УДК 519.21

О.А. Жуковська, Г.А. Глушаускене,
Л.С.Файнзільберг

ДОСЛІДЖЕННЯ ВЛАСТИВОСТЕЙ МОДИФІКОВАНОЇ ОЦІНКИ ДИСПЕРСІЇ ВИПАДКОВОЇ ВЕЛИЧИНИ ЗА ВИБІРКОЮ НЕЗАЛЕЖНИХ СПОСТЕРЕЖЕНЬ

Вступ

З теорії оцінок відомо [1], що для одного й того самого параметра можна отримати різні оцінки, спрямовані на досягнення певної мети, наприклад швидкості обчислень, стійкості (робастності), інваріантності щодо перетворення тощо. Це стосується й різних методів оцінки дисперсії випадкової величини за вибіркою незалежних спостережень, які розглядаються в багатьох наукових публікаціях, зокрема в працях [2–6].

Зрозуміло, що не існує методу обчислень, який був би однаково оптимальним у всіх випадках практичного застосування, і користувачу залишається знайти певний компроміс при виборі одного з відомих методів або розробляти свій метод під конкретну задачу.

Саме тому дослідження властивостей нових методів обчислення дисперсії випадкової величини за вибіркою незалежних спостережень досі залишається актуальною науковою задачею.

Постановка задачі

Відомо [6–8], що класичне співвідношення для обчислення незсуненої і спроможної оцінки дисперсії D має вигляд

$$\bar{D}_n = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{m})^2 , \tag{1}$$

де

$$\bar{m} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2}$$

є середнім значенням скінченного ряду незалежних спостережень $x_1, ..., x_n$.

Але обчислення оцінки дисперсії безпосередньо за формулою (1) потребує резервування комп'ютерної пам'яті для накопичення всього масиву даних $x_1, ..., x_n$, що при великій кількості спостережень обмежує практичне застосування цієї формули [5]. До того ж пряме обчислення суми великої кількості значень потенційно може призвести до переповнення розрядної сітки [6].

Ряд цих проблем вирішує рекурентна формула обчислення оцінки дисперсії [9, с. 232]

$$\bar{D}_n = \frac{n-2}{n-1}\bar{D}_{n-1} + \frac{1}{n}(x_n - \bar{m}_{n-1})^2 , \tag{3}$$

де $\bar{m}_{n-1}$ – оцінка математичного сподівання на $(n-1)$-му кроці. Проте при великому об'ємі вибірки рекурентне обчислення призводить до відомої проблеми накопичення помилок закруглення, що виникають на кожному кроці при виконанні арифметичних операцій.

Оскільки на практиці часто необхідно отримувати поточну оцінку дисперсії випадкової величини в міру накопичення даних, то замість (1) зручніше використовувати оцінку

$$\tilde{D}_n = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \tilde{m}_i)^2, \quad n = 2, 3, ... , \tag{4}$$

де $\tilde{m}_i$ – поточна оцінка математичного сподівання M випадкової величини X, обчислена на i-му кроці спостереження за формулою

$$\tilde{m}_i = \frac{1}{i}\sum_{k=1}^{i} x_k . \tag{5}$$

Однак виникає питання: наскільки правомірний перехід від стандартної (1) до модифікованої (4) оцінки.

Метою даної статті є дослідження властивостей оцінки (4) дисперсії випадкової величини та її модифікація, яка забезпечує незсуненість та спроможність поточної оцінки дисперсії в міру накопичення незалежних спостережень.

Основний результат

Нехай D – дисперсія випадкової величини X, а M – її математичне сподівання. Припустимо, що в нашому розпорядженні є скінченна вибірка незалежних спостережень X, тобто послідовність незалежних однаково розподілених величин $x_1, ..., x_n$.

Перевіримо, чи є оцінка (4) незсуненою, а саме чи збігається математичне сподівання оцінки (4) з істинною дисперсією D. Для цього запишемо (4) у вигляді

$$\tilde{D}_n = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \frac{1}{i}\sum_{k=1}^{i} x_k\right)^2 =$$

$$\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i^2 - 2x_i\frac{1}{i}\sum_{k=1}^{i} x_k + \left(\frac{1}{i}\sum_{k=1}^{i} x_k\right)^2\right)$$

та знайдемо математичне сподівання оцінки $\tilde{D}_n$:

$$M(\tilde{D}_n) =$$

$$\frac{1}{n-1}\sum_{i=1}^{n}\left(M[x_i^2] - 2M\left[x_i\frac{1}{i}\sum_{k=1}^{i} x_k\right] + M\left[\frac{1}{i}\sum_{k=1}^{i} x_k\right]^2\right).$$

Розглянемо окремо кожний доданок в останньому виразі. Оскільки оцінка дисперсії $\tilde{D}_n$ не залежить від вибору початку координат, за аналогією з [8, с. 316] виберемо його в точці m. Тоді $M[x_i^2] = D$. Далі, якщо врахувати, що величини $x_1, ..., x_n$ незалежні, а значить, $M[x_i x_j] = 0 \ \forall j \neq i$, то другий доданок можна записати так:

$$M\left[\left(x_i\frac{1}{i}\sum_{k=1}^{i} x_k\right)^2\right] =$$

$$= M\left[x_i\frac{x_1 + x_2 + ... + x_i}{i}\right] = M\left[\frac{x_i^2}{i}\right] = \frac{1}{i}D .$$

І, нарешті, третій доданок набуває вигляду

$$M\left[\left(\frac{1}{i}\sum_{k=1}^{i} x_k\right)^2\right] = M\left[\frac{1}{i^2}(x_1 + x_2 + ... + x_i)^2\right] =$$

$$= M\left[\frac{1}{i^2}(x_1^2 + x_2^2 + ... + x_i^2)\right] = \frac{1}{i^2}iD = \frac{1}{i}D .$$

Отже, маємо

$$M(\tilde{D}_n) = \frac{1}{n-1}\sum_{i=1}^{n}\left(D - 2\frac{1}{i}D + \frac{1}{i}D\right) = \frac{1}{n-1}\sum_{i=1}^{n} D\left(1 - \frac{1}{i}\right) =$$

$$= \frac{1}{n-1}\left(nD - D\left(1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n}\right)\right) =$$

$$= D\frac{\left(n - \left(1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n}\right)\right)}{n-1}.$$

Звідси видно, що математичне сподівання оцінки $\tilde{D}_n$ не збігається з істинним значенням дисперсії D, а значить, оцінка (4) зсунена.

Для отримання незсуненої оцінки достатньо помножити оцінку $\tilde{D}_n$ на величину

$$\frac{n-1}{n - \left(1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n}\right)} = \frac{n-1}{n - \Gamma(n)},$$

тобто

$$\breve{D}_n = \frac{n-1}{n - \Gamma(n)}\tilde{D}_n \tag{6}$$

або

$$\breve{D}_n = \frac{1}{n - \Gamma(n)}\sum_{i=1}^{n}\left(x_i - \frac{\sum_{k=1}^{i} x_k}{i}\right)^2, \tag{7}$$

де $\Gamma(n)$ – часткова сума гармонійного ряду:

$$\Gamma(n) = 1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n} =$$

$$= \ln n + \gamma_0 + o(1) \approx \ln n + \gamma_0 , \tag{8}$$

а $\gamma_0 = 0{,}57721564490...$ – стала Ейлера–Маскероні [10].

Перевіримо тепер, чи є оцінка (7) спроможною, тобто чи збігається за ймовірністю послідовність оцінок $\breve{D}_n$ до величини D, що оцінюється.

Традиційно перевірка спроможності оцінки проводиться з використанням закону великих чисел у будь-якій формі. Однак очевидно, що величини під знаком суми виразу (7), тобто $(x_i - \tilde{m}_i)^2$ та $(x_{i+1} - \tilde{m}_{i+1})^2$, не є незалежними, а всі класичні теореми про закон великих чисел розглядають послідовність незалежних величин, що стоять під знаком суми.

Отже, ми не в змозі застосувати закон великих чисел до нашої послідовності оцінок, тому вимушені вибрати один з обхідних шляхів для знаходження границі за ймовірністю нашої послідовності оцінок.

Один із можливих шляхів – це встановити зв'язок між виразами (1), (7) та використати відомий факт спроможності стандартної оцінки (1).

Спочатку встановимо взаємозв'язок між оцінками (1) і (4).

Безпосередньою перевіркою неважко переконатися в тому, що

$$\tilde{D}_2 = \frac{1}{2}\bar{D}_2,$$

$$\tilde{D}_3 = \frac{2}{3}\bar{D}_3 - \frac{1}{12}\bar{D}_2, \quad (9)$$

$$\tilde{D}_4 = \frac{3}{4}\bar{D}_4 - \frac{1}{18}\bar{D}_3 - \frac{1}{18}\bar{D}_2.$$

Вирази (9) дозволяють припустити, що для $n \geq 2$ справедливо співвідношення

$$\tilde{D}_n = \frac{n-1}{n}\bar{D}_n - \frac{1}{n-1}\left(\frac{n-2}{n(n-1)}\bar{D}_{n-1} + \frac{n-3}{(n-1)(n-2)}\bar{D}_{n-2} + \ldots + \frac{n-i-1}{(n-i+1)(n-i)}\bar{D}_{n-i} + \ldots + \frac{1}{2\cdot 3}\bar{D}_2\right). \quad (10)$$

Доведемо справедливість (10) методом математичної індукції. Правомірність бази індукції доводять за допомогою співвідношень (9). Тому залишається показати, що якщо рівність (10) справедлива для деякого довільного числа *n*, то вона справедлива і для значення $n+1$.

Для цього покажемо, що з (10) випливає така формула:

$$\tilde{D}_{n+1} = \frac{n}{n+1}\bar{D}_{n+1} - \frac{1}{n}\left(\frac{n-1}{(n+1)n}\bar{D}_n + \frac{n-2}{n(n-1)}\bar{D}_{n-1} + \ldots + \frac{n-i-1}{(n-i+1)(n-i)}\bar{D}_{n-i} + \ldots + \frac{1}{2\cdot 3}\bar{D}_2\right). \quad (11)$$

Оскільки згідно з (4)

$$\tilde{D}_n = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \frac{x_1 + \ldots + x_i}{i}\right)^2,$$

то

$$\tilde{D}_{n+1} = \frac{1}{n}\sum_{i=1}^{n+1}\left(x_i - \frac{x_1 + \ldots + x_i}{i}\right)^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{x_1 + \ldots + x_i}{i}\right)^2 + \frac{1}{n}\left(x_{n+1} - \frac{x_1 + \ldots + x_{n+1}}{n+1}\right)^2 =$$

$$= \frac{n-1}{n}\tilde{D}_n + \frac{1}{n}\left(x_{n+1} - \frac{x_1 + \ldots + x_{n+1}}{n+1}\right)^2.$$

Скористуємось (10) і продовжимо доведення:

$$\tilde{D}_{n+1} = \frac{(n-1)^2}{n^2}\bar{D}_n - \frac{1}{n}\left(\frac{n-2}{n(n-1)}\bar{D}_{n-1} + \frac{n-3}{(n-1)(n-2)}\bar{D}_{n-2} + \ldots + \frac{n-i-1}{(n-i+1)(n-i)}\bar{D}_{n-i} + \ldots + \frac{1}{2\cdot 3}\bar{D}_2\right) + \frac{1}{n}\left(x_{n+1} - \frac{x_1 + \ldots + x_{n+1}}{n+1}\right)^2. \quad (12)$$

Порівнюючи (11) та (12), бачимо, що (11) еквівалентно (12) тоді, коли

$$\frac{n}{n+1}\bar{D}_{n+1} - \frac{n-1}{n^2(n+1)}\bar{D}_n = \frac{(n-1)^2}{n^2}\bar{D}_n + \frac{1}{n}\left(x_{n+1} - \frac{x_1 + \ldots + x_{n+1}}{n+1}\right)^2$$

або

$$n\bar{D}_{n+1} = (n-1)\bar{D}_n + \frac{n+1}{n}\left(x_{n+1} - \frac{x_1 + \ldots + x_{n+1}}{n+1}\right)^2. \quad (13)$$

Перетворимо ліву частину виразу (13) таким чином:

$$W_{\text{л}} = n\frac{1}{n}\sum_{i=1}^{n+1}\left(x_i - \frac{x_1 + \ldots + x_{n+1}}{n+1}\right)^2 =$$

$$= \sum_{i=1}^{n+1}x_i^2 - \frac{1}{n+1}\left(\sum_{i=1}^{n+1}x_i\right)^2 =$$

$$= \sum_{i=1}^{n}x_i^2 + x_{n+1}^2 - \frac{1}{n+1}\left(\left(\sum_{i=1}^{n}x_i\right)^2 + 2x_{n+1}\sum_{i=1}^{n}x_i + x_{n+1}^2\right).$$

Отже, маємо

$$W_{\text{л}} = \sum_{i=1}^{n}x_i^2 + \frac{n}{n+1}x_{n+1}^2 - \frac{1}{n+1}\left(\left(\sum_{i=1}^{n}x_i\right)^2 + 2x_{n+1}\sum_{i=1}^{n}x_i\right). \quad (14)$$

Розглянемо тепер праву частину виразу (13):

$$W_{\text{пр}} = \sum_{i=1}^{n}x_i - \frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2 +$$

$$+\frac{1}{n(n+1)}\left((n+1)^2 x_{n+1}^2+\left(\sum_{i=1}^{n+1} x_i\right)^2-\right.$$

$$\left.-2(n+1)x_{n+1}\sum_{i=1}^{n+1} x_i\right)=$$

$$=\sum_{i=1}^{n} x_i-\frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2+\frac{1}{n(n+1)}\left((n+1)^2 x_{n+1}^2+\left(\sum_{i=1}^{n} x_i\right)^2+\right.$$

$$+x_{n+1}^2+2x_{n+1}\sum_{i=1}^{n} x_i-$$

$$\left.-2(n+1)x_{n+1}\sum_{i=1}^{n} x_i-2(n+1)x_{n+1}^2\right).$$

В результаті нескладних перетворень отримаємо

$$W_{\text{пр}}=\sum_{i=1}^{n} x_i^2+\frac{n}{n+1}x_{n+1}^2-$$

$$-\frac{1}{n+1}\left(\left(\sum_{i=1}^{n} x_i\right)^2+2x_{n+1}\sum_{i=1}^{n} x_i\right). \qquad (15)$$

Порівнюючи вирази (14) і (15), переконуємось в ідентичності лівої та правої частин виразів (13). Таким чином, співвідношення (13) доведено.

Отже, праві частини в (11) і (12) ідентичні, та з істинності (12) випливає істинність (11). Звідси формула (10) справедлива для будь-якого значення n.

Із співвідношення (6) з врахуванням (10) випливає співвідношення, що встановлює взаємозв'язок оцінок (7) і (4):

$$\breve{D}_n=\frac{(n-1)^2}{n(n-\Gamma(n))}\bar{D}_n-\frac{1}{n-\Gamma(n)}\left(\frac{n-2}{n(n-1)}\bar{D}_{n-1}+\right.$$

$$+\frac{n-3}{(n-1)(n-2)}\bar{D}_{n-2}+\ldots+$$

$$\left.+\frac{n-i-1}{(n-i+1)(n-i)}\bar{D}_{n-i}+\ldots+\frac{1}{2\cdot 3}\bar{D}_2\right). \qquad (16)$$

Оскільки оцінка $\bar{D}_n$ спроможна, то для доведення спроможності $\breve{D}_n$ достатньо показати, що

$$\lim_{n\to\infty}\breve{D}_n=\lim_{n\to\infty}\bar{D}_n. \qquad (17)$$

Для доведення використаємо теорему Теплиця [11] в такому формулюванні. Нехай є c_{nk}, $1\le k\le n$, такі, що:

$$c_{nk}\underset{n\to\infty}{\to}0 \ \ \forall k=1,2,\ldots,n, \qquad (18)$$

$$\sum_{k=1}^{n} c_{nk}\underset{n\to\infty}{\to}1, \qquad (19)$$

$$\exists C>0:\ \sum_{k=1}^{n}\left|c_{nk}\right|\le C \ \ \forall n\ge 1. \qquad (20)$$

Тоді, якщо деяка послідовність $\{a_n, n\ge 1\}$ збіжна, то і послідовність $\{b_n=\sum_{k=1}^{n} c_{nk}a_k\ ,\ n\ge 1\}$ також збіжна, причому $\lim_{n\to\infty} a_n=\lim_{n\to\infty} b_n$.

Застосуємо теорему Теплиця до таких послідовностей:

$$a_n=\bar{D}_n, \qquad b_n=\tilde{D}_n=\sum_{k=1}^{n} c_{nk}a_n. \qquad (21)$$

З (21) маємо

$$b_n=\sum_{k=1}^{n} c_{nk}\bar{D}_k=c_{n1}\bar{D}_1+c_{n2}\bar{D}_2+\ldots+c_{nn}\bar{D}_n,$$

де

$$c_{nk}=\begin{cases}0, \ k=1,\\ -\dfrac{1}{n-\Gamma(n)}\dfrac{k-1}{(k+1)k}, \ 1<k<n,\\ \dfrac{(k-1)^2}{k(n-\Gamma(n))}, \ k=n.\end{cases} \qquad (22)$$

Перевіримо виконання умов (18)–(20) для послідовностей (21). Очевидно, що $c_{nk}\underset{n\to\infty}{\to}0$ для будь-якого фіксованого k, тобто умова (18) виконана.

Перевіримо виконання умови (19):

$$\sum_{k=1}^{n} c_{nk}=\frac{(n-1)^2}{n(n-\Gamma(n))}-\frac{1}{n-\Gamma(n)}\left(\frac{n-2}{n(n-1)}+\right.$$

$$\left.+\frac{n-3}{(n-1)(n-2)}+\ldots+\frac{n-i-1}{(n-i+1)(n-i)}+\ldots+\frac{1}{2\cdot 3}\right)=$$

$$= \frac{(n-1)^2}{n(n-\Gamma(n))} - \frac{1}{n-\Gamma(n)}\left(\frac{n-1}{n(n-1)} - \frac{1}{n(n-1)} + \right.$$

$$+ \frac{n-2}{(n-1)(n-2)} - \frac{1}{(n-1)(n-2)} + \ldots +$$

$$+ \frac{n-i}{(n-i+1)(n-i)} - \frac{1}{(n-i+1)(n-i)} + \ldots +$$

$$\left. + \frac{2}{2\cdot 3} - \frac{1}{2\cdot 3}\right) =$$

$$= \frac{(n-1)^2}{n(n-\Gamma(n))} - \frac{1}{n-\Gamma(n)}\left(\frac{1}{n} + \frac{1}{n-1} + \ldots + \frac{1}{3}\right) +$$

$$+ \frac{1}{n-1}\left(\frac{1}{n(n-1)} + \frac{1}{(n-1)(n-2)} + \ldots + \frac{1}{2\cdot 3}\right).$$

Таким чином, матимемо

$$\sum_{i=1}^{n} c_{nk} = \frac{(n-1)^2}{n(n-\Gamma(n))} - A_n + B_n, \qquad (23)$$

де

$$A_n = \frac{1}{n-\Gamma(n)}\left(\frac{1}{n} + \frac{1}{n-1} + \ldots + \right.$$

$$\left. + \frac{1}{n-i} + \ldots + \frac{1}{3}\right),\ i = 0, \ldots, n-3,$$

$$B_n = \frac{1}{n-\Gamma(n)}\left(\frac{1}{n(n-1)} + \frac{1}{(n-1)(n-2)} + \ldots + \right.$$

$$+ \frac{1}{(n-i)(n-i-1)} + \ldots +$$

$$\left. + \frac{1}{2\cdot 3}\right),\ i = 0, \ldots, n-3.$$

Оскільки для гармонійного ряду $\Gamma(n)$ справедливе співвідношення (7), то матимемо

$$\frac{1}{3} + \ldots + \frac{1}{n-1} + \frac{1}{n} = \gamma_0 + \ln n - 1{,}5 + o(1) \sim \ln n.$$

Таким чином, отримуємо

$$A_n \sim \frac{\ln n}{n-\Gamma(n)} \underset{n\to\infty}{\to} 0.$$

Далі, із врахуванням того, що $i < n-2$, для загальних членів послідовностей A_n і B_n справедливо співвідношення

$$0 \le \frac{1}{(n-\Gamma(n))(n-i)(n-i-1)} \le$$

$$\le \frac{n-i-1}{(n-\Gamma(n))(n-i)(n-i-1)} = \frac{1}{(n-\Gamma(n))(n-i)},$$

та за ознакою порівняння $B_n \underset{n\to\infty}{\to} 0$.

Неважко показати, що

$$\lim_{n\to\infty} \frac{(n-1)^2}{n(n-\Gamma(n))} = 1.$$

Таким чином, маємо

$$\lim_{n\to\infty}\sum_{k=1}^{n} c_{nk} = \lim_{n\to\infty}\frac{(n-1)^2}{n(n-\Gamma(n))} - \lim_{n\to\infty} A_n + \lim_{n\to\infty} B_n = 1,$$

тобто і умова (19) теореми Теплиця виконана.

Перевіримо тепер виконання умови (20):

$$\sum_{k=1}^{n} |c_{nk}| = \sum_{k=1}^{n} c_{nk} = \frac{(n-1)^2}{n(n-\Gamma(n))} - \frac{1}{n-\Gamma(n)} \times$$

$$\times \left(\frac{n-2}{n(n-1)} + \frac{n-3}{(n-1)(n-2)} + \ldots + \frac{1}{2\cdot 3}\right) <$$

$$< \frac{(n-1)^2}{n(n-\Gamma(n))} < \frac{(n-1)^2}{n(n-(n+1)/2)} = 2\frac{(n-1)^2}{n(n-1)} < 2.$$

Інакше кажучи, умова (20) виконується для $C = 2$.

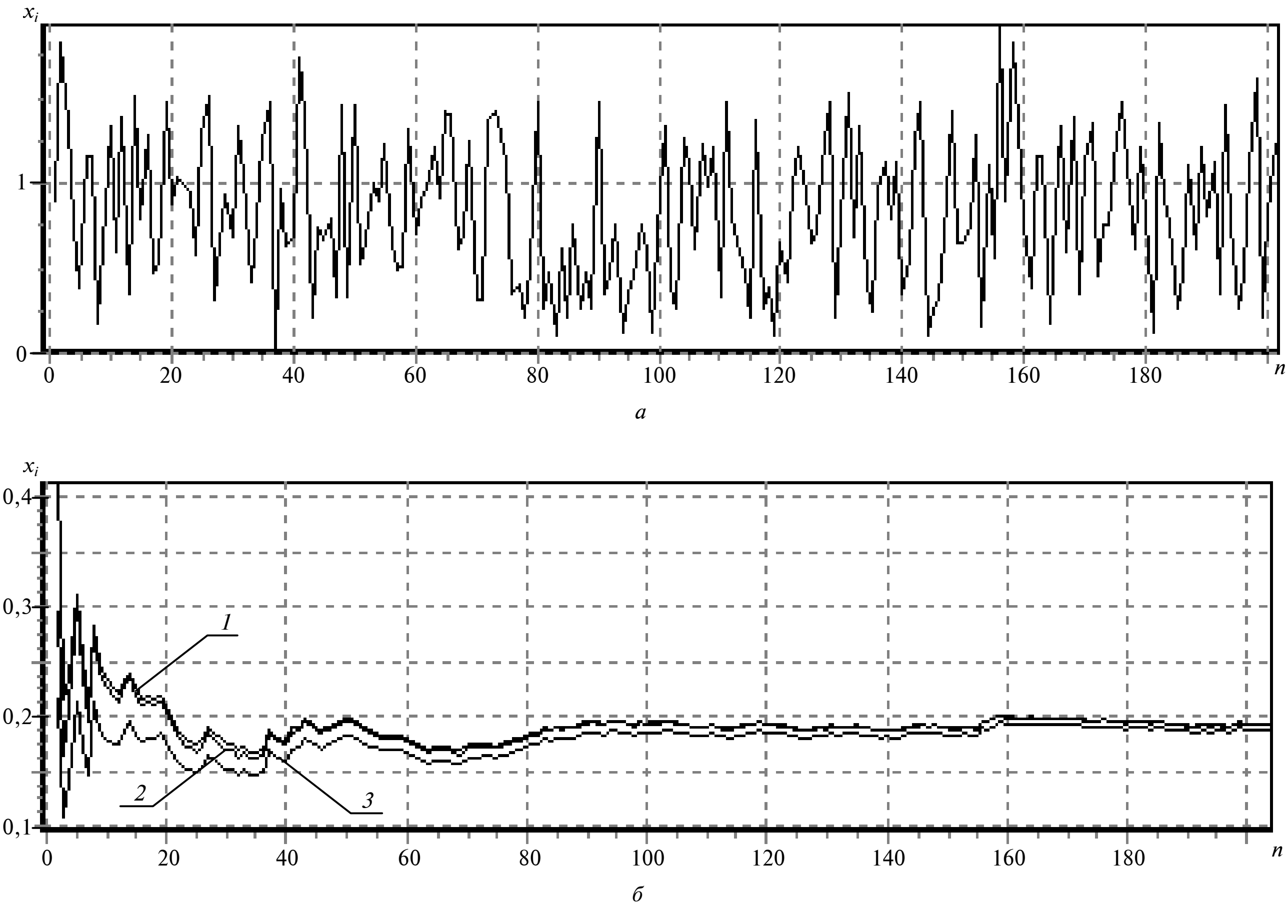
Оскільки всі умови теореми Теплиця виконуються, то $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$, де a_n, b_n визначені згідно з (21). Звідси випливає, що

$$\lim_{n\to\infty} \breve{D}_n = \lim_{n\to\infty} \bar{D}_n,$$

а отже, оцінка $\breve{D}_N$ спроможна. Більше того, оскільки відомо [1], що стандартна оцінка $\bar{D}_n$ навіть сильно спроможна, тобто збігається до величини, що оцінюється, не тільки за ймовірністю, а й з ймовірністю одиниця, то також сильно спроможними будуть оцінки $\breve{D}_n$ і $\breve{D}_N$.

На рисунку наведено результати чисельного моделювання. Вибірка незалежних спостережень $x_1, \ldots, x_n$ (рисунок, a), подана $n = 200$ реалізаціями випадкової величини X, що мала нормальний розподіл з математичним сподіванням $M = 1$ та дисперсію $D = 0{,}19$.

Як видно з результатів обчислення (рисунок, $б$), починаючи з кроку $n > 30$ модифікована оцінка $\breve{D}_n$, яка обчислювалась за формулою (7), майже збігається з традиційною оцінкою $\bar{D}_n$.



Результати чисельного моделювання: *а* — вибірка спостережень x_i; *б* — результати обчислення: *1* — оцінка $\bar{D}_n$; *2* — оцінка $\breve{D}_n$; *3* — оцінка $\tilde{D}_n$

Висновки

Традиційна оцінка дисперсії випадкової величини (1), що обчислюється за скінченною послідовністю незалежних спостережень, як це відомо, є незсуненою та спроможною. Запропонована модифікована оцінка (7), яка є більш зручною для обчислення поточного значення дисперсії в міру накопичення даних. Доведено, що модифікована оцінка також є незсуненою та спроможною.

Запропонована оцінка може бути використана при розв'язуванні практичних задач у техніці, медицині, економіці та інших галузях, коли треба прискорити оцінювання поточного значення дисперсії випадкової величині в міру накопичення незалежних даних.

О.А. Жуковская, Г.А. Глушаускене,
Л.С.Файнзильберг

ИССЛЕДОВАНИЕ СВОЙСТВ МОДИФИЦИРОВАННОЙ ОЦЕНКИ ДИСПЕРСИИ СЛУЧАЙНОЙ ВЕЛИЧИНЫ ПО ВЫБОРКЕ НЕЗАВИСИМЫХ НАБЛЮДЕНИЙ

Предложена модифицированная оценка, которая позволяет вычислять текущее значение дисперсии случайной величины по мере накопления не-

O.A. Zhukovska, G.A. Glushauskene,
L.S. Fainzilberg

RESEARCH OF THE MODIFIED ESTIMATION PROPERTIES OF RANDOM VARIABLE'S VARIANCE ON SAMPLE OF INDEPENDENT OBSERVATIONS

This study describes the modified estimator that allows determining the current value of a random variable's dispersion in the process of independent observation accumulation. Through experiments perзависимых наблюдений. Доказано, что такая оценка, так же как и традиционная, является несмещенной и состоятельной. Представлены результаты численного моделирования.

formed, we prove that such estimator as well as traditional is unbiased and consistent. Moreover, we provide the results of numerical modelling.

1. *Крамер Г.* Математические методы статистики. – М.: Мир, 1975. – 644 с.
2. *Reza A. Soltani*, *Moeanaddin R.* On Dispersion of Stable Random Vectors and Its Application in the Prediction of Multivariate Stable Processes // J. of Applied Probability. – 1994. – **31**, N 3. – P. 691–699.
3. *Brick J.M.*, *Kalton, G.* Handling Missing Data in Survey Research // Statistical Methods in Medical Research. – 1996. – N 5. – P. 215–238.
4. *Brick J.M.*, *Morganstein D.* WesVarPC: Software for Computing Variance Estimates from Complex Designs // Proc. of the 1996 Annual Research Conference. – Washington, U.S. Bureau of the Census, 1996. – P. 861–866.
5. *Kagan A.*, *Shepp L.* A. Why the variance? // Statistics and Probability Letters. – 1998. – **38**, N 4. – P. 329–333.
6. *Wolter K.M.* Introduction to variance estimation. – New York: Springer Verlag, 2007. – 450 p.
7. *Горбань І.І.* Теорія ймовірностей і математична статистика для наукових працівників та інженерів. – К.: Ін-т проблем мат. машин і систем НАН України, 2003. – 244 с.
8. *Вентцель Е.С.* Теория вероятностей. – М.: Наука, 1969. – 576 с.
9. *Knuth D.E.* The Art of Computer Programming. Vol. 2. Seminumerical Algorithms, 3rd edn. – Boston: Addison-Wesley, 1998. – 565 p.
10. *Корн Г.*, *Корн Т.* Справочник по математике. – М.: Наука, 1978. – 832 с.
11. *Дороговцев А.Я.* Математический анализ: Справ. пос. – К.: Вища шк., 1985. – 528 с.

Рекомендована Радою факультету менеджменту та маркетингу НТУУ "КПІ"

Надійшла до редакції 4 лютого 2008 року