

Е. А. ТРОИЦКИЙ  
(Ленинград)

**ОЦЕНКИ МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ  
И ДИСПЕРСИИ СЛУЧАЙНЫХ ПРОЦЕССОВ  
ПО ВЫБРОСАМ\***

Целью статьи является анализ оценок математического ожидания  $\mu$  и дисперсии  $\sigma^2$  эргодических нормальных случайных процессов по выбросам их реализаций и сравнение эффективности и удобства применения предлагаемых оценок с обычно используемыми, которые вычисляются по формулам:

$$\bar{x} = \frac{1}{T} \int_0^T \xi(t) dt; \quad (1)$$

$$S^2 = \frac{1}{T} \int_0^T [\xi(t) - \mu]^2 dt, \quad (2)$$

где  $\bar{x}$ ,  $S^2$  — соответственно оценки  $\mu$  и  $\sigma^2$  по мгновенным значениям процесса;  $T$  — длительность реализации процесса  $\xi(t)$ .

Выбросом реализации процесса за некоторый уровень  $a_j$  будем считать пересечение этого уровня снизу вверх. Среднее число выбросов в зависимости от значения нормированного уровня

$$\gamma_j = \frac{a_j - \mu}{\sigma} \quad (3)$$

с точностью до постоянного множителя равно одномерной плотности вероятностей для целого ряда стационарных случайных процессов. Если плотность вероятностей зависит от  $n$  параметров, то для выражения их через среднее число выбросов необходимо рассмотреть систему из  $(n+1)$  уравнений, каждое из которых представляет собой зависимость частоты выбросов  $E_j$  от выбранного уровня  $\gamma_j$ . Решая эту систему, можно получить выражения для каждого из  $n$  параметров. Очевидно, такой путь вычисления параметров процесса имеет смысл, если общее число пересечений в единицу времени уровня  $a_j$  конечно с вероятностью 1. Этим свойством обладают, в частности, нормальные эргодические случайные

\* Материал доложен на II Всесоюзном симпозиуме «Методы представления и аппаратный анализ случайных процессов и полей» (г. Сухуми, 1969 г.).

$$L_j = \frac{1}{2\pi} \int_{-\infty}^{\infty} R''(0) \exp\left(-\frac{1}{2} \gamma_j^2\right) d\gamma_j, \quad (4)$$

где  $R''(0)$  — вторая производная от корреляционной функции при значении аргумента, равном нулю. Из формул (3) и (4) следует, что в общем случае частота выбросов за уровень  $a_j$  зависит от параметров  $\mu$  и  $\sigma^2$  процесса, а также от коэффициента  $E_0 = \frac{1}{2\pi} [-R''(0)]^{0.5}$ , представляющего собой среднее число выбросов за уровень  $a_j = \mu$ .

Поэтому для выражения дисперсии и математического ожидания через среднее число выбросов необходима система из трех уравнений:

$$\begin{cases} E_1 = E_0 \exp\left(-\frac{1}{2} \gamma_1^2\right); \\ E_2 = E_0 \exp\left(-\frac{1}{2} \gamma_2^2\right); \\ E_3 = E_0 \exp\left(-\frac{1}{2} \gamma_3^2\right). \end{cases}$$

Если значения верхнего  $a_2$  и нижнего  $a_3$  уровней удовлетворяют условию

$$a_2 - a_1 = a_1 - a_3 = b,$$

то решение системы упрощается и нетрудно получить следующие выражения для  $\mu$  и  $\sigma^2$  через частоты выбросов процесса:

$$\mu = \frac{\Lambda_{13} - \Lambda_{12}}{\Lambda_{13} + \Lambda_{12}} \frac{b}{2} + a_1; \quad (5)$$

$$\sigma^2 = \frac{b^2}{\Lambda_{13} + \Lambda_{12}}, \quad (6)$$

где

$$\Lambda_{13} = \ln \frac{E_1}{E_3}; \quad \Lambda_{12} = \ln \frac{E_1}{E_2}.$$

По реализации процесса конечной длительности  $T$  можно лишь оценить среднюю частоту выбросов  $E$  и в качестве оценки данного параметра естественно принять отношение общего числа выбросов  $N_j(T)$  реализации  $\xi(t)$  за  $j$ -й уровень к ее длительности, т. е.

$$\bar{N}_j = \frac{N_j(T)}{T}.$$

В связи с этим, согласно (5) и (6), имеем следующие оценки математического ожидания дисперсии процесса по выбросам:

$$\bar{x}_3 = \frac{V_{13} - V_{12}}{V_{13} + V_{12}} \frac{b}{2} + a_1; \quad (7)$$

$$S_3^2 = \frac{b^2}{V_{13} + V_{12}}, \quad (8)$$

где  $\bar{x}_3, S_3^2$  — оценки параметров соответственно  $\mu$  и  $\sigma^2$  по выбросам реализации процесса за три уровня:  $V_{13} = \ln \frac{\bar{N}_1}{\bar{N}_3}$ ;  $V_{12} = \ln \frac{\bar{N}_1}{\bar{N}_2}$ . Для вычисления математического ожидания либо дисперсии процесса по выбросам, когда один из названных параметров известен, необходима система из двух уравнений, аналогичных (4). В этом случае получаем следующие оценки математического ожидания и дисперсии случайного процесса:

$$\bar{x}_2 = \frac{a_1 + a_2}{2} + \frac{V_{12} \sigma^2}{b}; \quad (9)$$

$$S_2^2 = \frac{c_2^2 - c_1^2}{2 V_{12}}, \quad (10)$$

где  $\bar{x}_2, S_2^2$  — оценки параметров соответственно  $\mu$  и  $\sigma^2$  по выбросам реализации процесса за два уровня:  $c_1 = a_1 - \mu$ ,  $c_2 = a_2 - \mu$ . Для определенности примем  $|c_1| < |c_2|$ . Если при этом априори известна еще и средняя частота выбросов  $E_0$ , то математическое ожидание либо дисперсию процесса можно оценить по выбросам реализации процесса за один уровень. Такой способ оценки дисперсии центрированного случайного процесса уже используется для решения ряда конкретных технических задач, например для измерения температуры [2]. Данная оценка имеет вид

$$S_1^2 = \frac{c^2}{2 V_0}, \quad (11)$$

где

$$V_0 = \ln \frac{E_0}{\bar{N}}; \quad c = a - \mu.$$

Рассмотрим требования по выбору уровней при расчете оценок параметров  $\mu$  и  $\sigma^2$  по полученным алгоритмам. Уровни, которые расположены вне области, ограниченной экстремальными значениями реализации процесса (уровни без пересечений), не могут быть использованы, так как они, естественно, не несут информации о вероятностных свойствах случайного процесса. Более того, уровни  $a, a_1, a_2$  и  $a_3$  целесообразно располагать между экстремальными значениями реализации таким образом, чтобы число их пересечений было достаточным для вычисления оценки частоты выбросов. С этой точки зрения уровни желательно выбирать вблизи среднего значения реализации, так как в этом случае происходит наибольшее число их пересечений.

Далее, любая оценка дисперсии случайного процесса должна быть строго положительной величиной. Для выполнения данного условия, согласно выражениям (8) и (10), расстояние между уровнями и их расположение относительно реализации должны быть такими, чтобы число выбросов за уровень  $a_1$  было больше, чем за другие уровни. Для оценки (11) имеем, что положение уровня  $a$  должно обеспечить выполнение условия  $E_0 > \bar{N}$ .

Таким образом, уровень  $a_1$  целесообразно располагать вблизи среднего значения реализации процесса, а уровни  $a, a_2, a_3$  должны быть расположены соответственно по отношению к уровням  $\mu$  и  $a_1$  по возможности на значительном расстоянии, но не вблизи экстремальных значений реализации. Более детально вопрос о рациональном расположении уровней, в частности о расстоянии между ними, при достаточно большой длительности реализации процесса будет рассмотрен ниже.

В соответствии с требованиями относительно выбора уровней получаем следующие ограничения на область возможных значений случайных величин  $\bar{N}$ ,  $\bar{N}_1$ ;  $\bar{N}_2$ ,  $\bar{N}_3$ , а именно:

$$\bar{N}_2 > 0; \bar{N}_3 > 0; \bar{N}_1 > \bar{N}_2; \bar{N}_1 > \bar{N}_3; \bar{N} < E_0.$$

Если при выбранных уровнях вычисленные значения оценок  $\bar{N}$ ,  $\bar{N}_1$ ,  $\bar{N}_2$ ,  $\bar{N}_3$  находятся вне области допустимых значений, то необходимо изменить положение уровней либо воспользоваться другой реализацией процесса.

Проверим состоятельность и несмещенность полученных оценок. Для рассматриваемого типа случайных процессов оценка среднего числа выбросов за уровень с вероятностью 1 стремится к оцениваемому параметру  $E_j$  при длительности реализации  $T \rightarrow \infty$  [1], т. е. оценка является состоятельной. Отсюда следует, что оценки (7)–(11) математического ожидания и дисперсии процесса по выбросам также обладают свойством сходимости по вероятности, так как они являются непрерывными функциями в допустимой области значений аргументов  $\bar{N}$ ,  $\bar{N}_1$ ,  $\bar{N}_2$ ,  $\bar{N}_3$ .

При достаточной длительности реализации процесса с учетом состоятельности оценки  $\bar{N}_j$  функции (7)–(11) можно считать приближенно линейными относительно случайных величин  $\bar{N}$ ,  $\bar{N}_1$ ,  $\bar{N}_2$ ,  $\bar{N}_3$  в допустимой области их значений, расположенной в окрестности точки  $(E, E_1, E_2, E_3)$ .

При этом условии математические ожидания оценок (7)–(11) равны оцениваемым параметрам  $\mu$  и  $\sigma^2$ .

Таким образом, получены асимптотически несмещенные и состоятельные оценки математического ожидания и дисперсии нормальных эргодических случайных процессов по выбросам. Этот метод имеет ряд преимуществ перед расчетом оценок тех же параметров по алгоритмам (1) и (2): во-первых, он проще, так как вместо измерения мгновенных значений случайного процесса необходимо произвести подсчет числа выбросов его реализации; во-вторых, требуется меньше времени при ручной обработке статистического материала. Далее, процесс расчета оценок параметров  $\mu$  и  $\sigma^2$  по выбросам можно легко автоматизировать путем использования счетных схем, определяющих число пересечений случайным процессом выбранных уровней. Эта аппаратура значительно проще, чем применяемая для вычисления оценок тех же параметров по формулам (1) и (2). Кроме того, выражение (8) в отличие от (2) позволяет оценить дисперсию процесса без предварительного расчета оценки математического ожидания.

Проанализируем зависимость дисперсий оценок параметров случайного процесса по выбросам от значений уровней и рассчитаем их эффективность относительно оценок (1) и (2). Анализ выполним на примере флюктуаций показаний прибора, который представляет собой динамическую систему второго порядка, возбуждаемую нормальным «белым» шумом. Выбранная модель представляет существенный интерес, так как описывает многие практически интересные процессы, имеющие различную физическую природу. Кроме того, в некоторых случаях поведение систем с порядком характеристического уравнения выше второго может быть приближенно описано характеристическим уравнением второго порядка. Выражение для частоты выбросов рассматриваемого типа случайного процесса имеет вид [3]

$$E_j = \frac{1}{T_0} \exp\left(-\frac{1}{2} \gamma_j^2\right), \quad (12)$$

где  $T_0$  — период собственных колебаний системы. Дисперсии оценок (7)–(11) при условии их линейности, т. е. при достаточно большом  $T$ , равны:

$$D(\bar{x}_3) = \frac{\sigma^2}{4\Delta\gamma^4} \left[ 16\gamma_1^2 \frac{D(\bar{N}_1)}{E_1^2} + (2\gamma_1 - \Delta\gamma)^2 \frac{D(\bar{N}_2)}{E_2^2} + \right. \\ \left. + (2\gamma_1 + \Delta\gamma)^2 \frac{D(\bar{N}_3)}{E_3^2} - 8\rho_{12} \frac{\gamma_1(2\gamma_1 - \Delta\gamma)}{E_1 E_2} - \right. \\ \left. - 8\rho_{13} \frac{\gamma_1(2\gamma_1 + \Delta\gamma)}{E_1 E_3} + 2\rho_{23} \frac{4\gamma_1^2 - \Delta\gamma^2}{E_2 E_3} \right]; \quad (13)$$

$$D(\bar{x}_2) = \frac{\sigma^2}{(\gamma_2 - \gamma_1)^2} \left[ \frac{D(\bar{N}_1)}{E_1^2} + \frac{D(\bar{N}_2)}{E_2^2} - 2\rho_{12} \frac{1}{E_1 E_2} \right]; \quad (14)$$

$$D(S_3^2) = \frac{\sigma^4}{\Delta\gamma^4} \left[ 4 \frac{D(\bar{N}_1)}{E_1^2} + \frac{D(\bar{N}_2)}{E_2^2} + \frac{D(\bar{N}_3)}{E_3^2} - \right. \\ \left. - 4\rho_{12} \frac{1}{E_1 E_2} - 4\rho_{13} \frac{1}{E_1 E_3} + 2\rho_{23} \frac{1}{E_2 E_3} \right]; \quad (15)$$

$$D[S_2^2] = \frac{4\sigma^4}{(\gamma_2^2 - \gamma_1^2)^2} \left[ \frac{D(\bar{N}_1)}{E_1^2} + \frac{D(\bar{N}_2)}{E_2^2} - 2\rho_{12} \frac{1}{E_1 E_2} \right]; \quad (16)$$

$$D[S_1^2] = \frac{4\sigma^4}{\gamma^4} \frac{D(\bar{N})}{E^2}, \quad (17)$$

где  $D(\bar{N})$ ,  $D(\bar{N}_1)$ ,  $D(\bar{N}_2)$ ,  $D(\bar{N}_3)$  — дисперсии оценок частот выбросов соответственно за уровни  $\gamma$ ,  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$ ;  $\rho_{12}$ ,  $\rho_{13}$ ,  $\rho_{23}$  — коэффициенты ковариации соответственно между  $\bar{N}_1$  и  $\bar{N}_2$ ,  $\bar{N}_1$  и  $\bar{N}_3$ ,  $\bar{N}_2$  и  $\bar{N}_3$ ;  $\Delta\gamma = \gamma_2 - \gamma_1$ .

Выбросы случайного процесса приближенно можно считать пуассоновским потоком событий, если в течение среднего времени между выбросами за уровень  $a_j = \mu$  корреляционная связь между ординатами процесса практически исчезает [4], т. е. значение  $1/E_0$  должно быть больше интервала корреляции  $\tau_k$ . Примем, что корреляционная связь между ординатами процесса отсутствует, если она меньше 0,10. Тогда при степени успокоения системы  $\beta > 1$  дисперсия оценки среднего числа выбросов процесса приближенно может быть представлена в виде

$$D(\bar{N}_j) = \frac{E_j}{T}. \quad (18)$$

Далее с целью упрощения анализа полагаем  $\rho_{12} = \rho_{13} = \rho_{23} = 0$ . Чем больше расстояние между уровнями, тем более оправданным является такое допущение. Как будет показано ниже, для достижения минимума дисперсий оценок параметров  $\mu$  и  $\sigma^2$  процесса верхний и нижний уровни должны отстоять друг от друга на значительное расстояние (не менее чем на  $3\sigma$ , где  $\sigma$  — среднее квадратическое отклонение).

Результаты расчетов дисперсий оценок по формулам (13)–(17) в зависимости от выбранных уровней для случая  $\beta > 1$  с учетом (12), (18) и принятых допущений, т. е.  $\rho_{12} = \rho_{13} = \rho_{23} = 0$ , приведены в таблице. Они даны в виде отношения дисперсий оценок параметров  $\mu$  и  $\sigma^2$  соответственно к  $q_1$  и  $q_2$ , где  $q_1 = \sigma^2 \frac{T_0}{T}$ ,  $q_2 = \sigma^4 \frac{T_0}{T}$ . Выбор  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma$ ,  $\Delta\gamma$  вне диапазонов, рассматриваемых в таблице, приводят, как показали

Дисперсии оценок параметров  $\mu$  и  $\sigma^2$  процесса по выбросам

|            |            | $\vartheta_1 = \frac{D(\bar{x}_2)}{q_1}$ |      |      | $\vartheta_2 = \frac{D(\bar{x}_3)}{q_1}$ |            |      | $\vartheta_3 = \frac{D(S_3^2)}{q_2}$ |      | $\vartheta_4 = \frac{D(S_3^2)}{q_2}$ |      | $\vartheta_5 = \frac{D(S_1^2)}{q_2}$ |  |
|------------|------------|--|------|------|--|------------|------|--------------------------------------|------|--------------------------------------|------|--------------------------------------|--|
| $\gamma_2$ | $\gamma_1$ | -1                                       | -1,5 | -2   | $\Delta\gamma$                           | $\gamma_1$ | 0    | $\pm 0,5$                            | 0    | $\pm 0,5$                            | 0    | 0                                    |  |
| 1          |            | 0,83                                     | 0,76 | 1,00 | 1  | 0,83       | 2,30 | 7,30                                 | 8,90 | 10,6                                 | 6,60 |                                      |  |
| 1,5        |            | 0,76                                     | 0,69 | 0,87 | 1,5                                      | 0,62       | 0,80 | 2,00                                 | 2,71 | 3,2                                  | 2,44 |                                      |  |
| 2          |            | 1,00                                     | 0,87 | 0,92 | 2  | 0,90       | 0,87 | 1,20                                 | 1,82 | 2,1                                  | 1,86 |                                      |  |
| 2,5        |            | 2,00                                     | 1,60 | 1,47 | 2,5                                      | 1,80       | 2,31 | 1,24                                 | 2,60 | 2,5                                  | 2,36 |                                      |  |

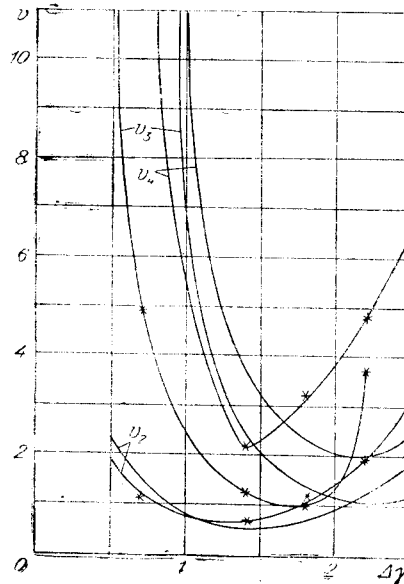
расчеты, к увеличению дисперсии оценок. Анализ данных таблицы показывает, что существуют оптимальные значения уровней, обеспечивающих минимум дисперсий оценок. Для оценки  $\bar{x}_2$  они близки к  $\gamma_1 = -1,5$ ,  $\gamma_2 = +1,5$ , для оценок  $\bar{x}_3$ ,  $S_3^2$ ,  $S_2^2$  и  $S_1^2$  минимум дисперсии оценок достигается при  $\gamma_1 = 0$  и  $\Delta\gamma$ , соответственно равном 1,5; 2,3; 2,0 и  $\gamma = 2,0$ . Отклонения нормированных уровней от приведенных значений до  $\pm 0,5\sigma$  приводят к незначительному увеличению дисперсий оценок, т. е. функции в окрестности экстремальных точек являются пологими. Кроме того, оценка дисперсии по выбросам за три уровня является более эффективной, чем за два.

На рисунке приведены графики теоретических и экспериментальных значений  $\vartheta_2$ ,  $\vartheta_3$ ,  $\vartheta_4$  в зависимости от  $\Delta\gamma$  при  $\gamma_1 = 0$ . Экспериментальные значения на графиках отмечены звездочками.

Экспериментальные данные получены в результате статистической обработки 20 реализаций длительностью  $T = 96 T_0$  записи шумов на выходе фотогальванометрического автокомпенсатора НФК-1. Флюктуации на выходе нановольтметра НФК-1, который может быть представлен в качестве динамической системы второго порядка, в основном обусловлены тепловым шумом сопротивления цепи рамки гальванометра. Следовательно, флюктуации показаний прибора НФК-1 практически являются рассматриваемым типом случайного процесса. Экспериментальные данные подтверждают наличие оптимальных значений уровней. Некоторое расхождение между экспериментальными и расчетными значениями дисперсий оценок, очевидно, объясняется ограниченностью статистического материала и приближенным характером расчета теоретических значений дисперсий оценок параметров  $\mu$  и  $\sigma^2$  по выбросам.

Рассмотрим теперь эффективность оценки (7) относительно оценки (1), а также оценки (8) относительно оценки (2) при  $\beta > 1$ ,  $\gamma_1 = 0$  и оптимальных значениях  $\Delta\gamma$ .

Согласно данным таблицы, имеем следующие выражения для относительных эффективностей:



$$\nu_1 = \frac{D(\bar{x})}{D(\bar{x}_3)} = \frac{D(\bar{x})}{0,62 q_1}; \quad \nu_2 = \frac{D(S^2)}{D(S_3^2)} = \frac{D(S^2)}{1,2 q_2}.$$

Для исследуемого типа случайного процесса в соответствии с выражениями для дисперсий оценок (1) и (2), приведенных в [4], считая  $T \gg \tau_k$ , относительно эффективности  $\nu_1$  и  $\nu_2$  равны:  $\nu_1 = \beta$ ,  $\nu_2 = 0,53\beta + 0,13$ . Отсюда следует, что при  $\beta > 1$  и  $\beta > 1,7$  оценки соответственно  $\bar{x}_3$  и  $S_3^2$  более эффективны, чем оценки  $\bar{x}$  и  $S^2$  при достаточно большой длительности реализаций. Следовательно, наряду с простотой вычисления, предлагаемые оценки обладают достаточно высокой эффективностью. Это позволяет рекомендовать оценки параметров  $\mu$  и  $\sigma^2$  случайного процесса по выбросам его реализации вместо оценок (1) и (2) при большом объеме статистического материала или для целей экспресс-анализа, так как в этом случае целесообразно использовать даже малоэффективные оценки, если они приводят к существенному упрощению расчетов.

### ЗАКЛЮЧЕНИЕ

Получены простые асимптотически несмещенные и состоятельные оценки математического ожидания и дисперсии нормального эргодического случайного процесса по выбросам его реализации в общем случае за три уровня. Применение этих оценок в случае большого объема статистического материала или для целей экспресс-анализа целесообразнее, чем вычисление оценок названных выше параметров по мгновенным значениям процесса.

### ЛИТЕРАТУРА

1. Г. Крамер, М. Лидбеттер. Стационарные случайные процессы. М., «Наука», 1969.
2. А. Д. Бродский, А. В. Саватеев. Новый метод абсолютного измерения температуры.— Измерительная техника, 1960, № 5.
3. С. Г. Рабинович. Анализ естественных флюктуаций показаний гальванометрических автокомпенсаторов.— Труды институтов Комитета стандартов, вып. 107 (167). М., 1969.
4. А. А. Свешников. Прикладные методы теории случайных функций. М., «Наука», 1968.

*Поступила в редакцию  
14 апреля 1970 г.,  
окончательный вариант —  
24 сентября 1970 г.*