ISSN 2713-1734

Научная статья
УДК 519.217
EDN YOMUWV
DOI 10.17150/2713-1734.2024.6(2).168-178

В.В. Братищенко
Байкальский государственный университет,
г. Иркутск, Российская Федерация

Скрытая марковская модель курса доллара

Аннотация. В статье предлагается использовать скрытые марковские модели для описания курса доллара. Скрытая цепь Маркова описывает свойства исходного ряда: возрастание, убывание или совпадение соседних членов ряда. Для курса доллара выявлено зависимость свойств от дня недели, поэтому цепь состояний описывается семью стохастическими матрицами. Проверка по критерию хи-квадрат подтвердила адекватность предложенной модели смены состояний. Каждое состояние марковской цепи для курса доллара приводит положительному, отрицательному или нулевому приращению. Значимой стохастической связи членами ряда приращений не выявлено. Математическое ожидание и дисперсия ряда приращений существенно изменяются для разных моментов времени, однако, для некоторых периодов времени приращения можно считать одинаково распределенными. Для таких периодов и при условии независимости приращений ряда получены формулы для вычисления прогноза и доверительного интервала. Предложенная модель может применяться для исследования других рядов наблюдений.

Ключевые слова. Скрытая марковская модель, нестационарный ряд, неоднородная цепь Маркова, прогноз ряда, доверительный интервал прогноза.

Информация о статье. Дата поступления: 13 марта 2024 г.; дата принятия к публикации: 23 мая 2024 г.; дата онлайн-размещения: 19 июня 2024 г.

Original article

V.V. Bratischenko
Baikal State University,
Irkutsk, Russian Federation

Hidden Markov Model of the Dollar Rate

Abstract. The article proposes to use hidden Markov models to describe the dollar exchange rate. A hidden Markov chain describes the properties of the original series: increasing, decreasing, or coincidence of neighboring members of the series. For the dollar exchange rate, the dependence of properties on the day of the week was revealed, so the chain of states is described by seven stochastic matrices. The chi-square test confirmed the adequacy of the proposed model of state changes. Each state of the Markov chain for the dollar exchange rate results in a positive negative or zero increment. No significant stochastic relationship between the terms of a number of increments was revealed. The mathematical expectation and dispersion of increments change significantly for different points in time, however, for some periods of time the increments can be considered equally distributed. For such periods and provided that the series increments are independent, formulas for calculating the forecast and confidence interval are obtained. The proposed model can be used to study other series of observations.

Keywords. Hidden Markov model, non-stationary series, inhomogeneous Markov chain, series forecast, forecast confidence interval.

Article info. Received 4 March, 2024; Accepted 23 May, 2024; Available online 19 June, 2024.

Стохастические ряды финансовых рынков отражают закономерности, влияющие на доходность отдельных финансовых активов. Знание таких закономерностей актуально для управления финансовыми потоками. Построение моделей финансовых рядов осложняется их нестационарностью, наличием пиков, «тяжелыми хвостами» распределений и другими факторами. С первых исследований финансовых рядов [1] была выявлена близость таких рядов процессам случайного блуждания. Поэтому классические модели авторегрессии и скользящего среднего (ARMA), а также их интегрированная версия (ARIMA) [2] не позволяют получить более точный прогноз по сравнению со случайным блужданием. Многочисленные работы используют самые разные методы описания финансовых рядов. В работе [3] основой модели ряда является дифференциальное уравнение, определяющее физическое явление турбулентного движения несжимаемой вязкой жидкости. Фрактальные модели [4] временной ряд описывают фракталом — фигурой, обладающей свойством автоподобия. Применяют для прогнозирования финансовых рядов искусственные нейронные сети [5]. Нейронные сети для отдельного ряда позволяют получать краткосрочные прогнозы, однако, закономерности, порождающие ряд, остаются скрытыми. В работах [6; 7] для прогнозирования финансовых рядов используются выделенные шаблоны поведения. В работе [8] выделение закономерностей предлагается на основе марковских моделей. Следует отметить, что и модель случайного блуждания, и процессы авторегрессии, и модели сглаживания по сути являются марковскими процессами. Преимущество марковских моделей заключается в явном описании вероятностных закономерностей. В данной работе для моделирования финансовых рядов предлагается использовать неоднородную марковскую цепь. Возможность применения такой модели демонстрируется на примере курса доллара.

Оценка вероятностного распределения случайного ряда часто является проблематичной из-за большого количества предположений о свойствах, вероятностных характеристиках, видах законов распределения и их параметрах. Даже оценка одномерного закона распределения, если доступно наблюдение одной реализации ряда, требует выполнения свойств стационарности и эргодичности. Наблюдаемые в экономике ряды, в основном, не являются стационарными. Такие ряды обычно представляют в виде композиции детерминированной и случайной составляющих. Предположение о существенном ослаблении зависимости между элементами ряда при увеличении расстояния между ними позволяет применять марковские модели. Один из вариантов такого применения заключается в использовании марковского процесса, который влияет на имеющийся ряд наблюдений. Такие модели получили название скрытых марковских моделей [9; 10]. В работе предлагается скрытая марковская модель для описания ряда курса доллара.

В теории вероятностей и математической статистики находят применение различные характеристики случайных процессов. Например, в известном критерии серий проверки ряда на независимость и случайность [11] исследуются серии значений больших и меньших медианы. Использование серий — векторов, составленных из L подряд идущих значений ряда — позволяет исследовать «локальные» свойства серий и на этом основании изучать свойства случайного ряда в целом.

Простым свойством соседних значения является факт возрастания или убывания. Для непрерывных распределений совпадение соседних членов ряда имеет нулевую вероятность. Однако, в распространенных рядах цен, курсов валют, котировках ценных бумаг такие случаи встречаются достаточно часто. Поэтому для соседних значений ряда ..., x_{i-1}, x_i, ... предлагается три состояния:

$$s_i = +1, \text{ если } x_{i-1} < x_i;\ s_i = 0, \text{ если } x_{i-1} = x_i;\ s_i = -1, \text{ если } x_{i-1} > x_i.$$

В качестве объекта моделирования выбран курс доллара с 01.01.1999 по 01.11.2021. Распределение состояний этого ряда существенно зависит от дня недели. В таблице приведены частоты состояний в зависимости от дня недели. Такой перекос имеет очевидное объяснение: по выходным курс не меняется, поэтому после выходного курс совпадает с предыдущим, а отклонения от этого правила, связаны с переносами выходных дней из-за праздников.

Частоты состояний по дням недели

День недели	Состояния		
	–1	0	+1
1	0,002	0,996	0,003
2	0,439	0,138	0,423
3	0,477	0,074	0,449
4	0,473	0,061	0,466
5	0,487	0,067	0,446
6	0,470	0,086	0,445
7	0,012	0,983	0,005

Частоты переходов из состояния в состояние также зависят от дней недели. Стохастические матрицы переходов для рабочих дней являются подобными. Матрица переходов из состояний среды в состояния четверга

$$P(3) = \begin{Vmatrix} 0,493 & 0,040 & 0,467 \\ 0,318 & 0,386 & 0,295 \\ 0,477 & 0,030 & 0,493 \end{Vmatrix}$$

определяет примерно равные частоты смены состояния возрастания на состояния убывания. Таким образом, отсутствуют явные тенденции к сохранению предыдущего состояния. Матрицы переходов для выходных дней также примерно одинаковы

$$P(6) = \begin{Vmatrix} 0,007 & 0,991 & 0,002 \\ 0,010 & 0,990 & 0,000 \\ 0,017 & 0,974 & 0,009 \end{Vmatrix}$$

и практически достоверно определяют переход к совпадению соседних состояний.

Если выполняется марковское свойство для последовательности состояний, то вероятность трех последовательных состояний можно вычислить, используя стохастические матрицы переходов

$$P\{(s_i = k)(s_{i+1} = m)(s_{i+2} = l)\} = P\{(s_i = k)\}P(d(i))_{km}P(d(i+1))_{ml},$$

где $d(i)$ — день недели i-го элемента ряда. Для проверки выполнения марковского свойства сравнивались эмпирические частоты троек подряд идущих значений и соответствующие теоретические вероятности. Различие частот и вероятностей оказалось незначительным, и совпадение частот с вероятностями прошло проверку по критерию хи-квадрат с доверительной вероятностью 0,9368. Таким образом, ряд состояний достаточно точно моделируется неоднородной марковской цепью.

Стохастические матрицы всех дней недели являются эргодическими [12]. Соответствующие им цепи Маркова не содержат невозвратных состояний. Каждая такая цепь характеризуется стационарным распределением. Оценки таких распределений приведены в таблице 1. Для эргодической цепи Маркова можно использовать следующее разложение стохастической матрицы P

$$P = X\Lambda X^{-1},$$

где $\Lambda = diag(\lambda_1, ..., \lambda_n)$ — диагональная матрица собственных чисел, X — матрица собственных векторов стохастической матрицы.

Эргодическая стохастическая матрица за t шагов сходится к единственному пределу при $t \to \infty$. Все строки предельной матрицы одинаковы и совпадают со стационарным распределением состояний. Сходимость стохастической матрицы $P^t = X\Lambda^t X^{-1}$ к предельному стационарному значению определяется соответствующими степенями собственных чисел. Одно из чисел стохастической матрицы всегда равно 1. Ему соответствует левый собственный вектор стационарного распределения.

Для матриц выходных дней все собственные числа, кроме единичного, близки к 0. Для матриц рабочих дней вторые по величине собственные числа образуют последовательность 0,1667; 0,3655; 0,3515; 0,3810; 0,6729. Собственные числа матрицы переходов за неделю не будут превышать произведение этих чисел, равное 0,005489. Это означает, что за неделю марковская цепь практически приходит к стационарному распределению вероятностей состояний, соответствующих дням недели (табл.).

Курс доллара является нестационарным. Для уменьшения влияния абсолютный величин рассмотрим ряд удельных разностей

$$y_i = \frac{x_i - x_{i-1}}{x_{i-1}}.$$

Автокорреляционная функция ряда y_t (рис. 1) косвенно свидетельствует о независимости элементов ряда.

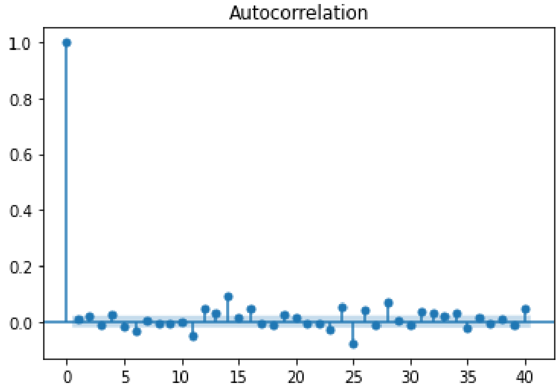

Рис. 1. Оценка автокорреляционной функции ряда y_t удельных разностей

Исследования не выявили значимых отклонений распределений удельных разностей по дням недели. Кроме этого, распределения положительных и отрицательных удельных разностей оказались практически совпадающими. На рис. 2 представлены соответствующие графики. Для наглядности гистограммы заменены на графики. Данные распределения ближе всего к гамма-распределению.

Ряд удельных разностей остается нестационарным — оценки его математических ожиданий и дисперсий существенно изменяются в некоторые периоды времени. На рис. 3 представлены усредненные оценки математического ожидания и среднеквадратичного отклонения (СКО) по декадам каждого месяца. Нестационарность ряда объясняется различными экономическими и политическими событиями, существенно влияющими на курсы валют. Описание таких закономерностей выходит за рамки данной работы.

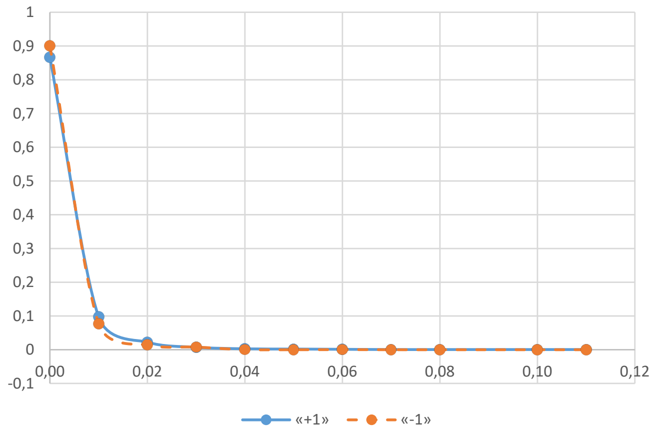

Рис. 2. Эмпирические плотности удельных разностей положительных («+1») и отрицательных («–1») приращений

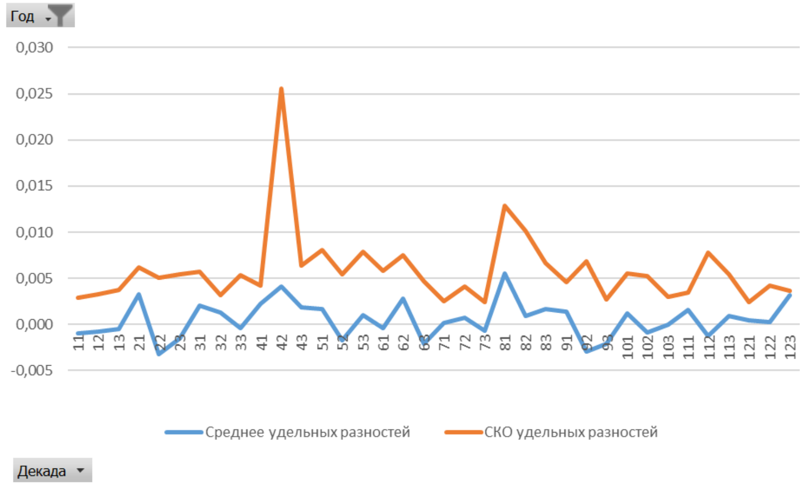

Рис. 3. Среднее и среднеквадратическое отклонение удельных разностей по декадам 2018 г. (ось ординат образована номерами месяца и декады)

Для практического применения выявленных закономерностей будем предполагать, что на протяжении некоторого периода времени удельные приращения независимы и одинаково распределены. В этом случае прогноз — случайную величину $Y_{t+\tau}$ — порождающую элемент ряда $y_{t+\tau}$ можно определить следующим образом

$$Y_{t+\tau} = y_t + \sum_{s_1, \ldots, s_\tau = -1}^{+1} P\big(d(t)\big)_{s_0 s_1} \ldots P\big(d(t+\tau-1)\big)_{s_{\tau-1} s_\tau} \times$$
$$\times \left(s_1 \varepsilon_{t+1} + \ldots + s_\tau \varepsilon_{t+\tau}\right),$$

где ε_t — независимые случайные величины, порождающие удельные разности, s_i-состояние в момент времени $t + i$. При условии независимости ε_t можно определить математические ожидания и дисперсии прогноза

$$M[Y_{t+\tau}] = y_t + \sum_{s_1,\ldots,s_\tau=-1}^{+1} P\big(d(t)\big)_{s_0 s_1} \ldots P\big(d(t+\tau-1)\big)_{s_{\tau-1} s_\tau} \times$$
$$\times\ (s_1 M[\varepsilon_{t+1}] + \ldots + s_\tau M[\varepsilon_{t+\tau}]),$$

$$D[Y_{t+\tau}] = \sum_{s_1,\ldots,s_\tau=-1}^{+1} P\big(d(t)\big)_{s_0 s_1} \ldots P\big(d(t+\tau-1)\big)_{s_{\tau-1} s_\tau} \times$$
$$\times\ (s_1{}^2 D[\varepsilon_{t+1}] + \ldots + s_\tau{}^2 D[\varepsilon_{t+\tau}]),$$

Графики изменения математического ожидания (рис. 4) и дисперсии (рис. 5) прогноза демонстрируют ожидаемую зависимость характеристик от дней недели.

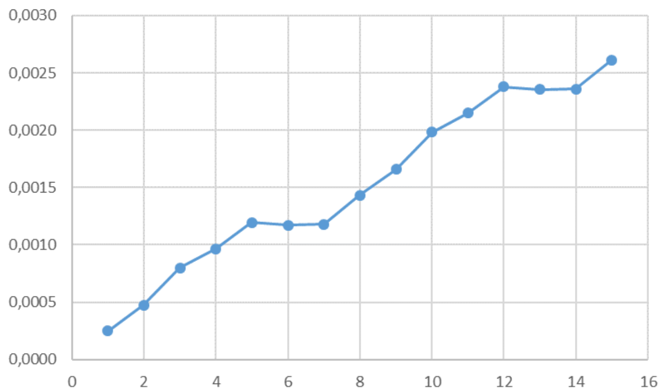

Рис. 4. Рост математического ожидания прогноза ряда y_t с понедельника

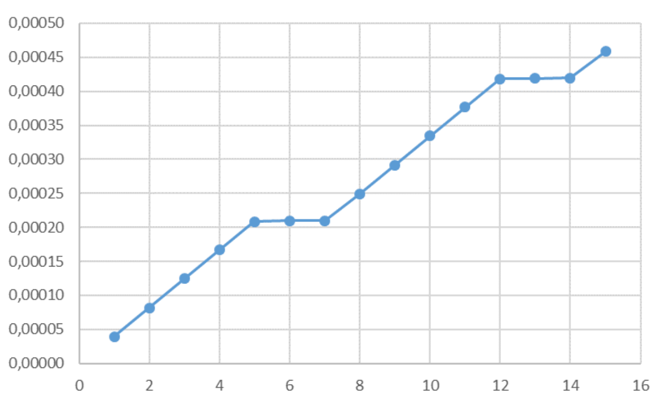

Рис. 5. Дисперсия прогноза ряда y_t с понедельника

На рис. 6 приведен фрагмент ряда y_t прогноз и границы прогноза, которые вычислялись как смещение прогноза на среднеквадратичное отклонение в положительную и отрицательную стороны. Доверительная вероятность такого интервала, оцененная по эмпирическому распределению отклонений ε, приблизительно равна 0,9.

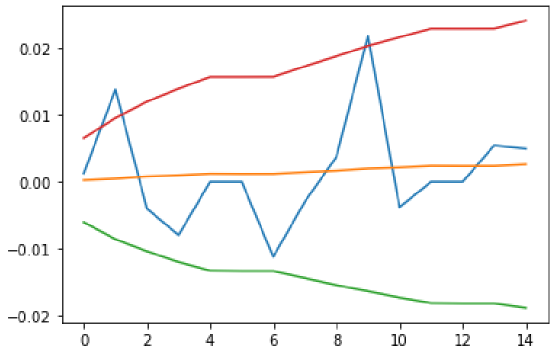

Рис. 6. Фрагмент наблюдений ряда y_t в сопоставлении с прогнозом, верхней и нижней границами прогноза

По соотношению рядов

$$x_t = (y_t + 1)\, x_{t-1}$$

можно определить прогноз ряда x_t как случайную величину

$$X_{t+\tau} = \sum_{s_1, \ldots, s_\tau = -1}^{+1} P\big(d(t)\big)_{s_0 s_1} \ldots P\big(d(t+\tau-1)\big)_{s_{\tau-1} s_\tau} \times$$

$$\times\, x_t (s_1 \varepsilon_{t+1} + 1) \ldots (s_\tau \varepsilon_{t+\tau} + 1),$$

а также, вычислить математические ожидания и дисперсии прогноза ряда X

$$M[X_{t+\tau}] = x_t \sum_{s_1, \ldots, s_\tau = -1}^{+1} P\big(d(t)\big)_{s_0 s_1} \ldots P\big(d(t+\tau-1)\big)_{s_{\tau-1} s_\tau} \times$$

$$\times\, (s_1 M[\varepsilon_{t+1}] + 1) \ldots (s_\tau M[\varepsilon_{t+\tau}] + 1),$$

$$M\big[X_{t+\tau}^{\;2}\big] = x_t^2 \sum_{s_1, \ldots, s_\tau = -1}^{+1} P\big(d(t)\big)_{s_0 s_1} \ldots P\big(d(t+\tau-1)\big)_{s_{\tau-1} s_\tau} \times$$

$$\times \big(s_1^{\,2} M\big[\varepsilon_{t+1}^{\;2}\big] + 2 s_1 M[\varepsilon_{t+1}] + 1\big) \ldots \big(s_\tau^{\,2} M\big[\varepsilon_{t+\tau}^{\;2}\big] + 2 s_\tau M[\varepsilon_{t+\tau}] + 1\big),$$

$$D[X_{t+\tau}] = M\big[X_{t+\tau}^{\;2}\big] - M[X_{t+\tau}]^2.$$

На Рис. 7 приведен график фрагмента ряда x_t, прогноза значений и границ прогноза с доверительной вероятностью 0,9.

Применение предложенной модели требует тщательной проверки условий независимости и одинаковости распределения приращений. Модель можно дополнить более сложными описаниями рядов приращений. Например, можно отдельно рассматривать ряды положительных и отрицательных приращений, выявлять зависимость законов распределения приращений от времени или других наблюдаемых факторов.

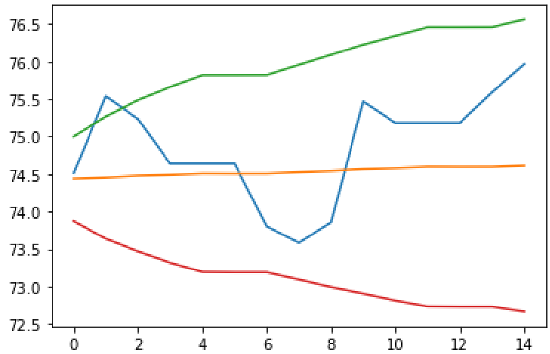

Рис. 7. Фрагмент наблюдений ряда x_t в сопоставлении с прогнозом, верхней и нижней границами прогноза

Проведенное исследование курса доллара выявило зависимость приращений курса от дня недели. Для описания таких зависимостей можно использовать скрытые марковские модели. Подобные модели можно применять для описания других свойств, связанных с особенностями функционирования экономических объектов и влияющих значения ряда.

Список использованной литературы

1. Samuelson P. Proof That Properly Anticipated Prices Fluctuate Randomly / P. Samuelson // Industrial Management Review. — 1965. — no. 6. — P. 41–49.

2. Бокс Дж. Анализ временных рядов: прогноз и управление / Дж. Бокс, Г. Дженкинс; под ред. В.Ф. Писаренко. — Москва : Мир, 1974. — 197 с.

3. Мусин А.Р. Применение математической модели турбулентного движение жидкости для прогнозирования значений обменных курсов / А.Р. Мусин. — EDN ZAOMHV // Азимут научных исследований: экономика и управление. — 2017. — № 2 (19). — С. 200–203.

4. Александровская Ю.П. Использование фрактальных методов для анализа финансовых рядов / Ю.П. Александровская. — EDN SXYHWF // Вестник Казанского технологического университета. — 2014. — № 18. — С. 256–261.

5. Казаковцева М.В. Прогнозирование котировок финансовых инструментов с помощью нейронных сетей / М.В. Казаковцева, Е.В. Конакова. — DOI 10.30914/2411-9687-2023-9-4-433-442. — EDN BEKPPK // Вестник Марийского государственного университета. Серия: Сельскохозяйственные науки. Экономические науки. — 2023. — № 4. — С. 433–442.

6. Балонишников А.М. Прогнозирование временных рядов методами Фармера-Сидоровича и Бокса-Дженкинса / А.М. Балонишников, В.А. Балонишникова, А.В. Копыльцов. — EDN OFUWHL // Известия Российского государственного педагогического университета им. А. И. Герцена. — 2011. — №141. — С. 7–16.

7. Amigó J.M. Detecting Determinism in Time Series with Ordinal Patterns: a Comparative Study / J.M. Amigó, S. Zambrano, M.A.F. Sanjuán. — DOI 10.1142/s0218127410027453 // International Journal of Bifurcation and Chaos. — 2010. — no. 20 (09). — P. 2915–2924.

8. Богомолов Р.О. Биномиальная байесовская модель бескупонной облигации / Р.О. Богомолов, В.М. Хаметов. — EDN WCDINV // Прикладная эконометрика. — 2016. — № 2 (42). — С. 100–120.

9. Моттль В.В. Скрытые марковские модели в структурном анализе сигналов / В.В. Моттль, И.Б. Мучник. — Москва : Физматлит, 1999. — 351 с.

10. Букреев В.Г. Выявление закономерностей во временных рядах в задачах распознавания состояний динамических объектов : монография / В.Г. Букреев, С.И. Колесникова, А.Е. Янковская. — EDN QJXPHT. — Томск : Изд-во Томского политехнического университета, 2010. — 254 с.

11. Айвазян С.А. Прикладная статистика: Основы моделирования и первичная обработка данных / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин. — Москва : Финансы и статистика, 1983. — 471 с.

12. Кемени Дж. Конечные цепи Маркова / Дж. Кемени, Дж. Снелл. —Москва : Наука, 1970. — 272 с.

References

1. Samuelson P. Proof That Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review*, 1965, no. 6, pp. 41–49.

2. Box D., George E.P. *Time Series Analysis; Forecasting and Control*. San Francisco, Holden-Day, 1970. 584 p. (Russ. ed.: Boks D., Dzhenkins G.; Pisarenko V.F. (ed.). *Time Series Analysis; Forecasting and Control*. Moscow, Mir Publ., 1974. 197 p.).

3. Musin A.R. Applying Mathematical Model of Fluid Turbulent Dynamics to Forecasting Exchange Rates Time Series. *Azimut nauchnykh issledovanii: ekonomika i upravlenie = Azimuth of Scientific Research: Economics and Administration,* 2017, no. 2, pp. 200–203. (In Russian). EDN: ZAOMHV.

4. Aleksandrovskaya Yu.P. Using fractal methods to analyze financial series. *Vestnik Kazanskogo tekhnologicheskogo universiteta = Herald of Kazan Technological University,* 2014, no. 18, pp. 256–261. (In Russian). EDN: SXYHWF.

5. Kazakovtseva M.V., Konakova E.V. Forecasting Financial Instruments Quotations Using Neural Networks. *Vestnik Mariiskogo gosudarstvennogo universiteta. Seriya: Sel'skokhozyaistvennye nauki. Ehkonomicheskie nauki = Bulletin of the Mari State University. Series: Agricultural Sciences. Economic Sciences,* 2023, no. 4, pp. 433–442. (In Russian). EDN: BEKPPK.

6. Balonishnikov A.M., Balonishnikova V.A., Kopyl'tsov A.V. Prediction of time series by Box-Jenkins and Farmer-Sidorovich methods. *Izvestiya Rossiiskogo gosudarstvennogo pedagogicheskogo universiteta im. A.I. Gertsena = Herzen University Journal of Humanities & Science,* 2011, no. 141, pp. 7–16. (In Russian). EDN: OFUWHL.

7. Amigó J.M., Zambrano S., Sanjuán M.A.F. Detecting Determinism in Time Series with Ordinal Patterns: a Comparative Study. *International Journal of Bifurcation and Chaos*, 2010, no. 20, pp. 2915–2924. DOI: 10.1142/s0218127410027453.

8. Bogomolov R.O., Khametov V.M. Binomial Bayesian zero-coupon bond model. *Prikladnaya ekonometrika = Applied econometrics,* 2016, no. 2, pp. 100–120. (In Russian). EDN: WCDINV.

9. Mottl' V.V., Muchnik I.B. *Hidden Markov models in structural signal analysis.* Moscow, Fizmatlit Publ., 1999. 351 p.

10. Bukreev V.G., Kolesnikova S.I., Yankovskaya A.E. *Identification of patterns in time series in problems of recognition of states of dynamic objects.* Tomsk, Tomskii politekhnicheskii universitet Publ., 2010. 254 p. EDN: QJXPHT.

11. Aivazyan S.A., Enyukov I.S., Meshalkin L.D. *Applied statistics: Fundamentals of modeling and primary data processing*. Moscow, Finansy i statistika Publ., 1983. 471 p.

12. Kemeny J.G., Snell D. *Finite Markov Chains*. Princeton, N.J., Van Nostrand, 1960. 232 p. (Russ. ed.: Kemeni D., Snell D. *Finite Markov Chains.* Moscow, Nauka Publ., 1970. 272 p.).

Информация об авторе

Братищенко Владимир Владимирович — кандидат физико-математических наук, доцент, доцент кафедры математических методов и цифровых технологий, Байкальский государственный университет, Байкальский государственный университет, г. Иркутск, Российская Федерация, e-mail: vvb@bgu.ru.

Information about the Author

Vladimir V. Bratischenko — PhD in Physics and Mathematical Sciences, Associate Professor, Department of Mathematical Methods and Digital Technologies, Baikal State University, Baikal State University, Irkutsk, Russian Federation, e-mail: vvb@bgu.ru.

Для цитирования

Братищенко В.В. Скрытая марковская модель курса доллара / В.В. Братищенко. — DOI 10.17150/2713-1734.2024.6(2).168-178. — EDN YOMUWV // System Analysis & Mathematical Modeling. — 2024. — Т. 6, № 2. — С. 168–178.

For Citation

Bratischenko V.V. Hidden Markov Model of the Dollar Rate. *System Analysis & Mathematical Modeling*, 2024, vol. 6, no. 2, pp. 168–178. (In Russian). EDN: YOMUWV. DOI: 10.17150/2713-1734.2024.6(2).168-178.