

ИДЕНТИФИКАЦИЯ НЕЛИНЕЙНОСТИ В РЕАЛЬНЫХ ДАНЫХ С ИСПОЛЬЗОВАНИЕМ УПРОЩЕННОГО ТЕСТА

П.И. БИДЮК, А.С. ГАСАНОВ, С.Е. ВАВИЛОВ

Рассмотрена задача идентификации нелинейности в экспериментальных данных с использованием статистических тестов. Представлен анализ известного статистического теста на нелинейность, основанного на соотношении Фишера, и предложен упрощенный тест, который может быть использован в условиях неполноты экспериментальных данных. Эмпирический критерий проверки данных на нелинейность рассчитывается на основе связи между размахом кумулятивной суммы и стандартным отклонением, вычисленным по экспериментальным данным. Эмпирически установлено существование тесной взаимосвязи между предложенным и известным тестами в смысле подобия окончательного результата тестирования. Для получения критических значений, необходимых для использования упрощенного теста, выполнены необходимые вычислительные эксперименты. Установлено, что предложенный упрощенный тест может быть использован как в условиях достаточной полноты, так и при неполных экспериментальных данных. Практическое применение различных тестов к фактическим данным свидетельствует о близости получаемых результатов.

ВВЕДЕНИЕ

Несмотря на то, что проблема выявления нелинейности, скрытой в экспериментальных данных, исследовалась в течение длительного времени, она все еще существует. Есть случаи, когда т. н. «классические подходы» не могут быть применены. В то же время для построения корректной модели очень важно иметь априорную информацию о наличии нелинейности в данных, полученных в процессе эксперимента. Описанные в литературных источниках тесты, посвященные определению нелинейности [1–5], применяются в условиях достаточности данных, т.е. когда каждая группа данных должна содержать достаточное количество наблюдений. В большинстве практических случаев эти условия проведения эксперимента не могут быть выполнены в связи с невозможностью проведения повторных экспериментов в одинаковых условиях. Это является основной причиной для создания нового метода идентификации (обнаружения) нелинейности в экспериментальных данных.

В большинстве случаев при определении нелинейности необходимо строить модель для экспериментального исследования факта нелинейности. Однако в некоторых случаях для этого необязательно строить модель, т.к. анализ на нелинейность, может успешно осуществляться с помощью соот-

ветствующих тестов. На рис. 1 приведена классификация методов определения нелинейности в экспериментальных данных.

Для оценивания типа связи между входом и выходом (линейная или нелинейная связь) можно воспользоваться спектральной функцией высокого порядка вида [5]:

$$X_{ij} = \frac{|S_{\omega}(\omega_i, \omega_j)|^2}{S_{\omega}(\omega_i)S_{\omega}(\omega_j)S_{\omega}(\omega_i / \omega_j)}, \quad (1)$$

где $S_{\omega}(\omega_i, \omega_j)$ — биспектральная плотность мощности; $S_{\omega}(\omega_i)$, $S_{\omega}(\omega_j)$ — спектральные плотности мощности временного ряда. При $S_{\omega}(\omega_i, \omega_j) = 0 \quad \forall \omega_i, \omega_j$ процесс будет линейным, а третий момент входного сигнала в этом случае $\mu_3 = 0$. Однако, если $X_{ij} = \text{const}$, то процесс линейный на рассматриваемом интервале, но при этом $\mu_3 \neq 0$.



Рис. 1. Методы идентификации нелинейности в экспериментальных данных

Такой подход к определению наличия нелинейностей имеет два недостатка. Во-первых, оценивание спектральной плотности мощности требует привлечения методов специальной предварительной обработки сигналов в виде применения временных окон, усреднения, цифровой фильтрации и т.д. Во-вторых, он не всегда может быть использован при решении задач идентификации систем, поскольку не дает возможности

получить оценки параметров модели в явном виде на основании недостаточности информативных значений входных последовательностей. В таких случаях информативный входной сигнал получают искусственно в виде специально генерируемых последовательностей, которые не всегда можно подавать на вход объекта вследствие особенностей его функционирования.

Возможно использование дисперсионного метода определения наличия нелинейностей, который основан на применении функции, вычисляемой с помощью достаточно сложного интегрального уравнения, если известны соответствующие плотности распределения вероятностей сигналов, которые также не всегда можно оценить с приемлемой точностью.

Известный критерий Фишера для обнаружения нелинейности в данных определяется выражением [1]:

$$\hat{F} = \frac{\frac{1}{k-2} \sum_{i=1}^k n_i (\bar{y}_i - \hat{y}_i)^2}{\frac{1}{n-k} \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}, \quad (2)$$

где y_{ij} — вектор измеренных значений; \bar{y} — среднее значение, определяемое по уравнению прямой; k — количество групп данных; n_i — число измерений в группе; n — общее число измерений. Этот статистический параметр может быть интерпретирован следующим образом:

$$\hat{F} = \frac{\text{Отклонение средних значений от линии регрессии}}{\text{Отклонение величины } y(k) \text{ от средних групп}}$$

Если \hat{F} статистика с $\nu_1 = k - 2$, $\nu_2 = n - k$ степенями свободы достигает или превосходит критический уровень, то гипотеза о линейности должна быть отклонена.

Вторая задача, посвященная вопросам идентификации нелинейности, скрытой в экспериментальных данных, может быть решена с помощью метода обобщенной переменной (МОП) [4]. Основная идея этого метода заключается в том, что по исходным эмпирическим данным, описывающим некоторую нелинейную гиперповерхность $y = f(x)$, восстанавливается линейная регрессия, позволяющая найти оценки направлений составляющих градиента функции отклика. С помощью этих функций можно синтезировать обобщенную переменную. Затем выборка трансформируется в другую выборку, что, в свою очередь, эквивалентно проецированию исходных многомерных данных в одномерное пространство. Соответственно, по этой выборке восстанавливается нелинейная степенная регрессия. Метод построения нелинейных моделей при наличии коротких выборок отличается относительной простотой алгоритмического и программного обеспечения.

Наиболее простым, не требующим сложных вычислений, является тест для определения наличия нелинейности, основанный на вычислении эмпирического критерия. Положительным качеством этого теста есть то, что он может быть использован также в условиях малой выборки. Описание предлагаемого теста для определения наличия нелинейности в экспериментальных данных приводится ниже.

Цель работы — по имеющимся статистическим (или экспериментальным) данным, описывающих процесс или явление предметной области, необходимо идентифицировать наличие нелинейности; в случае обнаружения нелинейности в экспериментальных данных построить соответствующую модель для анализа нелинейного процесса.

ПРЕДЛАГАЕМЫЙ ТЕСТ НА НЕЛИНЕЙНОСТЬ

Идея определения нелинейности основана на вычислении кумулятивной суммы, связанной с т.н. «кумулятивной кривой» отклонений [2]. Кривая накапливает отклонения экспериментальных данных от линии регрессии, параметры которой оцениваются методом наименьших квадратов. То есть строится линейная регрессия в форме: $\hat{y} = a_0 + a_1x$. Отклонения (остатки) e_i для всех точек от прямой линии определяются в соответствии с выражением: $e_i = y_i - a_0 - a_1x_i$ последовательно, в соответствии с возрастанием x_i , т.е., $x_1 < x_2 < \dots < x_n$. Тогда последовательные суммы отклонений (остатков) S_i определяются следующим образом:

$$s_0 = 0; s_1 = e_1; s_2 = e_1 + e_2, \dots; s_k = \sum e_k; s_n = 0. \quad (3)$$

На рис. 2 представлен график кумулятивной суммы отклонений, представляющий собой интегральную кривую.

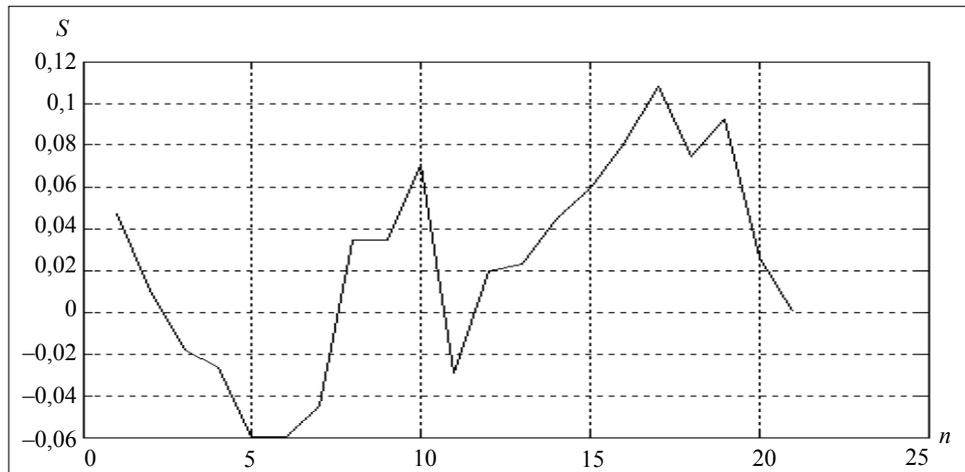


Рис. 2. График кумулятивной суммы отклонений

Основная идея метода заключается в определении диапазона, в котором изменяются отклонения (остатки) интегральной кривой и сравнение этого диапазона с критической величиной для различных уровней доверительной вероятности. Интегральная кривая представлена последовательностью сумм s_1, s_2, \dots, s_n , расположенных с равными интервалами. Эта кривая может пересекать линию регрессии в произвольном месте ближе к середине. Если эмпирические точки имеют ярко выраженную нелинейность, интегральная кривая пересечет линию регрессии только один раз (например, на рис. 3 показано, что точка пересечения находится приблизительно в середине интервала).

В результате выполнения анализа интегральной кривой определены ее максимальное и минимальное отклонения, а также установлен диапазон $R_{\max} = S_{\max} - S_{\min}$. То есть вычислена разность между максимальным и минимальным значением отклонений временного ряда от линейной аппроксимации.

Этот диапазон сравнивается с критической величиной для различных значений доверительных вероятностей. Эмпирический критерий проверки на нелинейность D рассчитывается как связь между R_{\max} и стандартным отклонением σ , вычисленным по экспериментальным данным, т.е. $D = R_{\max} / \sigma$.

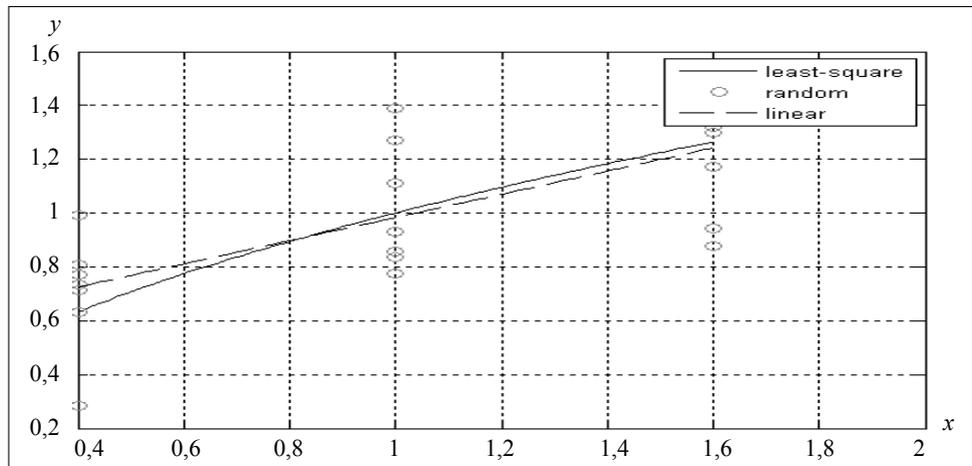


Рис. 3. График иллюстрирует выраженную нелинейность

Алгоритм применения предлагаемого критерия следующий.

1. Строится линейная аппроксимация данного ряда по МНК.
2. Вычисляется кумулятивная сумма и определяются S_{\max} и S_{\min} для интегральной кривой.
3. Вычисляется диапазон размаха $R_{\max} = S_{\max} - S_{\min}$.
4. Вычисляется оценка стандартного отклонения σ относительно прямой по формуле:

$$\sigma = \sqrt{\frac{1}{N-1} \sum [y(k) - \bar{y}]^2},$$

где N — длина ряда данных; \bar{y} — среднее значение, определяемое по уравнению прямой.

5. Рассчитывается эмпирическое значение критерия по формуле $D = R_{\max} / \sigma$.

6. Сравнивается полученное значение D с его критическим значением d_{crit} , которое можно найти из табл. 1.

Если $d_{\max} \leq d_{\text{crit}}$, то начальная гипотеза о линейности будет принята, а альтернативная гипотеза о нелинейности отклоняется. В противном случае, если $d_{\max} > d_{\text{crit}}$, экспериментальные данные должны быть описаны нелинейной функцией.

С использованием данных объемом 5–60 значений вычислены значения критической величины d_{crit} для значений вероятностей (уровней значимости) $p = 0,90$, $p = 0,975$, $p = 0,99$ [1]. В табл. 1 приведены значения критической величины d_{crit} для интегральной кривой при $\sigma = 0,01$.

ПРИМЕР ИСПОЛЬЗОВАНИЯ УПРОЩЕННОГО ТЕСТА

Для проверки предлагаемого теста на нелинейность были построены несколько моделей. Один из простейших примеров тестирования основан на использовании функции $y = x^{1/2}$ в выбранном интервале зависимой переменной $x \in [0,4; 1,6]$. В каждом сечении функция $y = x^{1/2}$ определена в точках: $x_1 = 0,4$, $x_2 = 1,0$ и $x_3 = 1,6$.

Таблица 1. Критические значения d_{crit} , полученные из стандартной таблицы Фишера для интегральной кривой при $\sigma = 0,01$

№	Число измерений N	d среднее	Критические величины d_{crit} при различных вероятностях p		
			$p = 0,90$	$p = 0,975$	$p = 0,99$
1	18	0,767	0,633	0,644	0,668
2	21	0,903	0,780	0,789	0,807
3	24	1,032	0,829	0,842	0,869
4	27	1,165	0,951	0,964	0,990
5	30	1,298	1,033	1,480	1,077
6	33	1,437	1,176	1,189	1,216
7	36	1,555	1,232	1,248	1,279
8	39	1,692	1,373	1,389	1,418
9	42	1,825	1,425	1,443	1,477
10	45	1,957	1,523	1,542	1,578
11	48	2,090	1,686	1,703	1,735
12	51	2,218	1,741	1,760	1,796
13	54	2,356	1,843	1,830	1,900

Зависимая переменная смоделирована в семи случайных точках (как это показано на рис. 4) с помощью случайной составляющей, которая сгенерирована по нормальному закону распределения:

$$y(k) = x^{1/2}(k) + \varepsilon(k), \quad (4)$$

где $\{\varepsilon(k)\} \sim N(0, \sigma_\varepsilon^2)$.

Таким образом, общее количество экспериментальных точек составило 21. На этом рисунке имеется точка пересечения кривых, показывающая факт наличия нелинейности в тестовых данных.

Для значения вероятности $p = 0,90$ вычисленное значение D составило 0,903, $d_{\text{crit}} = 0,780$. Следовательно, случайный процесс должен быть описан нелинейной функцией, т.к. 0,903 больше критического 0,780. Аналогично, для значения $p = 0,975$ полученное эмпирическое значение боль-

ше критического 0,789, а это подтверждает, что процесс должен быть описан нелинейной функцией.

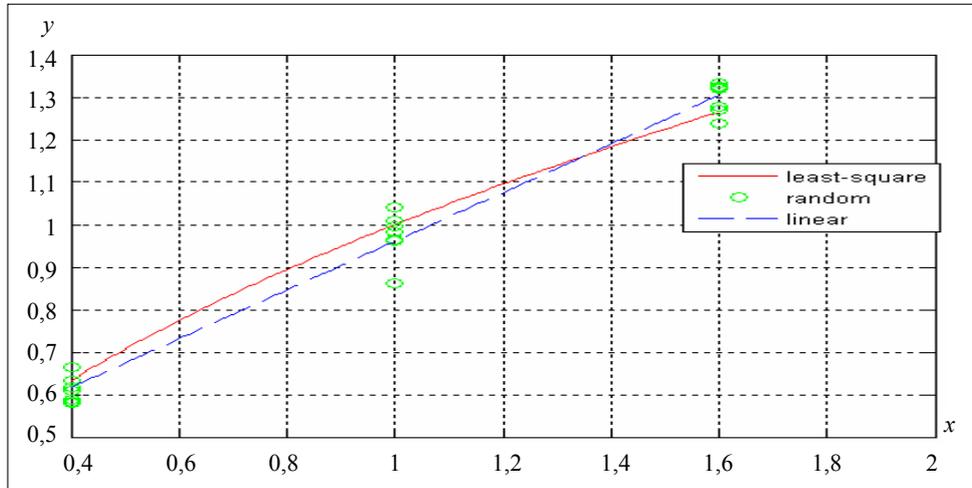


Рис. 4. График функции $y = x^{1/2}$ в выбранном интервале независимой переменной $x \in [0,4; 1,6]$

Такая схема была принята для получения возможности сравнения двух испытуемых тестов на нелинейность: классического теста Фишера и предлагаемого эмпирического критерия. Всего получено 100 реализаций случайного процесса $\varepsilon(k)$ для следующих оценок стандартных отклонений: $\sigma_{\varepsilon 1} = 0,02$, $\sigma_{\varepsilon 2} = 0,05$ и $\sigma_{\varepsilon 3} = 0,2$. В результате выполнения численных экспериментов получены эмпирические линейные связи между критерием Фишера и предлагаемым тестом.

На рис. 5 представлены кластеры экспериментально полученных точек по описанным двум методам для установления связи между двумя тестами при определении критических значений предлагаемого эмпирического критерия (зависимость оценок D от F , при возрастании F возрастает и D).

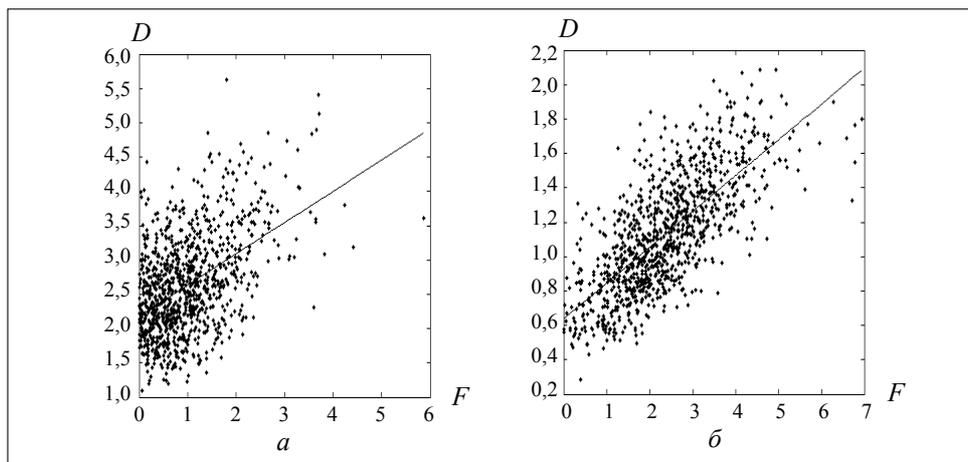


Рис. 5. Кластеры точек, показывающие наличие линейной связи между двумя различными тестами

Для того чтобы определить количественную оценку значения критерия D в зависимости от F было произведено 1000 измерений входного синтезированного сигнала для следующих значений стандартного отклонения: $a — \sigma = 0,02$; $b — \sigma = 0,05$; $c — \sigma = 0,2$.

Таким образом, удалось установить тесную связь между эмпирическим (относительным) отклонением D и квадратным корнем из статистики Фишера \sqrt{F} . Результаты моделирования представлены в табл. 2.

Таблица 2. Результаты моделирования по определению взаимосвязи эмпирического критерия и теста Фишера

№	Оценки стандартного отклонения σ_z для случайного процесса $\varepsilon(k)$ и значения коэффициента вариации V для различных стандартных отклонений		
	$\sigma_\varepsilon = 0,02$	$\sigma_\varepsilon = 0,05$	$\sigma_\varepsilon = 0,2$
1	$\sigma_\varepsilon = 0,02$	$\sigma_\varepsilon = 0,05$	$\sigma_\varepsilon = 0,2$
2	$V = 0,99$	$V = 0,893$	$V = 0,726$
3	$D = 3,485 + 0,6987\sqrt{F}$	$D = 3,442 + 0,713\sqrt{F}$	$D = 3,501 + 0,693\sqrt{F}$

Коэффициент вариации V , как мера относительного разброса случайной величины, определяется по формуле:

$$V = \frac{\sigma}{\bar{x}}, \quad (3)$$

где \bar{x} — выборочное среднее. Это соотношение принимает максимальное значение (коэффициент вариации составляет около 0,99), когда стандартное отклонение для шумовой составляющей σ минимальное. Сила связи минимальная в случае, когда σ принимает максимальное значение (коэффициент вариации составляет около 0,726). Другими словами, два теста тесно связаны друг с другом при этом классический подход является теоретически более обоснованным и для него имеется таблица критических значений, которая была использована авторами.

Однако предлагаемый тест проще классического и может быть применен в условиях неполноты экспериментальных данных. Так как упрощенный тест связан с классическим тестом соотношением $D = 3,485 + 0,6987\sqrt{F}$, то можно легко определить критические значения для d_{crit} . В табл. 3 приведены критические значения d_{crit} , полученные из стандартной таблицы Фишера для интегральной кривой при $\sigma = 0,02$.

Например, для вычисленного значения $d = 869$, $d_{\text{crit}} = 0,698$ — гипотеза о линейности отвергается с вероятностью 0,95. Соответственно, для значения $d = 869$, $d_{\text{crit}} = 0,719$ (гипотеза о линейности отвергается на уровне вероятности 0,975) и $d = 869$, $d_{\text{crit}} = 0,808$ (гипотеза о линейности отвергается на уровне вероятности 0,999).

Таблица 3. Критические значения d_{crit} , полученные из стандартной таблицы Фишера для интегральной кривой при $\sigma = 0,02$

№	Число измерений N	Критические величины d_{crit} при различных вероятностях p		
		$p = 0,90$	$p = 0,95$	$p = 0,99$
1	18	0,556	0,581	0,704
2	21	0,676	0,698	0,808
3	24	0,701	0,729	0,866
4	27	0,786	0,816	0,955
5	30	0,835	0,867	1,019
6	33	0,951	0,983	1,132
7	36	0,960	0,996	1,164
8	39	1,109	1,143	1,299
9	42	1,083	1,123	1,307
10	45	1,156	1,198	1,389
11	48	1,269	1,310	1,495
12	51	1,273	1,318	1,522

Полученные соотношения, связывающие вычисленные средние значения эмпирического критерия d_{mean} с критерием Фишера F для значений $\sigma = 0,02$, приведены в табл. 4.

Таблица 4. Соотношения, связывающие d_{mean} с критерием Фишера для значений стандартного отклонения $\sigma = 0,02$

№	d_{mean}	$d = a_0 + a_1 * \text{sqrt}(F)$
1	0,78	$d = 0,445 + 0,064 * \text{sqrt}(F)$
2	0,92	$d = 0,572 + 0,060 * \text{sqrt}(F)$
3	1,05	$d = 0,565 + 0,079 * \text{sqrt}(F)$
4	1,18	$d = 0,644 + 0,083 * \text{sqrt}(F)$
5	1,32	$d = 0,677 + 0,093 * \text{sqrt}(F)$
6	1,46	$d = 0,792 + 0,093 * \text{sqrt}(F)$
7	1,57	$d = 0,445 + 0,064 * \text{sqrt}(F)$
8	1,71	$d = 0,572 + 0,060 * \text{sqrt}(F)$
9	1,84	$d = 0,565 + 0,079 * \text{sqrt}(F)$
10	1,98	$d = 0,644 + 0,083 * \text{sqrt}(F)$
11	2,11	$d = 0,677 + 0,093 * \text{sqrt}(F)$
12	2,24	$d = 1,044 + 0,136 * \text{sqrt}(F)$
13	2,38	$d = 1,172 + 0,133 * \text{sqrt}(F)$

Поскольку оценки стандартного отклонения экспериментальных данных существенно влияют на конечный результат, его нужно определять по возможности точнее с использованием различных видов аппроксимации.

Критические значения d_{crit} , полученные из таблицы Фишера для заданных квантилей доверительной вероятности и заданного числа степеней свободы, служат для сравнения с реальными данными при идентификации нелинейностей. Для того чтобы вычислить отклонение, можно использовать приближение в виде параболы второго порядка или полиномиальной регрессии.

Сравнительный анализ метода, предложенного в [1] и в данной работе, показал, что корреляция между этими двумя подходами выявления нелинейности очень близка и составляет $r = 0,7 - 0,95$. Повторное моделирование позволяет утверждать, что эти два метода дают возможность получать очень близкие оценки, несмотря на различные алгоритмы их вычислений.

ВЫВОДЫ

В работе предложен упрощенный тест для анализа нелинейности, основанный на достаточно простых вычислениях и не требующий нескольких выборок данных для одного и того же процесса. В результате компьютерного моделирования и сравнительного анализа двух тестов установлено, что существует тесная связь между классическим тестом, предложенным Фишером, и упрощенным тестом, предложенным в работе. Это тесная взаимосвязь позволяет использовать упрощенный тест в условиях неполноты экспериментальных данных.

ЛИТЕРАТУРА

1. Закс Б. Статистическое оценивание. — М.: Статистика, 1976. — 598 с.
2. Кузьмин В.Н. Новый статистический метод идентификации нелинейности на основе эмпирических данных // Конференция по компьютерному анализу данных и моделированию. — Минск, 1998. — Т. 1. — С. 159–164.
3. Бидюк П.И., Баклан И.В., Гасанов А.С. Системный подход к анализу адекватности нелинейных моделей временных рядов // Кибернетика и системный анализ. — 2003. — № 3. — С. 147–158.
4. Бабак О.В., Гасанов А.С., Мельников С.В. Применение метода генеральной обобщенной переменной при идентификации нелинейных объектов // Вісн. КНУ ім. Т.Г. Шевченка. Сер. фіз.-мат. наук. — 2002. — № 5. — С. 9–12.
5. Rao M.J.N. Filtering and Control of Macroeconomic Systems. — Amsterdam: North Holland, 1987. — 360 p.

Поступила 04.04.2012