УДК 621.391

О.В. Івченко

Оцінювання дисперсії корельованої ексцесної завади методом максимізації полінома

Приведен практический пример использования нового алгоритма статистического оценивания параметров случайных негауссовских процессов при их моментно-кумулянтном описании. Найдена оценка дисперсии эксцесной коррелированной помехи. Проанализированы асимптотические свойства полученной оценки, показана ее эффективность с ростом степени стохастического полинома.

Examined of the use of new algorithm of statistical evaluation of parameters of the casual no Gaussian processes assumed that variate description of cumulant and moment . The assessment of a dispersion of statistically dependent no Gaussian exces that processes the correlated is found. Are analysed properties of assessments.

Ключевые слова: *оценка параметров, выборка, моментно-кумулянтное описание, негауссовская случайная величина, корреляция, метод максимизации полинома, статистическая зависимость, коефициент уменьшения дисперсии.*

Вступ

Актуальність побудови методів і алгоритмів обробки статистичних даних на основі спостережень обумовлена прагненням постійного підвищення точності вимірів (оцінювання) параметрів. Підвищення точності пов’язане з врахуванням у алгоритмах обробки всіх властивостей процесу, що спостерігається, зокрема наявності статистичних зв’язків між вибірковими значеннями [1]. Більшість класичних методів оцінювання оперують статистичними даними за умови відсутності статистичних зв’язків між ними. Так, наприклад, при оцінюванні дисперсії методом моментів [2] не враховуються значення багатомірних моментних функцій [3], що характеризують статистичні зв’язки. Оцінювання методом максимальної правдоподібності [2] вимагає знання багатомірних функцій розподілу, знаходження яких для реальних процесів є досить складним. Метод максимізації полінома [4, 5], на відміну від метода моментів, в своєму алгоритмі оперує всіма моментними і кумулянтними функціями, що дозволяє підвищити точність оцінки параметрів випадкових негауссівських процесів. При цьому, якщо вибіркові значення є статистично пов’язаними, то доцільно використовувати адаптований метод максимізації полінома [6].

Метою роботи є розширення практичного застосування адаптованого методу максимізації полінома для виміру параметрів негауссівських процесів, що дозволяє отримати кращі імовірнісні характеристики оцінок в порівнянні з результатами класичних підходів.

1. Постановка задачі

Як досліджувана модель сигналу використовується близький до гауссівського стаціонарний корельований випадковий процес $\xi(t,\vartheta)$ з нульовим математичним сподіванням $m_1(\vartheta)=0$, який описується послідовністю моментних функцій $m_{ij}^{(v,k)}(t_v,t_k;\vartheta)$ порядку $i+j$ ($i,j=\overline{0,s}$, $v,k=\overline{0,n}$) двомоментного розподілу. Вважається, що моментні функції $m_{ij}^{(v,k)}(t_v,t_k;\vartheta)$ є обмеженими, неперервними по $\tau=|t_v-t_k|$ функціями і в загальному випадку залежать від одного і того ж параметра ϑ, що приймає довільні значення на відкритому інтервалі. Як параметр ϑ можуть виступати статистичні характеристики випадкового процесу, наприклад дисперсія, коефіцієнт асиметрії, коефіцієнт ексцесу і т. д.

Задача оцінювання може бути сформульована таким чином. Нехай із випадкового сигнала $\xi(t,\vartheta)$ в моменти часу $t_1,t_2,...,t_n$ беруться значення $x_1=\xi(t_1,\vartheta)$, $x_2=\xi(t_2,\vartheta)$, ..., $x_n=\xi(t_n,\vartheta)$, причому дані значення однаково розподілені і статистично пов’язані між собою. Необхідно за однією реалізацією випадкового процесу $\xi(t,\vartheta)$, $t\in[-T;T]$, яка мала місце при значенні параметра $\vartheta=\vartheta_0$, побудувати таку функцію від вибіркових значень $\overline{\vartheta}=f(x_1,x_2,...,x_n)$, яка б була близькою до істинного значення параметра ϑ.

В даній роботі як випадковий сигнал $\xi(t,\vartheta)$ обраний ексцесний корельовний випадковий процес з нульовим математичним сподіванням. Як зазначалось в роботі [7], його статистичний опис представлений у вигляді одномірних кумулянтних функцій $\chi_2(\vartheta)$ і $\chi_4(\vartheta)$, а $\chi_3=0$ і двомірних кумулянтних функцій:

$\chi_{11}^{(v,k)}(t_v,t_k;\vartheta)$, $\chi_{13}^{(v,k)}(t_v,t_k;\vartheta)$ і $\chi_{22}^{(v,k)}(t_v,t_k;\vartheta)$, а $\chi_{12}^{(v,k)}(t_v,t_k)=0$, які залежать від оцінюваного параметра ϑ. Всі інші кумулянтні функції даного процесу можуть приймати довільні значення. При цьому статистичний зв'язок другого порядку, що спостерігається між вибірковими значеннями, характеризується кумулянтною функцією другого порядку (функція кореляції) $\chi_{11}^{(v,k)}(t_v,t_k;\vartheta)=\chi_2(\vartheta)R_{vk}(\tau)$, яка носить експоненційний характер, тобто нормована кумулянтна функція (коефіцієнт кореляції) дорівнює $R_{vk}(\tau)=e^{-A|\tau|}$, де $A>0$ – коефіцієнт.

Інтервал кореляції визначається як:

$$\tau_0=\int_0^\infty R_{vk}(\tau)d\tau=\int_0^\infty e^{-A|\tau|}d\tau=\frac{1}{A}. \qquad (1)$$

Необхідно за вибіркою $\bar{x}=(x_1,...,x_n)$ знайти оцінку дисперсії $\vartheta=\bar{\chi}_2$ за умови апріорної визначеності інших параметрів сигналу: $\chi_4(\vartheta)$, $\chi_{11}^{(v,k)}(t_v,t_k;\vartheta)$, $\chi_{13}^{(v,k)}(t_v,t_k;\vartheta)$ і $\chi_{22}^{(v,k)}(t_v,t_k;\vartheta)$.

В роботі використаний адаптований метод максимізації полінома з застосуванням стохастичних поліномів до четвертого порядку [6].

2. Кореляційний опис ексцесних корельованих випадкових процесів

Для знаходження асимптотичної дисперсії оцінки дисперсії, відповідно до методики, запропонованої в роботі [6], необхідно знайти матрицю коефіцієнтів кореляції, яка визначає значення кореляційних зв’язків між всіма елементами досліджуваної вибірки:

$$\mathbf{Z}=\begin{pmatrix} 1 & R_{12} & ... & R_{1n} \\ R_{21} & 1 & ... & R_{2n} \\ ... & & & \\ R_{n1} & R_{n2} & ... & 1 \end{pmatrix}, \qquad (2)$$

де $|\mathbf{Z}|$ — визначник матриці коефіцієнтів кореляції.

Згідно з адаптованим методом максимізації полінома, в його алгоритмі використовують функції $K_{ij}^{(v,k)}(t_v,t_k;\vartheta)$, які для стаціонарних процесів двомоментного розподілу визначаються наступним чином:

$$K_{ij}^{(v,k)}(t_v,t_k;\vartheta)=M(\xi^i(t_v)-m_i(\vartheta))(\xi^j(t_k)-$$
$$-m_j(\vartheta))=m_{ij}^{(v,k)}(t_v,t_k;\vartheta)-\alpha_i(\vartheta)\alpha_j(\vartheta), \qquad (3)$$

де $\alpha_i(\vartheta)\alpha_j(\vartheta)$ – добуток одномірних моментних функцій сигналу $\xi(t,\vartheta)$.

Тоді для ексцесних корельованих випадкових процесів, згідно з формулою (3) з урахуванням формул зв’язку моментних і кумулянтних функцій для одномоментного і двомоментного розподілу [3] і значень цих функцій для даного процесу [7], отримуємо наступні значення функцій $K_{ij}^{(v,k)}(t_v,t_k;\vartheta)$:

$$K_{11}^{(v,k)}(t_v,t_k;\vartheta)=m_{11}^{(v,k)}(t_v,t_k,\vartheta)-$$
$$-\alpha_1(\vartheta)\alpha_1(\vartheta)=\chi_{11}^{(v,k)}(t_v,t_k;\vartheta),$$
$$K_{12}^{(v,k)}(t_v,t_k;\vartheta)=0,$$
$$K_{13}^{(v,k)}(t_v,t_k;\vartheta)=\chi_{13}^{(v,k)}(t_v,t_k;\vartheta)+$$
$$+3\chi_2(\vartheta)\chi_{11}^{(v,k)}(t_v,t_k;\vartheta),$$
$$K_{14}^{(v,k)}(t_v,t_k;\vartheta)=0, \qquad (4)$$
$$K_{22}^{(v,k)}(t_v,t_k;\vartheta)=\chi_{22}^{(v,k)}(t_v,t_k;\vartheta)+$$
$$+2\chi_{11}^{2(v,k)}(t_v,t_k;\vartheta),$$
$$K_{23}^{(v,k)}(t_v,t_k;\vartheta)=0,$$
$$K_{33}^{(v,k)}(t_v,t_k;\vartheta)=6\chi_{31}^{(v,k)}(t_v,t_k;\vartheta)\chi_2(\vartheta)+$$
$$+9\chi_{11}^{(v,k)}(t_v,t_k;\vartheta)\chi_{22}^{(v,k)}(t_v,t_k;\vartheta)+$$
$$+9\chi_2^2(\vartheta)\chi_{11}^{(v,k)}(t_v,t_k;\vartheta)+6\chi_{11}^{3(v,k)}(t_v,t_k;\vartheta).$$

Слід зазначити, що при відсутності кореляції між вибірковими значеннями кумулянтні функції двомоментного розподілу перетворюються в кумулянти відповідного порядку одномоментного розподілу. Тоді функції (4) перетворюються у добре відомі функції для некорельованих ексцесних випадкових величин [4]:

$$K_{11}=\chi_2,\ K_{12}=0,\ K_{22}=\chi_4+2\chi_2^2,$$
$$K_{13}=\chi_4+3\chi_2^2,\ K_{23}=0,\ K_{33}=15\chi_2\chi_4+15\chi_2^3.$$

3. Оцінювання дисперсії ексцесних корельованих випадкових сигналів

Згідно з алгоритмом адаптованого методу максимізації полінома [6], досліджувані статистичні дані $x_1=\xi(t_1,\vartheta)$, $x_2=\xi(t_2,\vartheta)$, ..., $x_n=\xi(t_n,\vartheta)$ представляються у вигляді стохастичного поліному степеня s:

$$l_{sn}(\bar{x}/\vartheta;Z)=k_0(\vartheta;Z)+\sum_{i=1}^{s}k_i(\vartheta;Z)\sum_{v=1}^{n}x_v^i.$$

При цьому коефіцієнти полінома $k_0(\vartheta;Z)$, $k_i(\vartheta;Z)$ знаходяться по критерію мінімума дисперсії шуканої оцінки [5]. Слід зазначити, що оцінка дисперсії випадкового сигналу може бути знайдена з використанням стохастичних поліномів степеня $s=2$ і вище.

Тоді оцінка χ_2 при моментному описі випадкового процесу і степені полінома S знаходиться з рішення рівняння:

$$\sum_{i=1}^{s}\sum_{v=1}^{n} h_{iv}(\vartheta,R_{vk})\left[x_v^i-\alpha_i(\vartheta)\right]\Big|_{\vartheta=\chi_2}=0, \quad (5)$$

де x_v – статистично пов'язані і однаково розподілені вибіркові значення з досліджуваного випадкового процесу $\xi(t,\vartheta)$; $\alpha_i(\vartheta)$ — початкові моменти і-го порядку одномоментного розподілу ексцесного корельованого випадкового процесу [4]; $h_{iv}(\vartheta,R_{vk})$ — невідомі коефіцієнти, що крім оцінюваного параметра залежать від коефіцієнтів кореляції та знаходяться з рішення системи рівнянь:

$$\sum_{v=1}^{n}\sum_{j=1}^{s} h_{jv}(\vartheta,R_{vk})K_{ij}^{(v,k)}(t_v,t_k;\vartheta)=\frac{d}{d\vartheta}\alpha_i(\vartheta). \quad (6)$$

Для вирішення даної системи скористаємося формулами Крамера [8]:

$$h_{jv}(\vartheta,R_{vk})=\frac{\Delta_{isv}(\vartheta,R_{vk})}{\Delta_s(\vartheta,R_{vk})},\ \ i,j=\overline{1,s},\ \ v=\overline{1,n},$$

де $\Delta_s(\vartheta,R_{vk})=\left|K_{ij}^{(v,k)}(t_v,t_k;\vartheta)\right|$ — визначник матриці розмірністю s, елементами якої є функції $K_{ij}^{(v,k)}(t_v,t_k;\vartheta)$; $\Delta_{isv}(\vartheta,R_{vk})$ — визначник, що отримується з $\Delta_s(\vartheta,R_{vk})$ заміною i-го стовпця стовпцем, що складається з вільних членів системи рівнянь (6).

Тоді для другої степені стохастичного полінома маємо:

$$h_{1v}(\vartheta,R_{vk})=\frac{\Delta_{12v}(\vartheta,R_{vk})}{\Delta_2(\vartheta,R_{vk})};$$
$$h_{2v}(\vartheta,R_{vk})=\frac{\Delta_{22v}(\vartheta,R_{vk})}{\Delta_2(\vartheta,R_{vk})}, \quad (7)$$

де: $\Delta_2(\vartheta,R_{vk})=\begin{vmatrix}\mathbf{Z_{11}} & \mathbf{Z_{12}}\\ \mathbf{Z_{21}} & \mathbf{Z_{22}}\end{vmatrix}$, $\mathbf{Z_{11}}$, $\mathbf{Z_{12}}=\mathbf{Z_{21}}$, $\mathbf{Z_{22}}$ — матриці, що складені з елементів $K_{ij}^{(v,k)}(t_v,t_k;\vartheta)$, $i,j=1,2$, $v,k=\overline{1,n}$, відповідно.

Підставляючи коефіцієнти (7) в рівняння максимізації полінома (5), отримуємо:

$$\sum_{v=1}^{n}\frac{\Delta_{22v}(\vartheta,R_{vk})}{\Delta_2(\vartheta,R_{vk})}\left[x_v^2-\chi_2\right]=0.$$

Звідси отримуємо значення оцінки $\overline{\chi_2}$:

$$\overline{\chi}_2=\frac{\sum_{v=1}^{n}\Delta_{22v}(\vartheta,R_{vk})x_v^2}{\sum_{v=1}^{n}\Delta_{22v}(\vartheta,R_{vk})}. \quad (8)$$

В рівнянні максимізації полінома третього степеня, з якого знаходиться оцінка дисперсії ексцесного корельованого випадкового процесу, коефіцієнти знаходяться з рішення системи трьох рівнянь (6).

Використовуючи метод Крамера, знайдемо значення коефіцієнтів:

$$h_{1v}(\vartheta,R_{vk})=0;$$
$$h_{2v}(\vartheta,R_{vk})=\frac{\Delta_{23v}(\vartheta,R_{vk})}{\Delta_3(\vartheta,R_{vk})};$$
$$h_{3v}(\vartheta,R_{vk})=0,$$

де $\Delta_{23}(\vartheta,R_{vk})$ — визначник, що отримується з $\Delta_3(\vartheta,R_{vk})$ заміною 2-го стовпця стовпцем, що складається з вільних членів системи рівнянь (6).

Використовуючи вираз для початкового моменту $\alpha_2(\vartheta)$ та значення коефіцієнта $h_{2v}(\vartheta;R_{vk})$, легко отримати, що рівняння максимізації полінома третього степеня набуває вигляду:

$$\sum_{v=1}^{n}\frac{\Delta_{23v}(\vartheta,R_{vk})}{\Delta_3(\vartheta,R_{vk})}\left[x_v^2-\chi_2\right]=0.$$

З останнього рівняння знаходимо значення оцінки $\overline{\chi_2}$:

$$\overline{\chi}_2=\frac{\sum_{v=1}^{n}\Delta_{23v}(\vartheta,R_{vk})x_v^2}{\sum_{v=1}^{n}\Delta_{23v}(\vartheta,R_{vk})}. \quad (9)$$

Якщо вибірка є некорельованою, то $\Delta_{22v}(\vartheta,R_{vk})=\Delta_{23v}(\vartheta,R_{vk})=1$ і оцінка дисперсії, згідно з формулами (8) і (9), дорівнює класичній оцінці [8]:

$$\overline{\chi}_2=\frac{\sum_{v=1}^{n}x_v^2}{n}.$$

Розглянемо збільшення степеня стохастичного полінома.

В рівнянні максимізації полінома четвертого степеня, коефіцієнти (6) знаходяться з рішення системи чотирьох рівнянь.

Використовуючи формули Крамера, знаходимо значення коефіцієнтів:

$$\begin{aligned} h_{1v}(\vartheta,R_{vk}) &= \frac{\Delta_{14v}(\vartheta,R_{vk})}{\Delta_4(\vartheta,R_{vk})}; \\ h_{2v}(\vartheta,R_{vk}) &= \frac{\Delta_{24v}(\vartheta,R_{vk})}{\Delta_4(\vartheta,R_{vk})}; \\ h_{3v}(\vartheta,R_{vk}) &= \frac{\Delta_{34v}(\vartheta,R_{vk})}{\Delta_4(\vartheta,R_{vk})}; \\ h_{4v}(\vartheta,R_{vk}) &= \frac{\Delta_{44v}(\vartheta,R_{vk})}{\Delta_4(\vartheta,R_{vk})}, \end{aligned} \tag{10}$$

де $\Delta_{i4v}(\vartheta,R_{vk})$ — визначник, що отримується з $\Delta_4(\vartheta,R_{vk})$ заміною v-го стовпця стовпцем, що складається з вільних членів системи рівнянь, $i=\overline{1,...,4}$.

Розв'язок рівняння максимізації полінома при $s=4$ (5) можливий лише чисельними методами.

4. Асимптотичні властивості оцінки параметра дисперсії

Як відомо [4, 5], асимптотична дисперсія оцінки за методом максимізації полінома обернено пропорційна кількості добутої інформації про досліджуваний параметр, яка згідно з роботою [6] дорівнює:

$$J_{sn}(\vartheta) = \sum_{v=1}^{n}\sum_{i=1}^{n} h_{iv}(\vartheta;R_{vk})\frac{d}{d\vartheta}\alpha_i(\vartheta). \tag{11}$$

Підставивши в формулу (11) значення коефіцієнтів стохастичного полінома другого степеня (7), отримаємо значення асимптотичної дисперсії оцінки $\overline{\chi_2}$:

$$\sigma_{(\overline{\chi}_2)2} = J_{2n}^{-1}(\vartheta) = \frac{\Delta_2(\vartheta,R_{vk})}{\sum_{v=1}^{n}\Delta_{22v}(\vartheta,R_{vk})}. \tag{12}$$

При третій степені полінома асимптотична дисперсія оцінки $\overline{\chi_2}$ буде дорівнювати:

$$\sigma_{(\overline{\chi}_2)3} = J_{3n}^{-1}(\vartheta) = \frac{\Delta_3(\vartheta,R_{vk})}{\sum_{v=1}^{n}\Delta_{23v}(\vartheta,R_{vk})}. \tag{13}$$

Аналіз відношення чисельника і знаменника в виразах (12) і (13) показує їх еквівалентність, тобто при другій і третій степені стохастичного полінома оцінки кумулянта другого порядку мають однакову дисперсію.

При значеннях коефіцієнта кореляції, що прямують до нуля, дисперсії (12) і (13) оцінки параметра χ_2 прямують до дисперсії оцінки при незалежній вибірці [2]:

$$\sigma_{(\overline{\chi}_2)} = \chi_2^2\frac{\gamma_4+2}{n}.$$

Дисперсія оцінки $\overline{\chi_2}$ при четвертій степені стохастичного полінома дорівнює:

$$\sigma_{(\overline{\chi}_2)4} = J_{4n}^{-1}(\vartheta) = [\sum_{v=1}^{n}\frac{\Delta_{24v}(\vartheta,R_{vk})}{\Delta_4(\vartheta,R_{vk})} + \sum_{v=1}^{n}\frac{\Delta_{44v}(\vartheta,R_{vk})}{\Delta_4(\vartheta,R_{vk})}2\chi_2(\gamma_4+3)]^{-1}, \tag{14}$$

де γ_4 — коефіцієнт ексцесу випадкового сигналу $\xi(t,\vartheta)$.

На рис. 1 наведена залежність дисперсії оцінки параметра χ_2 при степенях полінома $s=2{,}3$ і $s=4$ від параметра A кореляційної функції при фіксованому значенні об'єму вибірки $n=500$ та коефіцієнта ексцесу γ_4.

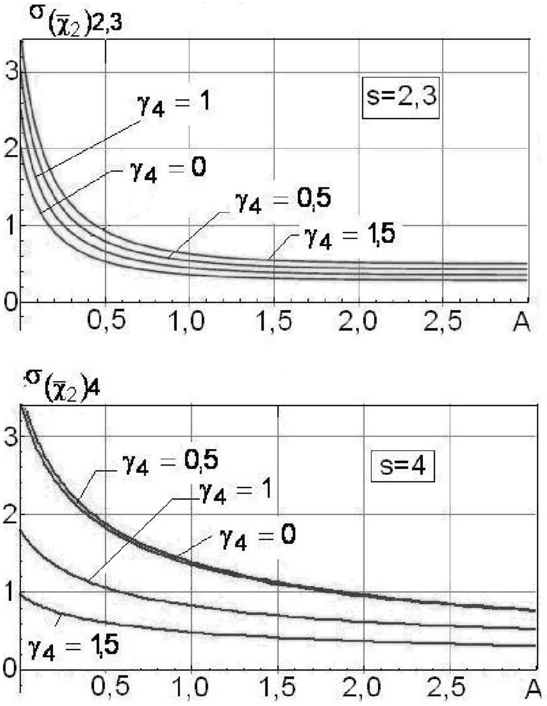

Рис. 1.

З графіка видно, що при збільшенні інтервалу кореляції τ_0 (1) між вибірковим значеннями, що взяті з досліджуваного процесу, дисперсія оцінки збільшується.

Для кількісного визначення зменшення дисперсії оцінки $\overline{\vartheta}$ довільного параметра ϑ, що знайдена методом максимізації полінома при степені полінома s, в порівнянні з дисперсією оцінки при степені полінома k введений коефіцієнт зменшення дисперсії цих оцінок:

$$g_{(\overline{\vartheta})} = \frac{\sigma_{(\overline{\vartheta})s}}{\sigma_{(\overline{\vartheta})k}}.$$

Графік залежності коефіцієнта зменшення дисперсії $g_{(\overline{\chi_2})}$ оцінки $\overline{\chi_2}$ при степенях полінома $s = 4$ і $k = 2$ від значень параметра A кореляційної функції при фіксованих значеннях параметра ексцесу γ_4 при об'ємі вибірки $n = 500$ наведений на рис. 2.

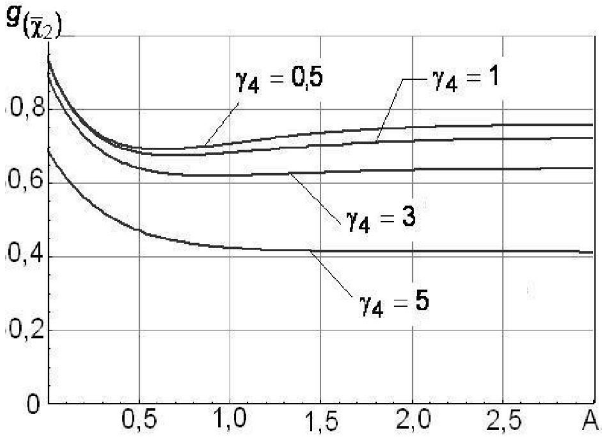

Рис. 2.

З графіка видно, що при степені полінома $s = 4$ дисперсія оцінки менша ніж при степені полінома $k = 2$. При цьому, при прямуванні значення γ_4 до правого краю області визначення (-0,327; 9,623) [3] величина $g_{(\overline{\chi_2})}$, а тому і дисперсія, що знайдена з рівняння (14), зменшується. Високі значення інтервалу кореляції τ_0 (1) між вибірковими значеннями, що взяті з досліджуваного процесу, призводять до збільшення дисперсії оцінок. При значеннях параметра A кореляційної функції більших 1,5 вплив кореляційного зв'язку на значення дисперсії відсутній.

Висновки

1. Отримані формули для знаходження оцінки дисперсії ексцесного корельованого випадкового сигналу при його моментно-кумулянтному описі з використаням стохастичних поліномів до четвертого степеня.
2. Отримані формули для визначення дисперсії знайденої оцінки. Показана залежність дисперсії оцінки від значень інтервалу кореляції між вибірковими значеннями.
3. Показано, що ефективність оцінки збільшується з ростом степеня полінома.
4. Показано, що на значення дисперсії оцінки впливають параметри негауссовості (коефіцієнт ексцесу), тому врахування в алгоритмі обробки випадкових даних всіх моментних функцій дозволяє підвищувати точність оцінювання.

Література

1. *Андерсон Т.* Введение в многомерный статистический анализ: Пер. с англ. Под ред. Б.В. Гнеденко – М.: Гос. изд-во физ.-мат. лит., 1963. – 500 с.
2. *Кендалл М.*, Стюарт А. Статистические выводы и связи: Пер. с англ. / Под ред. А.Н. Колмогорова – М.: Наука, 1973. — 900 с.
3. *Малахов А.Н.* Кумулянтный анализ случайных негауссовых процессов и их преобразование. – М.: Сов. радио, 1978. – 376 с.
4. *Кунченко Ю.П.* Полиномиальные оценки параметров близких к гауссовским случайных величин. Часть I. Стохастические полиномы, их свойства и применение для нахождение оценок параметров. – Черкассы: ЧИТИ, 2001. – 133 с.
5. *Кунченко Ю.П.* Стохастические полиномы. – К.: Наукова Думка, 2006. – 275 с.
6. *Палагін В.В.*, Івченко О.В. Адаптація методу максимізації полінома для оцінки параметрів випадкових величин за статистично-залежною вибіркою // Системи обробки інформації. – Харків, 2009. – Вип. 2(76). – С. 118–123.
7. *Палагін В.В.*, Івченко О.В. Особливості оцінювання параметрів статистично залежних випадкових величин // Вісник ЧДТУ. – 2009. – № 1. – С. 73–78.
8. *Крамер Г.* Математческие методы статистики: Пер. с англ. / Под ред. А.Н. Колмогорова. – М.: Мир, 1975. – 684 с.

Черкасский государственный технологический университет

Поступила в редакцию 18 марта 2010 г.