УДК 519.22./25

Т.А. Рыжкина, З.П. Старовойтова
Дальневосточный государственный технический рыбохозяйственный университет,
690087, г. Владивосток, ул. Луговая, 52б

МОДЕЛИРОВАНИЕ УСТОЙЧИВЫХ ФИЛЬТРОВ ДЛЯ СТОХАСТИЧЕСКИХ ПРОЦЕССОВ

На примерах генерации определенным образом случайных последовательностей описываются достаточно простые устойчивые модели с возможностью воспроизводства данных в условиях одного испытания. Авторегрессия как линейный фильтр последовательности случайных чисел позволяет на выходе получать свойства устойчивости или стационарности. Указанное свойство ставится в зависимость от длины входящего потока.

Ключевые слова: *последовательность случайных чисел, имитационная модель, стационарность, авторегрессия, автокорреляция, характеристический многочлен, теорема Руше.*

T.A. Ryzhkina, Z.P. Starovoytova

MODELING SUSTAINABLE FILTERS FOR STOCHASTIC PROCESSES

In the examples generate a certain way random sequences described is quite simple sustainable model with the possibility of reproduction of data in terms of a single trial. The autoregressive linear filter sequence of random numbers allows you to get the output properties of stability or stationarity. This property is dependent on the length of the incoming stream.

Key words: *the sequence of random numbers, simulation model, stationarity, autoregressive, the autocorrelation of the characteristic polynomial, theorem Ruše.*

Введение

Объектами исследования служат стохастические процессы и их реализации в виде временных рядов.

Набор случайных переменных $\boldsymbol{X(t)}$, где t – время (в общем случае – подмножество или множество действительных чисел) называется стохастическим процессом.

В работе не делается различия между стохастическим процессом $X(t)$ и порожденным с его помощью временным рядом, если процесс обладает свойством стационарности, хотя бы в слабом смысле.

Стационарным процессом в слабом смысле называют стохастический процесс, для которого среднее и дисперсия независимо от рассматриваемого периода времени имеют постоянное значение. Автоковариация зависит только от длины лага $\tau = t_1 - t_2$ между рассматриваемыми переменными t_1, t_2. С физической точки зрения такой процесс представляет колебания относительно некоторого постоянного значения, а стохастическая зависимость между двумя сечениями регулируется только расстоянием между ними.

Интерес в данной работе представляет получение моделей, лежащих в основе процедуры прогнозирования. Предполагается, прежде всего, получение информации о развитии процесса в настоящий момент, предсказание показателей процесса в течение кратковременного периода с учетом преемственности данных и степени их влияния на процесс. Применяется метод имитационного моделирования. Он позволяет строить модели, описывающие процессы так, как они проходили в действительности. При этом моделировании используются аналогии с сущностью физических явлений (например, случайные сигналы, «шумы»), описываемых моделируемыми законами распределения. Изучаемый реальный объект заменяется моделью с достаточной степенью точности его описания. Экспериментирование с моделью называют имитацией [1].

Работа является естественным продолжением и уточнением начатого исследования в [2]. Применяются методы композиции, генерации, регрессии, корреляции [1, 3, 4].

1. Цель исследования. Основные понятия и определения.

Целью является разработка достаточно устойчивой статистики. Под устойчивостью результатов имитации понимается степень их нечувствительности к изменению входных условий. Оценка устойчивости может быть выполнена разными способами. Чаще всего контролируют дисперсию результатов модели в зависимости от интервала моделирования. Если увеличения дисперсии не происходит, результаты применения модели считаются устойчивыми. На базе созданного случайного поля с помощью авторегрессионных моделей (преобразователей) ставится задача аппроксимации изучаемого явления с меньшим числом параметров и с более простыми свойствами.

Случайное поле $\boldsymbol{X(t)}$ генерируется методом композиции с помощью встроенных в ППП Excel статистических функций и надстройки «Пакет анализа».

В данном исследовании аппроксимируем теоретические характеристики процесса $\boldsymbol{X(t)}$, т.е. математическое ожидание $\boldsymbol{m_x(t)}$, дисперсию $\boldsymbol{D_x(t)}$, среднее квадратическое отклонение $\boldsymbol{\sigma_x(t)} = \sqrt{D_x(t)}$, состоятельными выборочными средними величинами $\bar{x}, \bar{d}$ **,** $\boldsymbol{\sigma}$ временного ряда $\boldsymbol{x_t}$ как реализации $\boldsymbol{X(t)}$ соответственно. Состоятельность выборочной оценки – это гарантированная сходимость по вероятности к теоретической оценке с увеличением массива наблюдений.

Для упрощения записи формул практикуется переход к центрированным стандартизованным случайным функциям $\boldsymbol{X(t)}$ с характеристиками $\boldsymbol{m_x(t)} = 0$, $\boldsymbol{\sigma_x(t)} = 1$. Реализации $\boldsymbol{x_t}$ случайных функций стандартизуются и центрируются аналогично.

Автокорреляционной функцией стандартизованного центрированного процесса $\boldsymbol{X(t)}$ называется выражение

$$\rho_x(t_1, t_2) = M[\overset{o}{X}(t_1) \cdot \overset{o}{X}(t_2)], \tag{1}$$

где $\overset{o}{X}(t_1)$, $\overset{o}{X}(t_2)$ – сечения центрированного стандартизованного $\boldsymbol{X(t)}$. Эта функция характеризует тесноту линейной связи между разными сечениями стандартизованного центрированного процесса.

Предполагается, что $\boldsymbol{X(t)}$ является **эргодическим**. Оценивание его характеристик возможно по одной достаточно длинной реализации $\boldsymbol{x_t}$ этого процесса. **Достаточным условием для эргодичности стационарного** $\boldsymbol{X(t)}$ по математическому ожиданию и автоковариации является сходимость автоковариации $K_x(\tau)$ к нулю при $\tau \to \infty$, [1–6].

Автоковариация и дисперсия связаны формулой

$$K_x(0) = D_x(t). \tag{2}$$

Очевидно, что дисперсия эргодического стационарного процесса также стремится к нулю.

Под **«гауссовским белым шумом»** в данной работе будем понимать нормальный стохастический процесс как совместное нормальное распределение двух его разных сечений.

Можно доказать, что стационарный «белый шум», представляющий собой систему одинаково распределенных нормальных, независимых случайных величин, не является нормальным. Действительно, по закону совместного распределения двух разных сечений этого процесса следует, что сечения, имеющие вероятность, отличную от нуля, совпадают. Это противоречит определению нормального процесса. Вероятность равенства двух случайных величин по нормальному закону равна нулю. Однако стационарность в слабом смысле у этого процесса имеется [1–9].

2. Генерация одномерного (многомерного) случайного поля

В качестве одномерного случайного поля (одномерной случайной величины) рассматриваем последовательность случайных чисел (ПСЧ), для которых моделируется нормальное распределение N(a, σ) методом композиции. Под многомерным случайным полем понимается система случайных величин (СВ), совместно характеризующих какое-либо случайное явление. Проще всего моделировать систему с независимыми компонентами. Дифференциальный закон распределения такой системы имеет вид

$$f(x_1, x_2, \dots x_n) = \mathop{\Pi}_{i=1}^{n} f_i(x_i)\,,$$

т.е. каждую из компонент системы можно моделировать независимо от других в соответствии с ее плотностью вероятностей (например, нормальной).

На основании предельных теорем теории вероятностей доказывается возможность представления одной СВ в виде комбинации достаточно большого числа СВ, имеющих более простые и легко реализуемые законы распределения. Известно, что распределение СВ $\boldsymbol{z}$ в виде преобразования

$$z = \frac{1}{\sqrt{\dfrac{k}{12}}}\left(\sum_{i=1}^{k} R_i \quad - \frac{k}{2}\right), \tag{3}$$

где R_i – равномерно распределенные на интервале [0;1] ПСЧ, при неограниченном возрастании k приближается к нормальному распределению N(0;1).

Действительно, равномерное распределение R имеет параметры m = 1/2, d = 1/12. Введем СВ $\boldsymbol{x} = \sum\limits_{i=1}^{k} R_i$, $m_x = \sum\limits_{i=1}^{k} \frac{1}{2} = \frac{k}{2}$, $\sigma_x = \sqrt{D_x} = \sqrt{\sum\limits_{i=1}^{k} \frac{1}{12}} = \sqrt{\frac{k}{12}}$. СВ $\boldsymbol{z}$, таким образом, является центрированной величиной, нормально распределенной по закону N(0;1), $\boldsymbol{z} = (\boldsymbol{x} - m_x) / \sigma_x$. В практике вычислений достаточно взять $k \geq 12$.

Кроме равномерно распределенных ПСЧ, по центральной предельной теореме [8], для моделирования нормального распределения возможно использование сумм достаточно большого числа СВ с одинаковым (неизвестно каким) распределением и конечными числовыми характеристиками. Параметры моделируемого распределения приближаются к величинам $\boldsymbol{a = km_x}$, $\boldsymbol{\sigma = k\sigma_x}$ при $k \to \infty$.

Технически генерация случайных (точнее, псевдослучайных) чисел может быть выполнена в «Пакете анализа» приложений Excel. Выбирается инструмент «Генерация случайных чисел», указывается одномерный (n-мерный с независимыми компонентами) случайный вектор заданной длины. В раскрывающемся списке выбирается тип распределения (например, нормальное распределение с заданными параметрами), на основе которого генерируются случайные числа. При таком построении ПСЧ имеет место псевдослучайность [1]. Это понимается в том смысле, что числа так или иначе являются детерминированными, дробная часть которых выражается конечной десятичной дробью. Следует также иметь в виду, что совокупность одинаково распределенных нормальных СВ как система распределена, вообще говоря, не по нормальному закону. В окне «Генерация случайных чисел» имеется поле «Случайное рассеивание». В этом поле указывается некоторое целое число для инициализации генератора случайных чисел. Следует это число фиксировать. Это позволяет генерировать каждый раз одну и ту же ПСЧ.

3. Общее описание авторегрессионной модели одномерного процесса

В этой модели, как известно, текущее значение временного ряда (случайного процесса или СВ) $\boldsymbol{x_t}$ выражается через конечную линейную совокупность предыдущих значений $\boldsymbol{x_{t-p}}$ $(p \geq 1)$ и остатки $\boldsymbol{\varepsilon_t}$, реализующие «белый шум». Для упрощения записи последующих операций центрируем $\boldsymbol{x_t}$, т.е. перейдем к отклонению $\boldsymbol{x_t}$ от среднего значения. Обозначим это отклонение символом $\tilde{x}_t = x_t - \bar{x}$. Формула авторегрессии принимает следующий вид:

$$\tilde{x}_t = \boldsymbol{\alpha} + \sum_{i=1}^{p} b_i \ \tilde{x}_{t-i} + \boldsymbol{\varepsilon}_t, \quad t = p+1,...,n, \quad b_i = const. \tag{4}$$

Иногда постоянный входящий поток $\boldsymbol{\alpha}$ не включают в формулу для получения прямой зависимости текущего значения от предыдущих значений, т.е.

$$\tilde{x}_t = \sum_{i=1}^{p} b_i \ \tilde{x}_{t-i} + \boldsymbol{\varepsilon}_t, \quad t = p+1,...,n, \quad b_i = const. \tag{5}$$

Коэффициенты формулы (5) могут быть получены методом наименьших квадратов (МНК) в простой форме при условии, что остатки $\boldsymbol{\varepsilon_t}$ образуют «белый шум».

Матричное уравнение для вычисления коэффициентов имеет вид

$$\boldsymbol{W^T W B = W^T X}, \tag{6}$$

где матрица $\boldsymbol{W} = \begin{pmatrix} x_1 & x_2 & \dots & x_p \\ x_2 & x_3 & \dots & x_{p+1} \\ \dots & \dots & \dots & \dots \\ x_{n-1} & x_{n-2} & \dots & x_{n-p} \end{pmatrix}$, $\boldsymbol{W^T}$ – транспонированная матрица по отношению к $\boldsymbol{W}$, матрица $\boldsymbol{B} = (\boldsymbol{b_1, b_2, \dots, bp})^{\boldsymbol{T}}$, $\boldsymbol{X = (x_{p+1}, \dots, x_{n-p})^T}$.

Другой вариант расчета коэффициентов регрессии с нулевым свободным членом возможен с помощью встроенной в ППП Excel функцией «Линейн».

Введем линейный оператор сдвига назад $\boldsymbol{L}$ (оператор лага) как в [1, 3], действующий так: $Lb_p\widetilde{x}_{t-p} = b_p L^p$, $p \geq o.$ Авторегрессионный оператор порядка $\boldsymbol{p}$ как функция оператора лага примет вид $\varphi(L) = 1 - \sum_{i=1}^{p} b_i\, L^i$. Формулы (4), (5) преобразуются соответственно:

$$\left(1 - \alpha - \sum_{i=1}^{p} b_i L^i\right) x_t = \varepsilon_t, \qquad \left(1 - \sum_{i=1}^{p} b_i L^i\right) x_t = \varepsilon_t. \tag{7}$$

Компактно формулы (4), (5), (7) записываются так:

$$\varphi(L)\widetilde{x}_t = \varepsilon_t. \tag{8}$$

Инструментом построения устойчивой (стационарной в слабом смысле) формулы (7) служит автокорреляционная функция (1). Величины $\widetilde{x}_{t-k}$ для разных $\boldsymbol{k}$ являются разными центрированными сечениями $\boldsymbol{x_t}$**.** Для получения формулы (1) последовательно выполняем следующие действия. Умножаем (5) на $\widetilde{x}_{t-k}$, $k > 0$; после умножения берем математическое ожидание; учитываем, что математическое ожидание совмещения $\widetilde{x}_{t-k} \cdot \varepsilon_t$ некоррелированных величин равно нулю; полученное разностное уравнение делим на $\boldsymbol{\sigma^2}$ (некоррелированными, для уточнения, являются величины $\boldsymbol{\varepsilon_{t-k}}$, $\boldsymbol{\varepsilon_t}$).

Таким образом, автокорреляционная функция выражается в виде линейной комбинации коэффициентов автокорреляции, с такими же коэффициентами, что и в формуле (8):

$$\rho_k = \sum_{i=1}^{p} b_i\; \rho_{k-i}\,,\;\; k = 1,...,n/4\,\mathbf{.} \tag{9}$$

Ограничиваем число $\boldsymbol{k}$ величиной $\boldsymbol{n}$ **/ 4**, чтобы не ослаблять поле корреляции [3, 6].

Используя автокорреляционный оператор, (9) можно записать так:

$$\varphi(L)\,\rho_k = 0. \tag{10}$$

Равенство нулю в соотношении (10) возможно только при условии, что норма $\|\varphi\| = 0$.

Норма – это конечная выпуклая неотрицательная числовая функция с некоторыми дополнительными свойствами, действующая в линейном, например, комплексном, пространстве. Для любого ограниченного оператора справедлива теорема [10]

$$\|\varphi\| = \sup_{|z|<1} \|\varphi(z)\| = \sup_{z \neq 0} \frac{|\varphi(z)|}{|z\,|}. \tag{11}$$

Это означает, что равенство нормы оператора нулю приводит к алгебраическому уравнению степени $\boldsymbol{p}$ вида

$$1 - \sum_{i=1}^{p} b_i\, z^{\mathrm{i}} = 0\,. \tag{12}$$

Это уравнение называют характеристическим [3, 6] по отношению к (8). Устойчивость формулы (8) зависит от корней характеристического уравнения (12). Обозначим его корни (действительные или комплексно-сопряженные) символами $\boldsymbol{(\zeta_i)^{-1}}$, $\boldsymbol{i = 1,\ldots,p}$ с учетом их кратности. Оператор $\boldsymbol{\varphi(L)}$ через корни характеристического уравнения можно записать следующим образом:

$$\varphi(L) = \prod_{i=1}^{p}(1-\zeta_i L), \quad \|L\| < 1. \tag{13}$$

Если все корни характеристического уравнения различны, то значения автокорреляционной функции представляются через $\boldsymbol{\zeta_i}$ в виде следующей комбинации действительных или комплекснозначных экспонент

$$\rho_k = \sum_{i=1}^{p} A_i\, \zeta_i^{\,k}\ ,\ k = 1,\ldots,n/4. \tag{14}$$

С другой стороны, для различных корней характеристического уравнения текущие значения процесса представляются в виде

$$\tilde{x}_t = \varphi^{-1}(L)\,\varepsilon_t = \sum_{i=1}^{p}\frac{K_i}{(1-\zeta_i L)}\cdot\varepsilon_t\quad . \tag{15}$$

Исходя из автокорреляционной функции вида (14) и принимая во внимание формулу (15), получаем сходящуюся последовательность значений $\boldsymbol{\rho_k}$ и самого процесса с ростом $\boldsymbol{k}$ при корнях уравнения (12) вне единичного круга, т.е.

$$|1/\zeta_i| > 1\,, i = 1,\ldots,p \quad , \quad |\zeta_i| < 1. \tag{16}$$

Например, для формул (14) первого и второго порядков соответственно имеем

$$\rho_1 = b_1\, d_1 = A\zeta\ \ ,\ \ \rho_2 = b_1\rho_1 + b_2\, d_2 = A_1\,\zeta_1^{\,2} + A_2\,\zeta_2^{\,2}.$$

В частности, текущие значения процесса в случае p-кратного корня **1 / ζ**, |ζ| < 1 приобретают вид

$$\tilde{x}_t = \sum_{i=1}^{p}(-1)^{i+1}C_p^{\,i}\zeta^i\tilde{x}_{t-i} + \beta\varepsilon\ \ , t = p+1,\ldots,n. \qquad \beta = \frac{\sigma_x}{\sigma_\varepsilon}\sqrt{\frac{(1-\zeta^2)^{2p-1}}{\sum_{i=0}^{p-1}(C^i{}_{p-1}\zeta^i)^2}}\,. \tag{17}$$

Формула (17) содержит некоторый коэффициент усиления устойчивости модели [11]. В общем случае автокорреляционная функция стационарного процесса состоит из смеси затухающих экспонент и затухающих синусоидальных волн [3].

4. Нули аналитической функции в замкнутой области

Однозначная дифференцируемая комплекснозначная функция $\boldsymbol{f(z)}$ в любой точке некоторой области называется аналитической функцией в этой области.

Известно, что число $\boldsymbol{\mu}$ нулей аналитической функции в замкнутой области, ограниченной контуром $\boldsymbol{C}$, определяется по принципу аргумента [12], а именно:

$$1 / 2\pi\ \Delta_C\ Arg\, f(z) = \mu,$$

где $Arg\, f(z)$ – аргумент функции; $f(z) = | f(z) |\ e^{iArgf(z)}$; Δ_C – приращение $Arg\, f(z)$ при обходе контура C.

Это утверждение может быть применено, в частности, к нахождению нулей характеристического многочлена $f(z) = 1 - \sum b_i z^i$, $i = 1,\ldots, p$, в круговой области. В другой формулировке стоит задача найти корни характеристического уравнения $f(z) = 0$ в круге $D = \{| z | <\}$.

Известно также, что максимум и минимум аналитической функции в замкнутой области достигается на границе области. Аналитическую функцию достаточно задать на границе замкнутой области. Значения функции в любой точке внутри области могут быть представлены через граничные значения функции по теории Коши.

Решение вопроса о количестве корней уравнения $1-\sum_{i=1}^{p} b_i z^i = 0$ может быть достигнуто на основании теоремы Руше [12].

Теорема Руше. Пусть функции $\boldsymbol{f(z)}$ и $\boldsymbol{\varphi(z)}$, аналитические в замкнутой области $\overline{D}$, ограниченной контуром $\boldsymbol{C}$, во всех точках этого контура удовлетворяют неравенству $| f(z) | > | \varphi(z) |$. Тогда их сумма $\boldsymbol{F(z) = f(z) + \varphi(z)}$ и функция $\boldsymbol{f(z)}$ имеют в области D одинаковое число нулей (с учетом их кратности).

Пример. Определить число корней уравнения $z^4 - 5z + 9 = 0$ внутри круга D=$\{|z| < 1\}$.

Решение. Пусть $\boldsymbol{f(z)} = 9$, $\boldsymbol{\varphi(z)} = z^4 - 5z$. На окружности $\boldsymbol{C} = \{|z|=1\}$ для введенных функций получаем следующие ограничения: $| \boldsymbol{f(z)} | = 9$; $| \boldsymbol{\varphi(z)} | \le | z^4 | + 5 | z | = 6$. Имеем на $\boldsymbol{C}$ неравенство $| \boldsymbol{f(z)} | > | \boldsymbol{\varphi(z)} |$. Функция $\boldsymbol{f(z)}$ не имеет корней внутри D=$\{|z| < 1\}$. Таким образом, по теореме Руше не имеет нулей в $D = \{|z| < 1\}$ функция $\boldsymbol{F(z)} = z^4 - 5z + 9$. Итак, корней данного уравнения внутри круга $D = \{|z| < 1\}$ нет.

5. Примеры генерирования ПСЧ и построения авторегрессионных моделей

В основу моделирования положена формула интегрального закона нормального распределения $N(a, \sigma)$ [13]

$$P(X < x) = 0{,}5 + \Phi(x - a / \sigma), \qquad (18)$$

где P – вероятность; $\Phi(x) = \int_0^x e^{-t^2/2} dt$ – функция Лапласа.

С помощью этой формулы по заданной вероятности P попадания нормально распределенной СВ X в интервал $(-\infty, x)$ возвращается точка $\boldsymbol{x}$. Для этого используется встроенная в ППП Excel функция «Норм. обр». ПСЧ для вероятностей генерируется с помощью инструмента «Генерация случайных чисел» в «Пакете анализа». Поле «Случайное рассеивание» в **первом примере** построения ПСЧ не было активировано.

Произвольно сформированная ПСЧ равномерного распределения $\boldsymbol{R}(\mathbf{0,1})$ была упорядочена, табл. 1, для придания процессу $\boldsymbol{x_t}$ некоторого прикладного смысла. Каждое число из ПСЧ понимается как вероятность $R_t = P(X < x_t)$, t – текущий момент времени.

Таблица 1

Случайное равномерное распределение

Table 1

Casual even distribution

t	R_t	t	R_t	t	R_t
1	0,988525	11	0,607685	21	0,450789
2	0,895962	12	0,607166	22	0,445692
3	0,858943	13	0,601764	23	0,37788
4	0,82284	14	0,585009	24	0,364452
5	0,808741	15	0,571184	25	0,352123
6	0,802606	16	0,563585	26	0,350291
7	0,783319	17	0,531663	27	0,303995
8	0,746605	18	0,519883	28	0,30195
9	0,710501	19	0,513535	29	0,193304
10	0,663045	20	0,479873	30	0,174108

Полученная последовательность показывает, что реализации случайного процесса $\boldsymbol{x_t}$ с ростом $\boldsymbol{t}$ становятся менее возможными. Попытаемся текущие значения $\boldsymbol{x_t}$, табл. 2, представить в виде взвешенной суммы предыдущих значений. Оператор сдвига назад $\boldsymbol{L^p}$ применим для $\boldsymbol{p}$ = 4; 3; 2; 1. Выбор числа наблюдений $\boldsymbol{n}$=30 сделан с учетом возможных значений лага $\boldsymbol{p}$ и свойств автокорреляционной функции.

На основании табл. 1 функция «Норм. обр» формирует ПСЧ нормального распределения с заданными параметрами $\boldsymbol{a = 1, \sigma = 0{,}1}$, табл. 2.

Относительно элементов табл. 2 вычислим их отклонения от среднего значения

$\tilde{x}_t = x_t - a = x_t - 1.$

Таблица 2

Случайное нормальное распределение

Table 2

Casual normal distribution

t	x_t	t	x_t	t	x_t
1	1,2274	11	1,0273	21	0,9876
2	1,1259	12	1,0272	22	0,9863
3	1,1076	13	1,0258	23	0,9689
4	1,0926	14	1,0215	24	0,9653
5	1,0873	15	1,0179	25	0,9620
6	1,0851	16	1,0160	26	0,9615
7	1,0783	17	1,0079	27	0,9487
8	1,0664	18	1,0050	28	0,9481
9	1,0555	19	1,0034	29	0,9134
10	1,0421	20	0,9950	30	0,9062

Таблица 3

Отклонения от среднего значения

Table 3

Deviations from a mean value

t	$\tilde{x}_t$	t	$\tilde{x}_t$	t	$\tilde{x}_t$
1	0,2274	11	0,0273	21	-0,0124
2	0,1259	12	0,0272	22	-0,0137
3	0,1076	13	0,0258	23	-0,0311
4	0,0926	14	0,0215	24	-0,0347
5	0,0873	15	0,0179	25	-0,0380
6	0,0851	16	0,0160	26	-0,0384
7	0,0783	17	0,0079	27	-0,0513
8	0,0664	18	0,0050	28	-0,0519
9	0,0555	19	0,0034	29	-0,0866
10	0,0421	20	-0,0050	30	-0,0938

Составим формулу (4) с лагом $\boldsymbol{p} = 4$ для отклонений в табл. 3, пользуясь инструментом «Регрессия» в надстройке «Пакет анализа». Значения $\tilde{x}_t$, $t = 5,\ldots$ выражаются через $\tilde{x}_t$, $t = 1$, 2, 3, 4:

$$\tilde{x}_t = -0,00690 + 0,91045\ \tilde{x}_{t-1} + 0,38030\ \tilde{x}_{t-2} - 0,31180\ \tilde{x}_{t-3} + 0,04339\ \tilde{x}_{t-4} + \varepsilon_t. \quad (19)$$

Стандартная ошибка этой формулы (с учетом случайности) $\boldsymbol{\sigma_{ост}} = 0,00786$. В поле корреляции формула (19) описывает явление по дисперсии результата на величину 0,9782.

Для исходного $\boldsymbol{x_t}$ имеем:

$$x_t = 0,01545 + 0,91045\, x_{t-1} + 0,38030\, x_{t-2} - 0,31180\ x_{t-3} + 0,04339\ x_{t-4} + \varepsilon_t. \quad (20)$$

В формуле (20) сохраняется постоянный входящий поток. Получим далее формулу вида (5), без постоянной величины, с лагом $\boldsymbol{p} = 4$ для отклонений в табл. 3. Коэффициенты этой формулы вычислим с помощью матричного уравнения (6) или функции «Линейн»:

$$\begin{pmatrix} 0,060403 & 0,06099 & 0,065327 & 0,077722 \\ 0.06099 & 0,064476 & 0,070039 & 0,085348 \\ 0.065327 & 0,070039 & 0,077632 & 0,096008 \\ 0,077722 & 0,085348 & 0,096008 & 0,126723 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} = \begin{pmatrix} 0,059149 \\ 0,058534 \\ 0,061468 \\ 0,072258 \end{pmatrix}.$$

Имеем: $\boldsymbol{b_1}$ =1,20494**;** $\boldsymbol{b_2}$ =0,54774; $\boldsymbol{b_3}$ =-0,81424; $\boldsymbol{b_4}$ =0,07916;

$$\tilde{x}_t = 1,20494\,\tilde{x}_{t-1} + 0,54774\,\tilde{x}_{t-2} - 0,81424\,\tilde{x}_{t-3} + 0,07916\,\tilde{x}_{t-4} + \varepsilon_t. \quad (21)$$

Стандартная ошибка формулы (21) равна $\boldsymbol{\sigma_{ост}}$ =0,00858082.

По аналогичной схеме составим формулы вида (5) без свободного члена, с лагом $\boldsymbol{p}$ = 3;2;1:

$$\tilde{x}_t = 1{,}30664\,\tilde{x}_{t-1} - 0{,}12935\,\tilde{x}_{t-2} - 0{,}01727\,\tilde{x}_{t-3} + \varepsilon_t, \qquad \boldsymbol{\sigma_{ocm}} = 0{,}009199. \tag{22}$$

$$\tilde{x}_t = 1{,}20446\,\tilde{x}_{t-1} - 0{,}22414\,\tilde{x}_{t-2} + \varepsilon_t, \qquad \boldsymbol{\sigma_{ocm}} = 0{,}00968. \tag{23}$$

$$\tilde{x}_t = 0{,}797481\,\tilde{x}_{t-1} + \varepsilon_t, \qquad \boldsymbol{\sigma_{ocm}} = 0{,}0167685. \tag{24}$$

Второй пример организуем следующим образом. Инициализируем генератор случайных чисел в поле «Случайное рассеивание» «Пакета анализа» цифрой 2. Формируем ПСЧ равномерного распределения $\boldsymbol{R}(\mathbf{0{,}1})$ и возвращаем по этой последовательности нормальное распределение $\boldsymbol{x_t}$ встроенной в ППП Excel функцией «Норм. обр.» с заданными параметрами $\boldsymbol{a} = \mathbf{1}$, $\boldsymbol{\sigma} = \mathbf{0{,}1}$, табл. 4, 5, 6. Решение задачи моделирования составим из двух частей.

Первая часть решения **второго примера** выполняется при условии сортировки вероятностей от большего к меньшему значению в табл. 4. В соответствии с упорядочиванием значений вероятностей происходит перестановка СВ $\boldsymbol{x_t}$ в табл. 5, 6.

Таблица 4

Случайное равномерное распределение (генератор 2)

Table 4

Casual even distribution (generator 2)

t	$\boldsymbol{R_t}$	t	$\boldsymbol{R_t}$	t	$\boldsymbol{R_t}$
1	0,001373	11	0,169591	21	0,336894
2	0,891629	12	0,860286	22	0,918149
3	0,738487	13	0,38612	23	0,577319
4	0,543077	14	0,018494	24	0,218726
5	0,899808	15	0,981262	25	0,454573
6	0,599689	16	0,544115	26	0,715354
7	0,445265	17	0,909848	27	0,181951
8	0,806635	18	0,926756	28	0,073977
9	0,326701	19	0,203192	29	0,906735
10	0,558977	20	0,883847	30	0,229255

Таблица 5

Случайное нормальное распределение (генератор 2)

Table 5

Casual normal distribution (generator 2)

t	$\boldsymbol{x_t}$	t	$\boldsymbol{x_t}$	t	$\boldsymbol{x_t}$
1	0,700525	11	0,904422	21	0,957904
2	1,123524	12	1,108161	22	1,139273
3	1,063869	13	0,971055	23	1,019504
4	1,010819	14	0,791411	24	0,92235
5	1,128046	15	1,208053	25	0,988588
6	1,025254	16	1,01108	26	1,056909
7	0,986237	17	1,133982	27	0,909205
8	1,086556	18	1,145205	28	0,85532
9	0,955096	19	0,916973	29	1,132092
10	1,014838	20	1,119444	30	0,92587

Таблица 6

Отклонения от среднего значения (генератор 2)

Table 6

Deviations from a mean value (generator 2)

t	$\widetilde{x}_t$	t	$\widetilde{x}_t$	t	$\widetilde{x}_t$
1	-0,29948	11	-0,09558	21	-0,0421
2	0,123524	12	0,108161	22	0,139273
3	0,063869	13	-0,02894	23	0,019504
4	0,010819	14	-0,20859	24	-0,07765
5	0,128046	15	0,208053	25	-0,01141
6	0,025254	16	0,01108	26	0,056909
7	-0,01376	17	0,133982	27	-0,0908
8	0,086556	18	0,145205	28	-0,14468
9	-0,0449	19	-0,08303	29	0,132092
10	0,014838	20	0,119444	30	-0,07413

Вторая часть решения задачи моделирования в **примере 2** упорядочивание вероятностей не предполагает.

В **первой части** получены модели:

$$x_t = -0{,}00547 + 1{,}666065\ x_{t-1} - 0{,}21468\ x_{t-2} - \\ - 0{,}48356\ x_{t-3} + 0{,}09019\ x_{t-4} + \varepsilon_t,\ \boldsymbol{\sigma_{ост}} = 0{,}016156, \tag{25}$$

$$x_t = 1{,}74641\ x_{t-1} - 0{,}17214\ x_{t-2} - 0{,}513\ x_{t-3} - 0{,}01605\ x_{t-4} + \varepsilon,\ \boldsymbol{\sigma_{ост}} = 0{,}01607, \tag{26}$$

$$x_t = 1{,}83844\ x_{t-1} - 0{,}44985\ x_{t-2} - 0{,}334383\ x_{t-3} + \varepsilon_t,\ \boldsymbol{\sigma_{ост}} = 0{,}015837, \tag{27}$$

$$x_t = 1{,}829673\ x_{t-1} - 0{,}75444808\ x_{t-2} + \varepsilon_t,\ \boldsymbol{\sigma_{ост}} = 0{,}017918, \tag{28}$$

$$x_t = 1{,}039321\ x_{t-1} + \varepsilon_t,\ \boldsymbol{\sigma_{ост}} = 0{,}0283364. \tag{29}$$

Во **второй части** имеем:

$$x_t = 0{,}018001 - 0{,}30482\ x_{t-1} + 0{,}004034\ x_{t-2} + \\ 0{,}237016\ x_{t-3} - 0{,}1443\ x_{t-4} + \varepsilon_t,\ \boldsymbol{\sigma_{ост}} = 0{,}010194, \tag{30}$$

$$x_t = -0{,}27168\ x_{t-1} + 0{,}055031\ x_{t-2} + 0{,}308748\ x_{t-3} - 0{,}2054\ x_{t-4} + \varepsilon_t\ ,\ \boldsymbol{\sigma_{ост}} = 0{,}10013, \tag{31}$$

$$x_t = -0{,}26784\ x_{t-1} + 0{,}087262\ x_{t-2} + 0{,}246727\ x_{t-3} + \varepsilon_t,\ \boldsymbol{\sigma_{ост}} = 0{,}099023, \tag{32}$$

$$x_t = -0{,}27872\ x_{t-1} - 0{,}07241\ x_{t-2} + \varepsilon_t,\ \boldsymbol{\sigma_{ост}} = 0{,}1005, \tag{33}$$

$$x_t = -0{,}29088\ x_{t-1} + \varepsilon_t,\ \boldsymbol{\sigma_{ост}} = 0{,}09745. \tag{34}$$

6. Проверка полученных моделей на устойчивость

Для формулы (19) учтем, что M(-0,0069 $\tilde{x}_{t-k}$)=0. Характеристическое уравнение этой модели принимает вид

$$z^4 - 7{,}18603 z^3 + 0{,}87650 z^2 + 20{,}98345 z - 23{,}04742 = 0. \tag{35}$$

Применим теорему Руше к этому уравнению. Левую его часть представим суммой

$\boldsymbol{F(z)} = \boldsymbol{f(z)} + \boldsymbol{\varphi(z)}$, где $\boldsymbol{f(z)} = -7{,}18603 z^3 + 20{,}98345 z - 23{,}04742$, $\boldsymbol{\varphi(z)} = z^4 - 0{,}87650\, z^2$.

Модули функций $\boldsymbol{f(z)}$, $\boldsymbol{\varphi(z)}$ удовлетворяют следующим неравенствам:

$9{,}25 \le |\, \boldsymbol{f(z)}\, | \le 51{,}2169, 0{,}1235 \le |\, \boldsymbol{\varphi(z)}\, | \le 1{,}8765$ на окружности $\boldsymbol{C} = \{|z| = 1\}$. Число корней уравнения (35) в $D = \{|z| < 1\}$ определяется числом нулей функции $\boldsymbol{f}(z)$ в круге D. Из трех нулей $\boldsymbol{f(z)}$ один, действительный, попадает внутрь D, два других, комплексно-сопряженных, оказываются вне окружности $\boldsymbol{C_1} = \{|z| = 2\}$. Таким образом, функция $\boldsymbol{F(z)}$ также имеет один нуль в $D = \{|z| < 1\}$. Другими словами, среди корней уравнения (35) один корень попадает внутрь круга $D = \{|z| < 1\}$. Формула (19) характеризует нестационарный процесс, не годится для прогноза значений случайного процесса вне поля корреляции. Последовательность коэффициентов автокорреляции расходится.

Характеристическое уравнение формулы (21) имеет вид

$$z^4 - 10{,}2860 z^3 + 6{,}9194 z^2 + 15{,}2216 z - 12{,}6326 = 0. \tag{36}$$

По теореме Руше левую часть (36) рассматриваем как сумму функций:

$\boldsymbol{f}(z) = z^4 + 6{,}9194 z^2 + 15{,}2216 z$, $\quad \boldsymbol{\varphi(z)} = -10{,}2860\, z^3 - 12{,}6326$.

Модули этих функций удовлетворяют неравенствам на окружности $\boldsymbol{C} = \{|z|=1\}$:

$7{,}3022 \le |\, \boldsymbol{f(z)}| \le 23{,}141$, $2{,}3466 \le |\boldsymbol{\varphi(z)}| \le 22{,}9186$, $|\, \boldsymbol{f(z)}| > |\, \boldsymbol{\varphi(z)}|$.

Это означает, что число корней уравнения (36) в $D = \{|z| < 1\}$совпадает с числом нулей $\boldsymbol{f}(z)$ в этой области. Функция $\boldsymbol{f}(z)$ имеет, по крайней мере, один нуль $z = 0$ в $D = \{|z| < 1\}$. Процесс (21) неустойчивый.

Для модели (22) с лагом 3 характеристическое уравнение принимает форму

$$z^3 + 7{,}4899 z^2 - 75{,}6600 z + 57{,}9039 = 0. \tag{37}$$

Левую часть (37) представим в виде суммы функций:

$\boldsymbol{f}(z) = -75{,}66 z$, $\quad \boldsymbol{\varphi(z)} = z^3 + 7{,}4899 z^2 + 57{,}9039$.

Модули этих функций на окружности $\boldsymbol{C} = \{|z| = 1\}$:

$|\, \boldsymbol{f(z)}\, | = 75$, $\quad |\, \boldsymbol{\varphi(z)}\, | \le 66{,}3938$, $\quad |\, \boldsymbol{f(z)}\, | > |\, \boldsymbol{\varphi(z)}|$.

Очевидно, что уравнение (37) имеет в $D = \{|z| < 1\}$ хотя бы один корень, так как хотя бы один корень в D имеет функция $\boldsymbol{f(z)}$**.** Формула (22) не обладает свойством устойчивости.

Модель (23) с лагом 2 имеет характеристическое уравнение вида

$$z^2 - 5{,}3737 z + 4{,}4615 = 0. \tag{38}$$

Корни этого уравнения $z_1 = 4{,}347$, $z_2 = 1{,}0262$. Таким образом, на основании (16) модель (23) устойчивая, на грани единичного корня [3, 6].

Для формулы (24) на основании ее характеристического уравнения

$$1 - 0{,}797481z = 0 \qquad (39)$$

и корня z = 1,2539484 справедливо суждение об устойчивости (24).

Таким образом, прогноз будущих значений результата $\boldsymbol{x_t}$ в **примере 1** на несколько шагов вперед при условии, что ПСЧ равномерного распределения была упорядочена, возможен только с лагом $\boldsymbol{p}$ = 1,2.

По теореме Руше, примененной к моделям (25)–(29), устанавливается их неустойчивость. Стохастичность в этих формулах ограничена упорядочиванием реализаций случайного процесса по вероятностям. Аппроксимация изучаемого явления с помощью формул (25)–(29) возможна только на один шаг вперед.

Проверка по той же методике моделей (30)–(34), полученных без ограничений на реализации случайного процесса (**пример 2)**, показывает их устойчивость. Однако случайные ошибки аппроксимации на несколько шагов вперед в этом случае имеют больший порядок, чем в формулах (25)–(29)

Формальный устойчивый линейный фильтр на основе ПСЧ **примера 2** дает формула (17). Характеристическое уравнение для нее имеет кратный корень $z = 1 / \zeta$, $|z| > 1$. Пусть $z = 5$. Тогда (17) с лагом $\boldsymbol{p}$ = 4, $\sigma_x = 0{,}0112$, $\sigma_{ост} = 0{,}0161$ имеет вид

$$x_t = 0{,}8x_{t-1} - 0{,}24x_{t-2} + 0{,}032x_{t-3} - 0{,}0016x_{t-4} + \beta\varepsilon_t\ , \quad \beta = 0{,}379. \qquad (40)$$

Формула (40), фильтруя входящий поток, табл. 6, не показывает тенденции в изменении уровней $\boldsymbol{x_t}$ и представляет колебательный процесс относительно нуля.

При необходимости в полученных стационарных формулах можно вернуться к исходному среднему значению.

Заключение

Итоги проведенных исследований на основании приведенных примеров могут быть представлены следующим образом:

• используются известные механизмы построения случайных полей с точки зрения, определенной авторами;

• достаточно строго математически описан инструмент проверки модели авторегрессии на стационарность в виде характеристического уравнения;

• предложен способ проверки стационарности модели с помощью теоремы Руше;

• установлено, что чем больше элемент случайности, тем проще сформировать стационарную модель лаговой длины $p \geq 4$ с постоянным входным потоком или без него;

• для физических случайных процессов подтверждается оптимальная длина стационарной модели авторегрессии $p \leq 2$.

Список литературы

1. Эконометрика: учебник / под ред. В.Б. Уткина. – М.: Изд-во «Дашков и К°», 2009. – 564 с.

2. Рыжкина, Т.А. Преобразования плоского «белого шума» с определенными свойствами выходных характеристик / Т.А. Рыжкина, З.П. Старовойтова // Науч. тр. Дальрыбвтуза. – Владивосток: Дальрыбвтуз, 2012. – Вып.27. – С.71–82.

3. Лукашин, Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов: учеб. пособие / Ю.П. Лукашин. – М.: Финансы и статистика,2003. – 416 с.

4. Вентцель, Е.С. Теория случайных процессов и ее инженерные приложения: учеб. пособие / Е.С. Вентцель, Л.А. Овчаров. – М.: Высш. шк., 2000. – 383 с.

5. Хрущева, И.В. Основы математической статистики и теории случайных процессов: учеб. пособие / И.В. Хрущева, В.И. Щербаков, Д.С. Леванова. – СПб.: Изд-во «Лань», 2009. – 336 с.

6. Эконометрика: учебник / И.И. Елисеева, С.В. Курышева, Т.В. Костеева и др.; под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2007. – 576 с.

7. Сборник задач по теории вероятностей, математической статистике и теории случайных функций: учеб. пособие / под общ. ред. А.А. Свешникова. – СПб.: Изд-во «Лань», 2008. – 448 с.

8. Вентцель, Е.С. Теория вероятностей: учебник для вузов / Е.С. Вентцель. – М.: Высш. шк., 1999. – 575 с.

9. Гурский, Е.И. Теория вероятностей с элементами математической статистики: учеб. пособие / Е.И. Гурский. – М.: Высш. шк., 1971. – 328 с.

10. Колмогоров, А.Н. Элементы теории функций и функционального анализа: учебник / А.Н. Колмогоров, С.В. Фомин. – М.: Наука,1972. – 496 с.

11. Кловский, Д.Д. Обработка пространственно-временных сигналов (в каналах передачи информации) / Д.Д. Кловский, В.А. Сойфер. – М.: Связь, 1976. – 208 с.

12. Лаврентьев, М.А. Методы теории функций комплексного переменного: учеб. пособие / М.А. Лаврентьев, Б.В. Шабат. – М.: Наука, 1965. – 716 с.

13. Гмурман, В.Е. Теория вероятностей и математическая статистика / В.Е. Гмурман. – М.: Высш. шк., 1972. – 368 с.

Сведения об авторах: Рыжкина Тамара Александровна,
кандидат физико-математических наук, доцент;
Старовойтова Зоя Павловна, доцент.