

## МЕТОДЫ И СИСТЕМЫ ОБРАБОТКИ ИЗМЕРИТЕЛЬНОЙ ИНФОРМАЦИИ НА ЭВМ

УДК 681.833 : 519.2

Р. Н. ВИЛЬДАНОВ, А. Н. ДОМАРАЦКИЙ, Л. Н. ИВАНОВ,  
В. А. ПОПОВ, А. Н. СМОРОДИНОВ, Ю. И. ЮРЛОВ

(Новосибирск)

### ВЫБОР ПАРАМЕТРОВ АЛГОРИТМОВ ПРИКЛАДНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА

В научных исследованиях прочно утвердился прикладной статистический анализ стационарных случайных сигналов, в котором основную роль играет определение корреляционных функций, спектральных плотностей, одномерных амплитудных и временных распределений. Практика последних лет остро ставит вопрос о проведении этого анализа в ритме с экспериментом, когда получаемые статистические характеристики могут быть использованы не только для идентификации объекта исследования, но и для управления течением эксперимента. Такой анализ не всегда удается проводить на ЭВМ. Поэтому представляется целесообразным и своевременным исследовать алгоритмы получения корреляционных функций, спектральных плотностей, одномерных распределений и выбор их параметров с целью повышения эффективности и точности анализа при сохранении низких затрат оборудования на реализацию специализированных вычислителей. Не представляет сомнений, что в связи с достигнутым уровнем современной микроэлектроники и повышенными требованиями к точности анализа наиболее перспективными являются цифровые алгоритмы.

Начнем рассмотрение с алгоритма определения корреляционных функций. Как известно [1, 2], корреляционные функции определяются по алгоритму

$$K_x(\mu\Delta t) = \frac{1}{N} \sum_{v=1}^N x(t_v) x(t_v + \mu\Delta t), \quad (1)$$

где  $\Delta t$  — период дискретизации реализации стационарного случайного сигнала  $x(t)$ ;  $\mu = 0, 1, 2, \dots, l$ ;  $l\Delta t = \tau_{\max}$  — максимальное значение аргумента корреляционной функции;  $x(t_v)$  — дискретизированные по уровню значения реализации сигнала в дискретные моменты времени  $t_v$ . По поводу выбора  $\Delta t$  и количества ступеней дискретизации сигнала по уровню имеется достаточное количество работ [2—4], в которых можно найти соответствующие рекомендации. Здесь мы остановимся лишь на выборе параметра  $\mu\Delta t$  алгоритма (1).

При использовании естественного течения времени для реализации временной задержки  $\mu\Delta t$  возможны два основных варианта построения специализированного вычислителя [5, 6], которые представлены на рис. 1, а и б ([5], [6]). В [5] за время  $\Delta t$  производится  $l$  сдвигов в ре-

гистре, т. е. за время одного сдвига необходимо произвести все вычислительные и логические операции над кодами чисел  $x(t_v)$  и  $x(t_v + \mu\Delta t)$ , а минимальное значение аргумента корреляционной функции  $\Delta\tau_{\min}$  определяется величиной  $\Delta t = l\delta t$ , ограниченной длительностью операций  $\delta t$ . В [6]  $\Delta\tau_{\min} = \Delta t = \delta t$ , т. е. в  $l$  раз меньше, чем в [5]. Это говорит о том, что реализация [6] алгоритма (1) расширяет его возможности, однако вызывает значительное увеличение затрат оборудования.

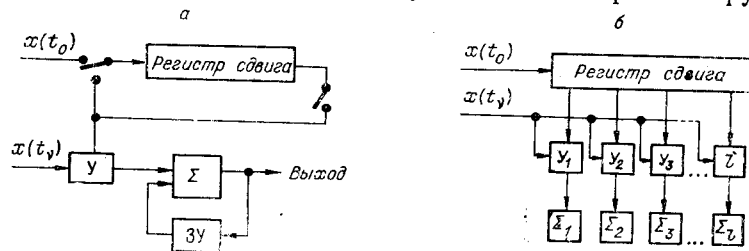


Рис. 1.

Возможен компромиссный вариант [7], при котором без значительного увеличения оборудования шаг  $\Delta\tau_{\min}$  можно сделать принципиально сколь угодно малым. Действительно, если (1) представить в виде

$$K_x(i\delta\tau + \mu\Delta t) = \frac{1}{N} \sum_{v=1}^N x(t_v) x(t_v + i\delta\tau + \mu\Delta t), \quad (2)$$

где  $i=0, 1, 2, \dots, m$ ;  $\mu=0, 1, 2, \dots, l/m$ , и определять  $K_x(\tau)$  в отдельных каналах, для которых фиксировано  $i$ , то начальный сдвиг  $\delta t$  между каналами практически может быть любым. Затраты оборудования на реализацию алгоритма (1) и (2) в зависимости от длины сдвигающего регистра приведены на рис. 2, где кривая затрат на реализацию алгоритма (2) в зависимости от  $l$  располагается в заштрихованной области. Преимущество алгоритма (2) заключается в том, что минимальный шаг дискретизации по времени  $\Delta\tau_{\min}$  корреляционной функции можно сделать много меньше  $\Delta t$  ( $\Delta\tau_{\min}$  определяется лишь временем переходных процессов входных устройств). Это дает возможность повысить верхнюю граничную частоту обрабатываемых сигналов в десятки раз при низких скоростях обработки в  $i$ -х каналах.

Таким образом, применение алгоритма (2) в специализированном вычислителе не влечет за собой больших затрат оборудования и расширяет возможности статистического анализа во временной области без привлечения специальных разработок. Последнее позволяет сделать вывод о целесообразности вычисления корреляционных функций путем

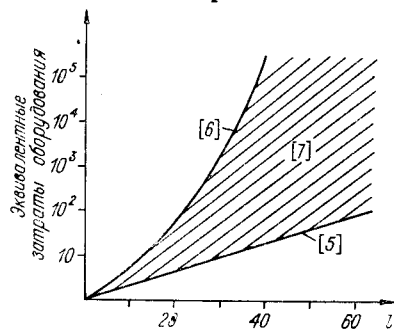


Рис. 2.

обработки реализаций сигналов во времени, а не путем преобразования Фурье спектральных плотностей. Кроме того, специализированный вычислитель, работающий по алгоритму (2), дает возможность использовать  $m$  каналов для одновременного определения других характеристик прикладного статистического анализа по одному (одновременно не менее  $l/m$  точек) или  $m$  каналам, что, с одной стороны, повышает быстродействие алгоритмов в  $m$  раз (например, классического алгорит-

ма преобразования Фурье), с другой — дает пространственную или временную картину изменения определяемых характеристик.

Итак, при многоцелевом статистическом анализе спектральные плот-

$$a_k = \frac{2}{N} \sum_{r=0}^{N-1} f_r \cos \frac{\pi}{N} rk; \quad (3)$$

$$b_k = \frac{2}{N} \sum_{r=1}^{N-1} f_r \sin \frac{\pi}{N} rk, \quad (4)$$

где  $a_k$ ,  $b_k$  — приближенные значения коэффициентов Фурье;  $f_r$  — значения преобразуемой функции;  $N$  — количество отсчетов преобразуемой функции.

При вычислениях по классическому алгоритму для нахождения одного коэффициента Фурье необходимо произвести  $N$  считываний величин  $f_r$  и  $\cos \frac{\pi}{N} rk$  из памяти и выполнить столько же операций умножения. (Суммирование можно не учитывать, так как суммирование парных произведений легко совместить с суммированием частичных произведений при перемножении.) Всего для вычисления  $N$  коэффициентов понадобится  $N^2$  перемножений.

Относительно алгоритма БПФ [8] следует отметить, что он основан на рекурсивной процедуре ряда комплексных подпоследовательностей, первой из которых является множество значений преобразуемой функции  $\{f_r\}$ . Для нахождения  $N$  коэффициентов здесь требуется всего  $2N \log_2 N$  перемножений. Так как умножение наиболее емкая по времени операция, то быстроедействие БПФ при прочих равных условиях намного больше, чем быстроедействие классического алгоритма. Однако сложность логической структуры БПФ, а также его рекурсивность приводят к значительному усложнению управляющего и арифметического устройств вычислителя. Простота логической структуры классического алгоритма и его нерекурсивность позволяют создать вычислитель, быстроедействие которого может быть повышено за счет конструктивных упрощений.

Основная операция при вычислении коэффициентов — получение парных произведений  $f_r \cos \frac{\pi}{N} rk$ . Поэтому время работы алгоритма и затраты оборудования на его реализацию существенно зависят от разрядности сомножителей при представлении чисел с фиксированной запятой. При ограниченном числе разрядов значений  $f_r$  можно найти ту оптимальную разрядность представления  $\cos \frac{\pi}{N} rk$ , которая, с одной стороны, не приведет к значительной потере точности вычисления коэффициентов Фурье, с другой — не вызовет чрезмерного увеличения количества оборудования арифметического устройства вычислителя. Представление  $\cos x$  двоичными разрядами с фиксированной запятой эквивалентно квантованию на  $2^m$  равноотстоящих уровней или аппроксимации  $\cos x$  ступенчатой функцией. На рис. 3 показаны некоторые способы аппроксимации, причем их отличие сводится к разным способам округления значений  $\cos x$  до  $m$  разрядов. Существует много способов квантования тригонометрических функций по уровню. Можно, например, квантовать  $\cos x$  на  $P$  уровней так, что  $P$  гармоник (после первой) в разложении

аппроксимирующей ступенчатой функции в ряд Фурье станут нулями. Однако при этом величины уровней будут представляться значительно большим числом разрядов, чем  $\log_2 P$  [9]. Квантование на  $P$  равноотстоящих уровней дает возможность представить все  $P$  значений уровней минимумом разрядов.

Анализ погрешностей, возникающих при квантовании  $\cos x$  ( $\sin x$ ) на  $P = 2^m$  уровней, показывает, что абсолютная погрешность  $\Delta A_k$  ( $\Delta B_k$ ) вычисления коэффициентов Фурье может быть оценена следующим образом:

$$\Delta A_k^-, (\Delta B_k) \leq \frac{2}{\pi P k} \sum_{r=0}^{N-1} |f_r|. \quad (5)$$

Моделирование специализированного вычислителя на УВМ показало, что при  $m$ -разрядном представлении преобразуемой функции  $f_r$  тригонометрические функции достаточно задавать  $m - 2$  разрядами. Тогда

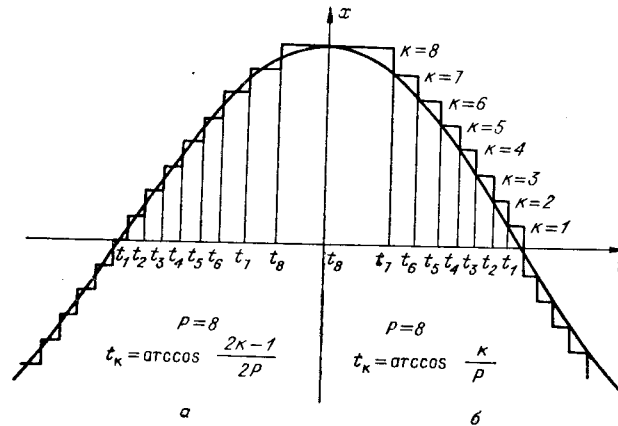


Рис. 3.

погрешность вычисления практически определяется ограниченностью разрядности функции  $f_r$ , т. е. нерекурсивность классического алгоритма позволяет использовать тригонометрические функции, аппроксимированные ступенчатыми функциями с небольшим числом уровней квантования. Это значительно упрощает устройство хранения и выдачи  $\cos x$  ( $\sin x$ ) в специализированном вычислителе и дает возможность при определении коэффициентов Фурье оперировать с малоразрядными числами.

Основные блоки анализатора, реализующего алгоритмы (2) — (4), могут быть использованы и для вычисления одномерных плотностей вероятностей. В этом случае наиболее важным для обеспечения достаточной точности вычислений является правильный выбор таких параметров алгоритма, как ширина дифференциального коридора и время анализа. В [10] рассмотрено влияние конечности интервала наблюдения на погрешность вычисления плотности вероятности для нормальных случайных сигналов. Определим точность оценки плотности вероятности в зависимости от времени анализа  $T$  и ширины дифференциального коридора  $\Delta x$  для случайных сигналов с произвольными законами распределения.

Пусть  $x(t)$  — некоторая реализация стационарного эргодического случайного сигнала, заданная на интервале  $(0, T)$ , а  $f(x)$  — его одномерная плотность вероятности. Экспериментальное определение плотности вероятности за фиксированное время  $T$  основано на измерении от-

носительного времени пребывания реализации сигнала в дифференциальном коридоре. Образовав из реализации  $x(t)$  новый процесс  $\eta(t)$ , определим оценку  $f_T(x)$  как

$$f_T(x) = \frac{1}{T\Delta x} \int_0^T \eta(t) dt, \quad (6)$$

где

$$\eta(t) = \begin{cases} 0, & \text{если } x(t) < x - \frac{\Delta x}{2}; \\ 1, & \text{если } x - \frac{\Delta x}{2} < x(t) < x + \frac{\Delta x}{2}; \\ 0, & \text{если } x(t) > x + \frac{\Delta x}{2}. \end{cases} \quad (7)$$

Определим смещенность оценки  $f_T(x)$ , вычислив ее математическое ожидание

$$M[f_T(x)] = \frac{1}{T\Delta x} \int_0^T M[\eta(t)] dt, \quad (8)$$

или с учетом (7) получим

$$M[f_T(x)] = \frac{1}{\Delta x} \int_{x - \frac{\Delta x}{2}}^{x + \frac{\Delta x}{2}} f(x) dx = \frac{F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right)}{\Delta x}, \quad (9)$$

где  $F(x)$  — функция распределения  $x(t)$ . Из (9) следует, что оценка  $f_T(x)$  смещенная.

Для характеристики точности оценки (6) определим отдельно среднее значение смещения оценки и ее дисперсию. Среднее значение смещения оценки равно

$$m(x) = M[f_T(x) - f(x)] = \frac{F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right)}{\Delta x} - f(x) \quad (10)$$

и с уменьшением  $\Delta x$  стремится к нулю.

Дисперсия оценки  $f_T(x)$  определяется как

$$\sigma^2(x) = M[f_T(x) - M[f_T(x)]]^2 = M[f_T^2(x)] - M^2[f_T(x)]. \quad (11)$$

Учитывая условие эргодичности  $\eta(t)$ , выражение (11) можно привести к виду [1]

$$\sigma^2(x) = \frac{2}{T\Delta x^2} \int_0^T \left(1 - \frac{\tau}{T}\right) R_\eta\left(x \pm \frac{\Delta x}{2}, \tau\right) d\tau - M^2[f_T(x)], \quad (12)$$

где  $R_\eta\left(x \pm \frac{\Delta x}{2}, \tau\right)$  — смешанный начальный момент второго порядка процесса  $\eta(t)$ . Для центрированного процесса

$$\sigma^2(x) = \frac{2}{T\Delta x^2} \int_0^T \left(1 - \frac{\tau}{T}\right) K_\eta\left(x \pm \frac{\Delta x}{2}, \tau\right) d\tau, \quad (13)$$

где  $K_\eta\left(x \pm \frac{\Delta x}{2}, \tau\right)$  — корреляционная функция процесса  $\eta(t)$ . Заметим, что время корреляции процесса  $\eta(t)$  изменяется в зависимости от уровня анализа и не может быть больше, чем время корреляции  $\tau_k$  процесса  $x(t)$  [1]; тогда, завышая (13), будем иметь

$$\sigma_1^2(x) = \frac{2\tau_k}{T\Delta x^2} \sigma_\eta^2, \quad (14)$$

где  $\sigma_{\eta}^2$  — дисперсия сигнала  $\eta(t)$ ;  $\tau_k$  — интервал корреляции сигнала  $x(t)$ . При этом с учетом (9)

$$\sigma_{\eta}^2 = \int_{x - \frac{\Delta x}{2}}^{x + \frac{\Delta x}{2}} \eta^2(t) f(x) dx - M^2[\eta(t)] = F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right) - \left[F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right)\right]^2. \quad (15)$$

Тогда вместо (14) получим

$$\sigma_1^2(x) = \frac{2 \left\{ F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right) - \left[ F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right) \right]^2 \right\}}{\Delta x^2} \frac{\tau_k}{T}. \quad (16)$$

В отличие от (10) дисперсия оценки (6) увеличивается с уменьшением  $\Delta x$ . Результирующая относительная среднеквадратическая ошибка оценки  $f_T(x)$  равна

$$\delta(x) = \sqrt{\frac{\sigma_1^2(x) + m^2(x)}{f^2(x)}}. \quad (17)$$

Подставив (10) и (16) в (17), имеем

$$\delta(x) = \sqrt{\frac{\frac{2\tau_k}{T f^2(x) \Delta x^2} \left\{ F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right) - \left[ F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right) \right]^2 \right\} + \frac{\left[ F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right) \right]^2}{f^2(x) \Delta x^2}}{\frac{2 \left[ F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right) \right]}{f(x) \Delta x}} + 1}. \quad (18)$$

Разложим  $F\left(x + \frac{\Delta x}{2}\right)$  и  $F\left(x - \frac{\Delta x}{2}\right)$  в окрестности точки  $x$  в ряд Тейлора, тогда разность этих функций будет равна

$$F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right) = 2 \sum_{n=1}^{\infty} \frac{F^{(2n-1)}(x)}{(2n-1)!} \left(\frac{\Delta x}{2}\right)^{2n-1}. \quad (19)$$

Бесконечную сумму в (19) можно заменить на  $f(t) \Delta x$  для нормального закона распределения при  $\Delta x \leq 0,39\sigma$ , для релеевского — при  $\Delta x \leq 0,26$  и для экспоненциального — при  $\Delta x \leq \frac{0,28}{\lambda}$  с погрешностью 0,001.

Для равновероятного закона эта замена производится точно. Поэтому (18) можно привести к виду

$$\delta(x) = \sqrt{\frac{2\tau_k}{T} \left( \frac{1}{f(x) \Delta x} - 1 \right)}. \quad (20)$$

На рис. 4—6 соответственно представлены зависимости относительной среднеквадратической погрешности от отношения  $\frac{T}{\tau_k}$  при разных

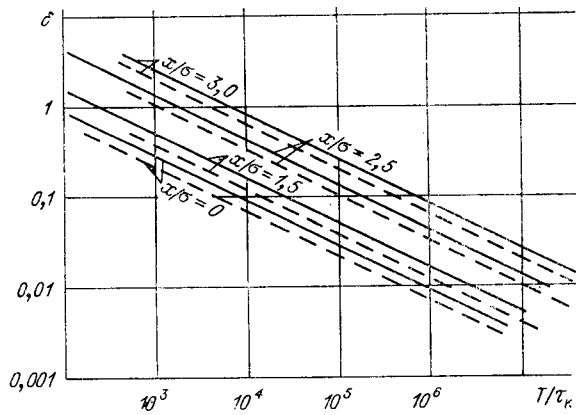


Рис. 4.

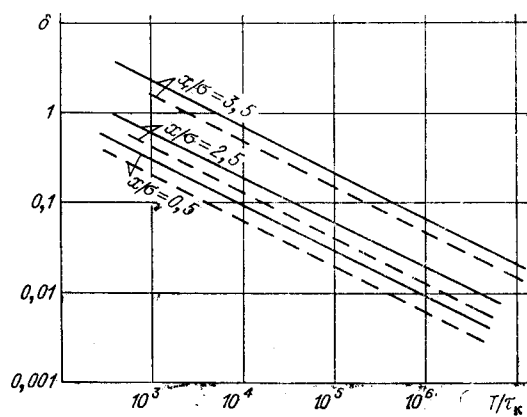


Рис. 5.

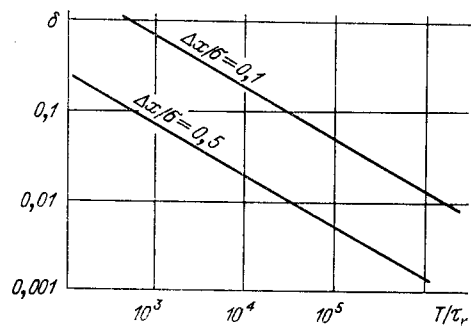


Рис. 6.

уровнях анализа и ширине дифференциального коридора для нормального (непрерывные кривые при  $\frac{\Delta x}{\sigma} = 0,06$ , штриховые —  $\frac{\Delta x}{\sigma} = 0,1$ ), релейского (непрерывные кривые при  $\frac{\Delta x}{\sigma} = 0,05$ , штриховые —  $\frac{\Delta x}{\sigma} = 0,1$ ) и равновероятного ( $f(x) = \frac{1}{2a}$ ) законов распределения, вычисленные по (20). Из (20) следует, что

$$\Delta x = \frac{2\tau_k}{Tf(x)\left(\delta^2 + \frac{2\tau_k}{T}\right)}. \quad (21)$$

Формула (21) имеет ясный физический смысл. Если  $\frac{\tau_k}{T}$  и  $\delta$  фиксированы, то  $\Delta x$  и  $f(x)$  связаны обратно пропорциональной зависимостью, т. е. при вычислении малых вероятностей [ $f(x)$  близко к 0]  $\Delta x$  должно быть большим в пределах, указанных при выводе (20). Для фиксированных  $\Delta x$  и  $\delta$  при вычислении малых вероятностей необходимо увеличивать отношение  $\frac{T}{\tau_k}$ . Полученная оценка точности (20) формулы (6) раскрывает эмпирическое соотношение  $\Delta x T \approx \text{const}$  [10] и показывает, что величину дифференциального коридора необходимо выбирать не только в зависимости от интервала анализа, но и от вида плотности распределения. Это подтверждает общее правило, по которому при оценке точности статистических характеристик надо знать их вид и точные значения.

В заключение отметим, что правильный выбор параметров алгоритмов прикладного статистического анализа позволяет не только повысить его эффективность, но и создавать относительно простые многоцелевые анализаторы.

#### ЛИТЕРАТУРА

1. Н. А. Лившиц, В. Н. Пугачев. Вероятностный анализ систем автоматического управления, т. I. М., «Советское радио», 1963.
2. А. Н. Домарацкий, Л. Н. Иванов, Е. Н. Карышев, Б. С. Сеницын. Дискретная измерительная корреляционная система (ДИКС). Новосибирск, «Наука», 1965.
3. А. А. Косякин, Г. Ф. Филаретов. Построение цифровых корреляторов.— Идентификация и аппаратура для статистических исследований. М., «Наука», 1970.
4. D. G. Watts. A General Theory of Amplitude Quantisation with Application to Correlation Determination.— Proc. IEE, v. 109, № 15, 1962.
5. Ю. Я. Зотов. Быстродействующие корреляторы с синхронными накопителями.— Труды II Всесоюзного симпозиума «Методы представления и аппаратурный анализ случайных процессов и полей», т. II. Новосибирск, 1969.
6. Любау. Корреляционный анализ сигналов, выполняемый в реальном масштабе времени.— Электроника (русс. перевод), 1966, № 22.
7. А. Н. Домарацкий, Л. Н. Иванов, Е. Н. Карышев, В. А. Попов, А. И. Скурлатов, А. Н. Смородинов. Коррелятор. Авторское свидетельство № 292169.— ОИПОТЗ, 1971, № 4.
8. W. T. Cochran, J. W. Cooley. What is the Fast Forier Transform?— IEEE Trans. Audio and Electroacoustics, 1967, v. AU—15, juni.
9. С. С. Тонкаль. К гармоническому анализу ступенчатой функции, аппроксимирующей синусоидальное напряжение.— Проблемы технической электродинамики, вып. 24. Киев, 1970.
10. Г. Я. Мирский. Аппаратурное определение характеристик случайных процессов. М.— Л., «Энергия», 1967.
11. В. В. Губарев, Ю. К. Постоенко. Неравенство для среднеквадратических погрешностей оценок вероятностных характеристик случайных процессов.— Труды II Всесоюзного симпозиума «Методы представления и аппаратурный анализ случайных процессов и полей», т. II. Новосибирск, 1969.

Поступила в редакцию  
27 мая 1971 г.