УДК 621.301

О. В. Брягин[1], **А. К. Егоров**[2], **Г. Н. Розоринов**[2]
[1]Министерство внутренних дел Украины
ул. Богомольца 10, 01024 Киев, Украина
[2]Национальный технический университет Украины
«Киевский политехнический институт»
проспект Победы, 37, 03056 Киев, Украина

Об оценке многомерных функций распределения вероятностей речевых сигналов

Предложены алгоритмы оценивания многомерных функций распределения вероятностей с усреднением по ансамблю реализаций и по времени. Приведены структурные схемы устройств, реализующих эти алгоритмы. Показано, что оценки многомерных функций распределения вероятностей — несмещенные и состоятельные.

Ключевые слова: *случайные процессы, многомерные функции, доверительная вероятность.*

Введение

Одним из наиболее актуальных и сложных вопросов в современной статистической радиотехнике, электроакустике, теории связи, статистическом анализе, теории проверки статистических гипотез и во многих других прикладных областях науки и техники является вопрос о виде многомерных распределений вероятностей анализируемых или обрабатываемых случайных процессов. Как правило, эту проблему не решают, а ограничиваются допущениями о том, что рассматриваемые процессы либо имеют нормальное распределение, либо процессы марковские, либо отсчеты процессов независимы [1, 2]. Связано это с тем, что практически единственным известным многомерным распределением вероятностей является многомерное нормальное распределение, а предположение о марковости или независимости отсчетов случайного процесса позволяет заменить многомерное распределение произведением одномерных или условных распределений вероятностей. Вместе с тем, имеется дольно широкий круг задач, для которых предположение, например, о нормальности или о марковости процесса может оказаться не совсем корректным. К таким процессам относятся прежде всего нестационарные случайные процессы, например, речевые и видео сигналы.

В статье предлагается физически реализуемый способ оценки многомерных функций распределения вероятностей случайных процессов.

На рис. 1 показана структура ранее предложенного в [3] устройства для оценки многомерных распределений, которую в дальнейшем будем называть устройством совмещения событий.

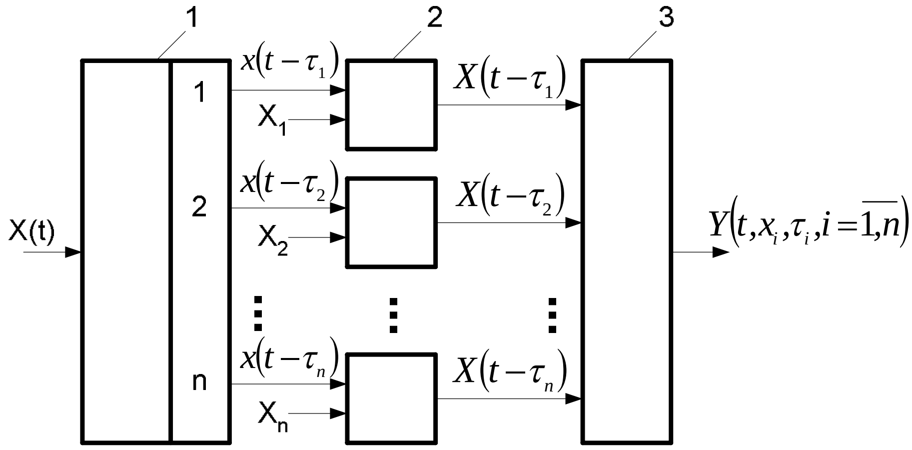

Рис. 1. Структура устройства совмещения событий

При поступлении на вход элемента задержки 1, имеющего n выходов случайного процесса $x(t)$, на его выходах формируются n процессов $x(t-\tau_i),\ i=\overline{1,n}$. Эти процессы подаются на пороговые элементы 2 (компараторы), где сравниваются с величинами $x_i,\ \ i=\overline{1,n}$, в соответствии с алгоритмом:

$$X(t-\tau_i)=\begin{cases}1 & при\ \ x(t-\tau_i)\le x_i \\ 0 & при\ \ x(t-\tau_i)> x_i\end{cases},\qquad i=\overline{1,n}. \tag{1}$$

Полученные в результате сравнений случайные последовательности единиц и нулей поступают на входы n-входового элемента **И**3, на выходе которого формируется процесс, представляющий собой совмещение процессов $X(t-\tau_i),\ i=\overline{1,n}$, т.е.

$$y(t,x_i,\tau_i)=\bigcap_{i=1}^{n}X(t-\tau_i)=\begin{cases}1, & при\ x(t-\tau_i)\le x_i,\forall\ i=\overline{1,n}\\ 0, & при\ невыполнении\ хотя\ бы\ одного\ условия\ (1).\end{cases} \tag{2}$$

С учетом (1) и (2) математическое ожидание процесса $y(t,x_i,\tau_i,i=\overline{1,n})$ будет равно:

$$E\{y(t,x_i,\tau_i,i=\overline{1,n})\}=P\{x(t-\tau_1)\le x_1,x(t-\tau_2)\le x_2,...,x(t-\tau_n)\le x_n\}. \tag{3}$$

Основываясь на определении многомерной функции распределения вероятностей [4, 5], приходим к выводу, что математическое ожидание, получаемое на выходе устройства совмещения событий, совпадает с многомерной функцией распределения вероятностей случайного процесса $x(t)$. Следовательно, оценка многомерной функции распределения вероятностей случайного процесса $x(t)$ сводится к статистическому усреднению отклика $y(t, x_i, \tau_i, i = \overline{1,n})$ устройства совмещения событий.

Рассмотрим возможные процедуры оценки многомерной функции распределения вероятностей.

Оценка с усреднением по ансамблю реализаций

Структура устройства для оценки многомерной функции распределения вероятностей путем статистического усреднения по ансамблю реализаций показана на рис. 2.

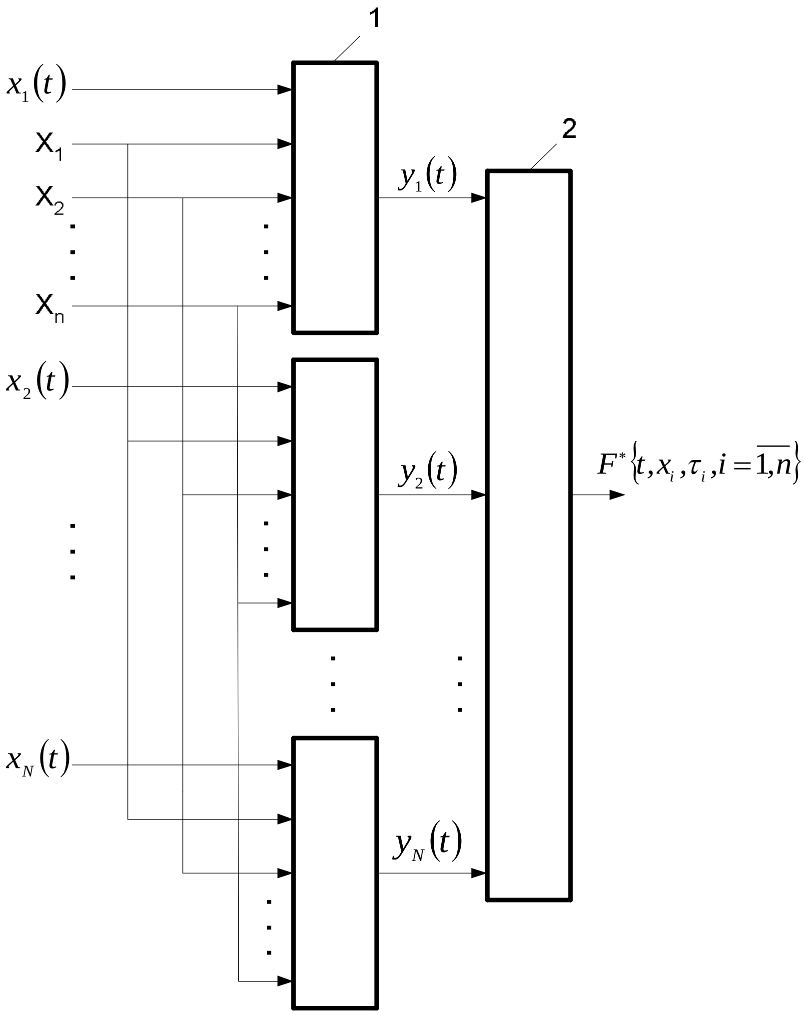

Рис. 2. Устройство для оценки n-мерной функции распределения вероятностей с усреднением по ансамблю реализаций

Каждая i-я реализация $x_i(t), i=\overline{1,N}$, исследуемого процесса поступает на вход соответствующего устройства 1 совмещения событий, на выходах каждого из которых формируются процессы $y_j(t, x_i, \tau_i, i=\overline{1,n}),\ j=\overline{1,N}$, типа (2). Процессы $y_j(t, x_i, \tau_i, i=\overline{1,n}),\ j=\overline{1,N}$, подаются на сумматор 2 (аналоговый или цифровой), на выходе которого формируется оценка n-мерной функции распределения вероятностей процесса $x(t)$ вида

$$F_N^*(t, x_i, \tau_i, i=\overline{1,n}) = \frac{1}{N}\sum_{i=1}^{N} y_i(t, x_i, \tau_i, i=\overline{1,n}). \qquad (4)$$

Найдем математическое ожидание и дисперсию оценки (4). С учетом (3) для математического ожидания имеем:

$$E\left\{F_N^*(t, x_i, \tau_i, i=\overline{1,n})\right\} = \frac{1}{N}\sum_{i=1}^{N} E\left\{y_i(t, x_i, \tau_i, i=\overline{1,n})\right\} = F\left\{x(t-\tau_i) \le x_i,\ i=\overline{1,n}\right\}, \qquad (5)$$

где $F\left\{x(t-\tau_i) \le x_i, i=\overline{1,n}\right\} = P\left\{x(t-\tau_1 \le x_1, x(t-\tau_2) \le x_2, \ldots, x(t-\tau_n) \le x_n\right\}$ — n-мерная функция распределения вероятностей процесса $x(t)$. Следовательно, оценка (4) — несмещенная. С учетом (4), (5) и в случае, когда процессы $x_i(t), i=\overline{1,n}$, представляют собой ансамбль независимых реализаций, для дисперсии оценки (4) получаем:

$$D\{F_N^*(t, x_i, \tau_i,\ i=\overline{1,n})\} = \frac{1}{N}\left(F\left\{x(t-\tau_i) \le x_i, i=\overline{1,n}\right\} - F^2\left\{x(t-\tau_i) \le x_i, i=\overline{1,n}\right\}\right). \qquad (6)$$

Отметим, что так как для любой, в том числе и многомерной функции распределения вероятностей, имеет место соотношение $0 \le F\{x\} \le 1$, то из (6) следует, что дисперсия оценки $F^*(t, x_i, \tau_i, i=\overline{1,n})$ удовлетворяет условию

$$D\left\{F_N^*(t, x_i, \tau_i,\ i=\overline{1,n})\right\} \le \frac{0{,}25}{N}, \qquad (7)$$

то есть не зависит от вида распределения исследуемого процесса.

Оценка с усреднением по времени

На рис. 3 показана структура устройства оценки многомерной функции распределения вероятностей случайного процесса путем усреднения отклика устройства 1 совмещения событий по времени.

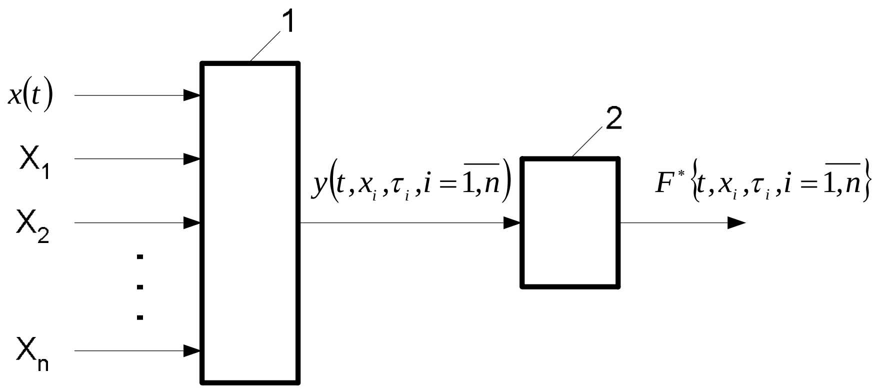

Рис. 3. Устройство для оценки *n*-мерной функции распределения вероятностей с усреднением по времени

Для этого требуется устройство накопления 2, например интегратор с постоянной интегрирования T . В этом случае оценка многомерной функции распределения вероятностей равна:

$$F_T^*(t, x_i\tau_i, i=\overline{1,n}) = \frac{1}{T}\int_0^T y(t, x_i, \tau_i, i=\overline{1,n})dt\,. \tag{8}$$

Тогда математическое ожидание оценки $F_T^*(t, x_i, \tau_i, i=\overline{1,n})$ будет:

$$E\left\{F_T^*(t, x_i, \tau_i, i=\overline{1,n})\right\} = \frac{1}{T}\int_0^T E\left\{y(t, x_i, \tau_i, i=\overline{1,n})\right\}dt\,, \tag{9}$$

что для стационарных эргодических процессов приводит к равенству

$$E\left\{F_T^*(t, x_i, \tau_i, i=\overline{1,n})\right\} = F\left\{x(t-\tau_i) \le x_i, i=\overline{1,n}\right\}. \tag{10}$$

Как видим и при усреднении по времени оценка многомерной функции распределения вероятностей — несмещенная, по крайней мере, для стационарных случайных процессов. Для дисперсии оценки (8), в случае стационарных случайных процессов получим выражение:

$$D\left\{F_T^*(t, x_i, \tau_i, i=\overline{1,n})\right\} =$$

$$= \frac{1}{T^2}\int_0^T\int_0^T (E\left\{y(t_1, x_i, \tau_i, i=\overline{1,n})y(t_2, x_i, \tau_i, i=\overline{1,n})\right\} - F^2\left\{x(t-\tau_i) \le x_i, i=\overline{1,n}\right\})dt_1 dt_2\,, \tag{11}$$

где, с учетом (2) и (3):

$$E\left\{y(t_1,x_i,\tau_i,i=\overline{1,n})\,y(t_2,x_i,\tau_i,i=\overline{1,n})\right\}=F\left\{x(t_1-\tau_i)\le x_i,\;\; x(t_2-\tau_i)\le x_i, i=\overline{1,n}\right\} \quad (12)$$

— неизвестная $2n$-мерная функция распределения вероятностей процесса $x(t)$. Это делает расчет дисперсии оценки весьма сложным. Однако, как показано в [7], при $T\to 0$ дисперсия $D\left\{F_T^*(t,x_i,\tau_i,i=\overline{1,n})\right\}$ стремится к нулю. Там же отмечается, что, как правило, интегрирование в (8) не может быть выполнено аналитически и в качестве альтернативы предлагается численное интегрирование выборок процесса $y(t,x_i,\tau_i,i=\overline{1,n})$, равноотстоящих через промежутки времени ΔT. Таким образом, если $y_1=y(\Delta T,x_i,\tau_i,i=\overline{1,n}),\;\; y_2=y(2\Delta T,x_i,\tau_i,i=\overline{1,n}),\ldots,y_N=(N\Delta T,x_i,\tau_i,i= =\overline{1,n})$ и $N\Delta T=T$, то оценка многомерной функции распределения $F^*(t,x_i,\tau_i, i=\overline{1,n})$ может быть представлена в виде:

$$F^*(t,x_i,\tau_i,i=\overline{1,n})=\frac{1}{N}\sum_{i=1}^{N}y_i\,, \qquad (13)$$

что для статистически независимых выборок приводит к аналогичным (6) и (7) выражениям для дисперсии оценки (13).

В заключение этого раздела отметим возможность статистического усреднения отклика устройства совмещения событий с помощью предложенного в [8] t-текущего интегратора, позволяющего отслеживать характер изменений многомерной функции распределения вероятностей во времени. При этом оценка находится в соответствии с правилом

$$F_T^*\left\{t,x_i,\tau_i,i=\overline{1,n}\right\}=\frac{1}{T}\int_t^{t+T}y(t,x_i\tau_i,i=\overline{1,n})\,dt\,. \qquad (14)$$

Очевидно, что для стационарных эргодических процессов математическое ожидание и дисперсия этой оценки будут совпадать с аналогичными параметрами оценки (8).

Точность оценки многомерных функций распределения вероятностей

Доверительные вероятности для предлагаемых оценок многомерных функций распределения можно определить двумя способами. Первый из них заключается в использовании для определения доверительной вероятности результатов измерений неравенства Чебышева [6]. При этом для доверительной вероятности рассматриваемых оценок можно записать:

$$P\left(\left|F_N^*(t,x_i,\tau_i,i=\overline{1,n})-F\{x(t-\tau_i)\le x_i,i=\overline{1,n}\}\right|>\delta\right)\le\frac{D\left\{F_N^*(t,x_i,\tau_i,i=\overline{1,n})\right\}}{\delta^2}, \quad (15)$$

где δ — максимально допустимое отклонение оценки от истинного значения n-мерной функции распределения вероятностей, или с учетом (7), более жесткое условие

$$P\left(\left|F_N^*(t,x_i,\tau_i,i=\overline{1,n})-F\{x(t-\tau_i)\le x_i,i=\overline{1,n}\}\right|>\delta\right)\le\frac{0{,}25}{N\delta^2}, \tag{16}$$

в котором N равно либо числу реализаций в ансамбле, по которому производится статистическое усреднение, либо числу выборок на выходе устройства совмещения событий при усреднении по времени. Отметим, что этот критерий накладывает довольно жесткие условия на величину N, особенно при нахождении малых значений функции распределения. Так, например, для обеспечения условия

$$P\left(\left|F_N^*(t,x_i,\tau_i,i=\overline{1,n})-F\{x(t-\tau_i)\le x_i,i=\overline{1,n}\}\right|>0{,}01\right)\le 0{,}1$$

из (16) получаем $N\ge 1000$, что иногда довольно сложно реализовать практически.

Второй подход к решению вопроса о точности оценки многомерных функций распределения вероятностей основан на том, что при достаточно больших N можно полагать отклонение полученной оценки от ее математического ожидания распределенным по нормальному закону [1, 4, 7]. При этом для доверительной вероятности получаем:

$$P\left(\left|F_N^*(t,x_i,\tau_i,i=\overline{1,n})-F\{x(t-\tau_i)\le x_i,i=\overline{1,n}\}\right|>\delta\right)\le 1-2\Phi\left(\frac{\delta}{\sqrt{D\left\{F_N^*(t,x_i,\tau_i,i=\overline{1n})\right\}}}\right), \tag{17}$$

где $\Phi(x)$ — интеграл вероятностей, или, с учетом (16):

$$P\left(\left|F_N^*(t,x_i,\tau_i,i=\overline{1,n})-F\{x(t-\tau_i)\le x_i,i=\overline{1,n}\}\right|>\delta\right)\le 1-2\Phi\left(2\delta N\right). \tag{18}$$

При этом для приведенного выше примера условие

$$P\left(\left|F_N^*(t,x_i,\tau_i,i=\overline{1,n})-F\{x(t-\tau_i)\le x_i,i=\overline{1,n}\}\right|>0{,}01\right)\le 0{,}1$$

будет выполняться уже при $N\ge 350$. Как видим, условия (15), (16) предъявляют к величине N гораздо более жесткие требования.

В качестве примера в табл. 1, 2 приведены результаты измерения семимерной функции распределения вероятностей с усреднением по ансамблю и по времени соответственно для 10000 независимых реализаций или отсчетов гауссовского случайного процесса.

Таблица 1. Усреднение по ансамблю реализаций

x_1	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	–0,2500	1,2500
x_2	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	0,0000	1,0000
x_3	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	0,2500	0,7500
x_4	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	0,5000	0,5000
x_5	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	0,7500	0,2500
x_6	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	1,0000	0,0000
x_7	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	1,2500	–0,2500
F_7^*	0,0014	0,0080	0,0275	0,0739	0,1660	0,2973	0,4664	0,6232	0,0484	0,0493
${}_0F_7^*$	0,0017	0,0078	0,0260	0,0733	0,1643	0,2963	0,4619	0,6200	0,0479	0,0478
${}_0F_7$	0,0017	0,0078	0,0276	0,0756	0,1655	0,2984	0,4577	0,6163	0,0483	0,0483
δ	0,0003	0,0002	0,0001	0,0017	0,0005	0,0011	0,0087	0,0069	0,0001	0,0010

Таблица 2. Усреднение по времени

x_1	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	–0,2500	1,2500
x_2	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	0,0000	1,0000
x_3	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	0,2500	0,7500
x_4	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	0,5000	0,5000
x_5	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	0,7500	0,2500
x_6	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	1,0000	0,0000
x_7	–0,2500	0,0000	0,2500	0,5000	0,7500	1,0000	1,2500	1,5000	1,2500	–0,2500
F_7^*	0,0027	0,0096	0,0281	0,0766	0,1692	0,2982	0,4616	0,6071	0,0483	0,0497
${}_0F_7^*$	0,0016	0,0078	0,0288	0,0770	0,1675	0,3025	0,4643	0,6076	0,0487	0,0487
${}_0F_7$	0,0020	0,0078	0,0276	0,0756	0,1655	0,2984	0,4577	0,6163	0,0495	0,0495
δ	0,0007	0,0018	0,0005	0,0010	0,0037	0,0002	0,0039	0,0092	0,0012	0,0002

В таблицах приняты такие обозначения:

$$F_7^* = F^*\left(x_1, t;\ x_2, t-\tau;\ x_3, t-2\tau;\ \ldots,\ x_7, t-6\tau\right) = P^*\left\{x(t) < x_1, x(t-\tau) < x_2,\ x(t-3\tau) < x_3, \ldots,\ x(t-6\tau) < x_7\right\}$$

— оценка семимерной функции распределения вероятностей;

$${}_0F_7^* = \prod_{i=1}^{7} F_1^*\left(x_i, t-(i-1)\tau\right) = \prod_{i=1}^{7} P^*\left\{x(t-(i-1)\tau) < x_i\right\}$$

— произведение оценок одномерных функций распределения вероятностей для соответствующих отсчетов анализируемого процесса;

$${}_0F_7 = \prod_{i=1}^{7} F_1\left(x_i, t-(i-1)\tau\right)$$

— произведение расчетных (теоретических) значений одномерных функций распределения вероятностей для соответствующих отсчетов анализируемого процесса;

$\delta = \left| \; F_7^* - {}_0F_7 \; \right|$ — отклонение оценки семимерной функции распределения вероятностей от расчетного значения.

Выводы

1. Разработаны алгоритмы оценивания многомерных функций распределения вероятностей с усреднением по ансамблю реализаций и по времени.
2. Приведены структурные схемы устройств, реализующих эти алгоритмы.
3. Показано, что получаемые оценки многомерных функций распределения вероятностей — несмещенные и состоятельные.
4. Приведены формулы для расчета доверительных вероятностей и доверительных интервалов оценок.
5. Полученные результаты могут быть использованы не только для многомерного вероятностного анализа случайных процессов, но и для проверки гипотез: о стационарности случайного процесса; о марковости случайного процесса; о независимости отсчетов случайного процесса.

1. *Кендалл М. Дж., Стюарт А.* Многомерный статистический анализ и временные ряды. — М.: Наука, 1976. — 736 с.
2. *Левин Б.Р.* Теоретические основы статистической радиотехники. Кн. 2. — М.: Сов. радио, 1968. — 552 с.
3. Авт. свид. СССР. Устройство для измерения многомерных функций распределения вероятностей / Белов С.В., Егоров А.К., Железняк В.К. — № 234924; Опубл. 01.04.86.
4. *Левин Б.Р.* Теоретические основы статистической радиотехники. Кн. 1. — М.: Сов. радио, 1974. — 504 с.
5. *Корн Г., Корн Т.* Справочник по математике. — М.: Наука, 1984. — 832 с.
6. Унковский В.А. Теория вероятностей. — М.: Военно-морское изд-во, 1953. — 320 с.
7. *Купер Дж., Макгиллем К.* Вероятностные методы анализа сигналов и систем. — М.: Мир, 1989. — 376 с.
8. *Козанне А., Флере Ж., Метр Г., Руссо М.* Оптика и связь. — М.: Мир, 1984. — 504 с.

Поступила в редакцию 27.07.2004