемо позначення

$$Y_i = \sum_{j=1}^{N_i} Y_{ij}, \quad S_i = (Y_i, N_i^1, N_i^2), \quad i = 1, 2, \dots, K.$$

Нехай $t = (t_1, t_2)$ – вектор можливих значень невідомого параметра $\mu = (\mu_1, \mu_2)$ з множини $[0, 1]^2$, а $s = (x_1, x_2, y)$. Введемо в розгляд функцію $f(s, t) = P(Y_1 = y | N_1^1 = x_1, N_1^2 = x_2, \mu = t)$. Тепер ми можемо записати логарифмічну функцію вірогідності:

$$L(S, t) = \sum_{i=1}^K \ln f(S_i, t).$$

Оцінкою методу максимальної вірогідності буде значення параметра t , що максимізує логарифмічну функцію вірогідності:

$$\hat{\mu} = \arg \max_{t \in [0, 1]^2} L(S, t).$$

При моделюванні використовувалися наступні значення параметрів. Середні значення характеристик у класах дорівнювали $\mu_1 = 0.2$ та $\mu_2 = 0.5$. Загальна кількість об'єктів у групі була фіксована $N_i = 5$, а кількість об'єктів першого класу з рівними ймовірностями приймала значення на множині $\{0, 1, \dots, 5\}$.

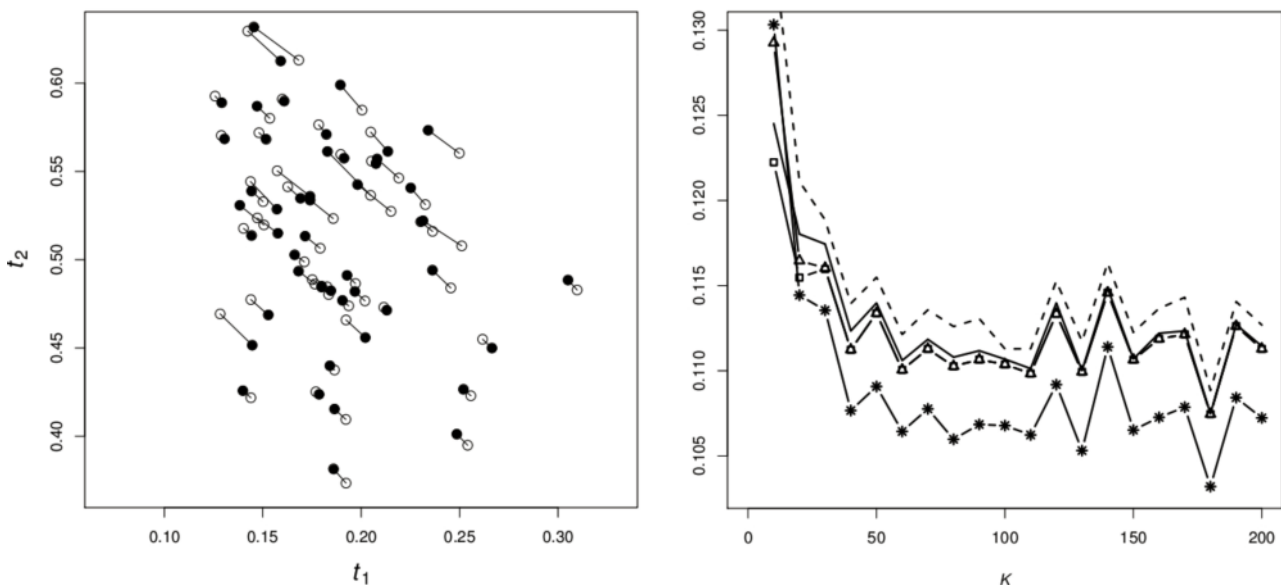


Рис. 3. Порівняння зважених оцінок та оцінок максимальної вірогідності

Для візуалізації зв'язку між зваженими оцінками та оцінками максимальної вірогідності для $K = 50$ було побудовано оцінки за обома методами. На рис. 3 зліва зображені зважені оцінки з найкращими коефіцієнтами та оцінки максимальної вірогідності. Порожні кружечки відповідають зваженим оцінкам, а повні – оцінкам максимальної вірогідності. Лінією з'єднані оцінки, що розраховані на спільних даних.

На рис. 3 справа зображені нормовані середньоквадратичні відхилення для різних оцінок. Оцінкам максимальної вірогідності відповідають лінії зі зірочками.

Таким чином, параметричні оцінки виявляються помітно кращими, ніж розглянуті нами непараметричні. Тому виникає питання про можливість поліпшення непараметричних оцінок за допомогою більш повного використання інформації, наявної у вибірці.

5. Висновки

Досліджено задачу оцінювання середнього значення у моделі суміші зі змінними концентраціями. Проведено порівняння різних оцінок на модельованих даних. Показані переваги застосування техніки адаптивного оцінювання у реальних обстеженнях. Розгляд часткового випадку з використанням параметричної моделі показав перевагу методу найбільшої вірогідності над зваженими оцінками.

1. Майборода Р. Є. Оцінка розподілів компонентів сумішей що змінюються // Укр. мат. журнал. – 1996. – Т. 48, № 4. – С. 562–566. 2. Майборода Р. Є. Статистичний аналіз сумішей. – К.: ВПЦ "Київський університет", 2003. – 176 с. 3. Майборода Р. Є., Сугакова О. В. Оцінювання та класифікація за спостереженнями із суміші. – К.: ВПЦ "Київський університет", 2008. – 213 с. 4. Щербіна А. М. Оцінювання середнього у моделі суміші зі змінними концентраціями // Теорія ймовірностей та математична статистика. 5. Kubaychuk O. O. Estimation of moments by observations from mixtures with varying concentrations // Theory of Stochastic Processes. – 2002. – Vol. 8 (24), № 3–4. – P. 226–232. 6. McLachlan G. J., Pell D. Finite Mixture Models. – NY: Wiley, 2000. 7. Newcomb S. A generalized theory of the combination of observations so as to obtain the best result // Amer. J. Math. – 1894. – Vol. 8. – P. 343–366. 8. Pearson K. Contribution to the mathematical theory of evolution // Trans. Roy. Soc. A. – 1894. – Vol. 185. – P. 71–110. 9. Shcherbina A., Maiboroda R. Merging data from anonymous and open surveys: two-population problems // Proceedings of the Baltic-Nordic-Ukrainian Summer School on Survey Statistics. – Kyiv: ТВІМС, 2009. – 177 p. 10. Titterton D. M., Smith A. F., Makov O. E. Analysis of Finite Mixture Distributions. – New York: Wiley.