

Особенности исследования методом сингулярного спектрального анализа случайных процессов, распределенных по экспоненциальному закону

Албкеират Д.М.; Якимов Е.А.
Кафедра АСУ, электротехнический факультет
ГУВПО «Белорусско-Российский университет»
г. Могилев, Беларусь
e-mail: e-soft@bk.ru

Аннотация — Рассматриваются результаты исследования последовательностей данных с экспоненциальным шумом путем преобразования методом сингулярного спектрального анализа (SSA-методом). Показываются особенности восстановления шумовой составляющей, предложен алгоритм корректировки восстановленной шумовой составляющей с учетом стандартного отклонения..

Ключевые слова: SSA-метод, последовательность данных; экспоненциальное распределение; коррекция

I. ВВЕДЕНИЕ

Одной из основных проблем при подготовке и проведении стохастических экспериментов в имитационном моделировании является правильный выбор модели входных данных, т.е. закона распределения вероятностей, который точно соответствует поведению случайных входных процессов. При этом стандартные законы распределения могут не соответствовать вероятностному поведению реальных входных процессов, особенно на границах распределения.

Несмотря на проведенные исследования, задача установления типа выравнивающей кривой распределения и нахождения оценок параметров последовательностей данных, хранящихся в информационной системе предприятия, к настоящему времени до конца не решена. Такая ситуация предполагает дальнейшее исследование новых методов обработки данных, в частности, метода сингулярного спектрального анализа (SSA-метода) [1].

II. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

Обычно исследуют поведение сложной системы при полной неопределенности, состоящей из двух компонентов: во-первых, изменчивость, определенная случайными процессами в системе и именуемая стохастической неопределенностью, случайной неопределенностью или непреодолимой неопределенностью; во-вторых, неопределенность, обусловленная нехваткой знаний или достаточной информации о системе. Поэтому предлагается учитывать оба компонента при формировании последовательности входных данных для имитационной модели, например, предприятия общественного питания:

$$Exp([3; 10]) \rightarrow Norm([2; 5], 8) \rightarrow Norm([4; 7], [2; 8]),$$

где $Exp([3; 10])$ – источник поступления заявок, $Norm([2; 5], 8)$ – время обслуживания заявки узла 1; $Norm([4; 7], [2; 8])$ – время обслуживания заявки узла 2.

Для исследования по методике [2] принята последовательность данных (временной ряд) G с экспоненциальным шумом

$$G = Exp(\gamma; \beta)_n, n = 43.$$

где $Exp(\gamma; \beta)_n$ – функция, возвращающая n случайных чисел, распределенных по экспоненциальному закону с параметром смещения $\gamma = 0$ и параметром масштаба $\beta = 1$. Последовательность данных G получена в пакете *Minitab* (рис. 1).

После преобразования временного ряда G SSA-методом получен восстановленный для шумовой составляющей ряд \hat{G}_N , и трендовая составляющая \hat{G}_T

$$G \sim \hat{G} = \hat{G}_N + \hat{G}_T.$$

Следует обратить внимание на гистограммы распределения случайных чисел исходного временного ряда (рис. 1) и восстановленной шумовой составляющей для этого ряда (рис. 2).

На рис. 1 представлен явно выраженный экспоненциальный закон распределения случайных чисел. Восстановленная шумовая составляющая (рис. 2) отличается визуально от исходного экспоненциального типа распределения.

Проверка по критерию Колмогорова – Смирнова [3] соответствия полученных случайных чисел теоретическому экспоненциальному распределению показала, что шумовая составляющая \hat{G}_N распределена по экспоненциальному закону.

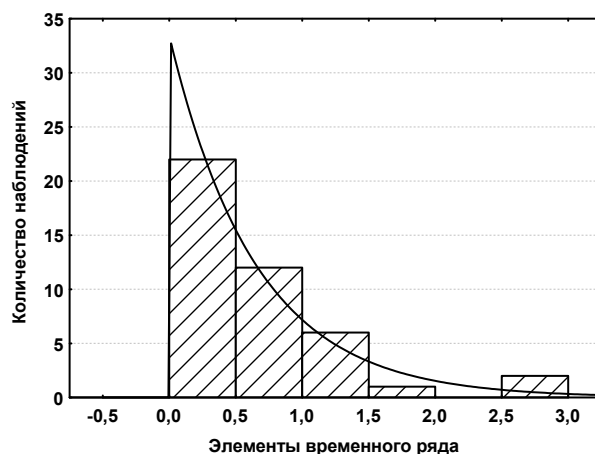


Рис. 1. Распределение случайных чисел исходного временного ряда

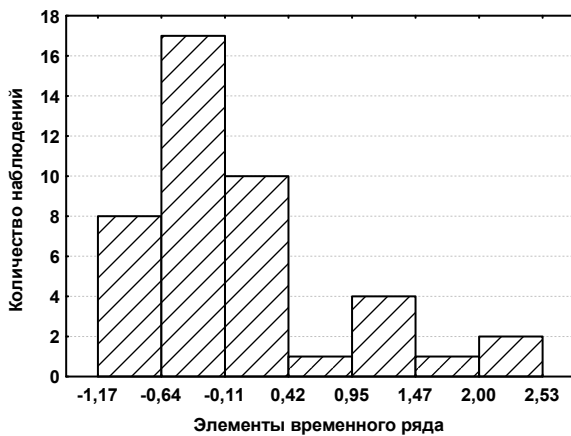


Рис. 2. Распределение случайных чисел восстановленной шумовой составляющей ряда

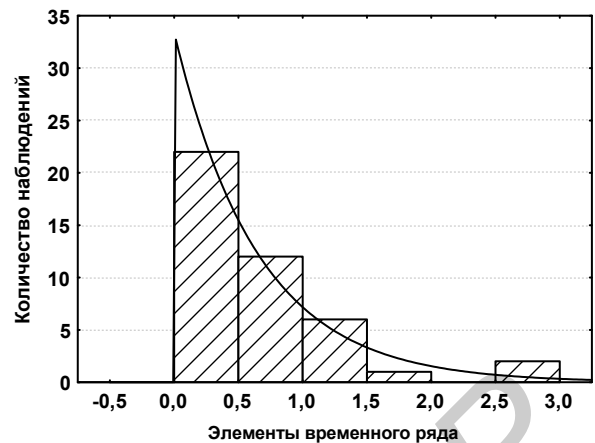


Рис. 3. Восстановленный ряд с данными, перенесенными из отрицательной области значений в положительную область

Следует отметить, что трендовая составляющая \hat{G}_T восстановленного временного ряда \hat{G} совпадает с параметром масштаба β исходного временного ряда G .

III. КОРРЕКТИРОВКА ШУМОВОЙ СОСТАВЛЯЮЩЕЙ

В связи с явными отличиями представленных гистограмм на рис. 1 и рис. 2, становится проблематичным определение типа распределения. Поэтому ставится задача исследования восстановленной шумовой составляющей для поиска алгоритма определения типа распределения. Так как шумовая составляющая расположена в отрицательной и положительной области вещественных чисел, рассмотрим восстановленный шум ряда G в следующем виде:

$$\hat{G}_{Nj} = \hat{G}_{Nj}^- + \hat{G}_{Nj}^+,$$

где \hat{G}_{Nj}^- , \hat{G}_{Nj}^+ – случайные числа соответственно в отрицательной и положительной области (рис. 2).

Исследовано соответствие восстановленных случайных величин \hat{G}_{Nj}^- и \hat{G}_{Nj}^+ теоретическому экспоненциальному распределению по критерию Колмогорова – Смирнова. Исследована также и сумма этих временных рядов по формуле

$$\hat{G}'_{Nj} = |\hat{G}_{Nj}^-| + \hat{G}_{Nj}^+,$$

где $|\hat{G}_{Nj}^-|$ – модули случайных чисел в отрицательной области. \hat{G}'_{Nj} представлен на рис. 3.

В результате проведенных исследований предлагается алгоритм для коррекции восстановленной шумовой составляющей методом сингулярного спектрального анализа, основанный на понятии динамической трендовой составляющей и стандартного отклонения элементов временного ряда.

В соответствии с алгоритмом выполняют следующие действия:

Шаг 1. Находят стандартное отклонение восстановленной шумовой составляющей, которое определяет диапазон корректирующих значений Range_{jT}^p динамического тренда по формуле

$$\text{Range}_{jT}^p \approx 0,8 \cdot s(\hat{G}_{jN}).$$

Шаг 2. Находят среднее значение диапазона коррекции Range_{jT}^p .

Шаг 3. Диапазон Range_{jT}^p относительно середины разбивается для получения n (в исследованиях $n = 43$) значений, при этом начальные значения полученного ряда являются положительными, конечные – отрицательными.

Шаг 4. Полученные на шаге 3 значения ряда суммируются с восстановленной шумовой составляющей [4].

Дополнительные исследования с дискретным равномерным шумом подтверждают полученные результаты по корректировке восстановленной шумовой составляющей.

[1] Golyandina, N. Analysis of Time Series Structure: SSA and Related Techniques / N. Golyandina, V. Nekrutkin, A. Zhigljavsky. – Boca Raton: Chapman & Hall/CRC, 2001. – 310 p.

[2] Якимов, Е. А. Исследование SSA-метода на основе комплексного применения информационных технологий / Е. А. Якимов // Доклады БГУИР. – 2010. – № 2(48). – С. 77–83.

[3] Большев, Л. Н. Таблицы математической статистики / Л. Н. Большев, Н. В. Смирнов. – М.: Наука, 1983. – 416 с.

[4] Якимов, Е. А. Особенности преобразования временных рядов методом сингулярного спектрального анализа / Е. А. Якимов, А. А. Ковалевич, Д. М. Албкеират // Вычислительный интеллект (результаты, проблемы, перспективы): материалы 1-й Междунар. науч.-техн. конф. (10–13 мая 2011 г., Черкассы). – Черкассы: Маклаут, 2011. – С. 504–505.