МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

МАТЕРИАЛЫ

VI Международной молодежной научной конференции

«МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ИНФОРМАЦИОННЫХ, ТЕХНИЧЕСКИХ И ЭКОНОМИЧЕСКИХ СИСТЕМ»

Томск, 24–26 мая 2018 г.

Под общей редакцией
кандидата технических наук И.С. Шмырина

Томск
Издательский Дом Томского государственного университета
2018

10. *Зенкова Ж.Н., Муравлева М.А.* Расчет стоимости оборотных средств предприятия по цензурированным интервалом данным // Логистические системы в глобальной экономике [Электронный ресурс] : материалы Междунар. науч.-практ. конф. (3–4 марта 2014 г., Красноярск). Вып. 1. – Электрон. сб. – Сиб. гос. аэрокосмич. ун-т. – Красноярск, 2014. С. 103–107.

11. *Зенкова Ж.Н., Макеева О.Б.* Применение методов обработки цензурированных данных при анализе оборачиваемости // Вестник науки Казахского агротехнического университета им. С. Сейфуллина. Астана. – 2014. – № 3 (82). – С. 21–30.

12. *Зенкова Ж.Н., Муравлева М.А.* Асимптотическая несмещенность оценки функции распределения по однократно интервалом цензурированным данным с привлечением информации о симметрии // Новые информационные технологии в исследовании сложных структур: материалы X российской конференции с международным участием. – Томск, Издательский Дом Томского государственного университета, 2014. – С. 114–115.

13. *Дмитриев Ю.Г., Зенкова Ж.Н.* Об одной оценке симметричной функции распределения по цензурированной выборке // Материалы IV Всероссийской научно-практической конференция "Информационные технологии и математическое моделирование". Анжеро-Судженск. 2005. – Томск: Изд-во Том. ун-та, 2005. – Ч. 2. – С. 8–10.

ПРОГНОЗИРОВАНИЕ ТЕНДЕНЦИИ МАКСИМАЛЬНОЙ ЦЕНЫ АКЦИИ

Ю.Г. Дмитриев, Ю.Е. Пальянова
Томский государственный университет
dmit@mail.tsu.ru, pje0360@gmail.com

Введение

В работе рассматривается применение математической модели растущих систем [1–3] к моделированию и прогнозированию цен акций. Под растущей системой понимается любая совокупность объектов одной природы, которые вместе участвуют в процессах возникновения, создания и развития системы, подчиняющихся законам, одинаковым для всех объектов. Описание растущих систем возможно с помощью вероятностно-статистических моделей [1,2]. Системы подобного типа используются в геологии для изучения совокупности месторождений полезных ископаемых [4,5], для выявления закономерностей численности населения различных территорий [6], определения экстремального расхода воды в реках [7], расчета финансовых показателей и в других сферах практической деятельности.

В данной работе, исходя из предположения, что совокупность цен акций за определенный период времени является растущей системой, проводится подбор математической модели, которая отражает зависимость максимальных цен акций за разные периоды времени. Полученная модель приводит к специфической зависимости между ценами акций, позволяющей предсказывать тенденцию изменения их максимальных значений. Результаты применения модели иллюстрируются на модельных данных, а также данных стоимостей акций публичного акционерного общества «Газпром».

1. Описание модели для анализа цен акций

Пусть имеется N временных периодов с одинаковым количеством n дней (в рамках растущей системы это означает, что имеется N ячеек накопления). Наблюдается максимальное значение цены акции в каждый из n дней. Максимальные цены между днями будем считать независимыми и одинаково распределенными случайными величинами с функцией распределения $F(x)$. Обозначим максимальную цену акции за n дней в k-м периоде через $X_k^{(n)}$ и рассмотрим последовательность $X_1^{(n)},...,X_N^{(n)}$, отвечающую N периодам. При этом функция распределения $G_{(n)}(x) = P\left\{X_k^{(n)} < x\right\} = F^n(x)$.

Выберем в качестве $F(x)$ двойной экспоненциальный закон $F(x;\mu,\sigma) = \exp\left\{-\exp\left\{-(x-\mu)/\sigma\right\}\right\}$, $-\infty < x < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$, предложенный в [7] для описания максимальных значений наблюдаемой величины. Упорядочим на-

блюдения по возрастанию: $X_{(1)}^{(n)} < X_{(2)}^{(n)} < ... < X_{(N)}^{(n)}$. Порядковые статистики $X_{(k)}^{(n)}$ можно рассматривать как оценки квантилей закона распределения $G_{(n)}(x)$ уровней $p_k = k/(N+1)$, $k = \overline{1,N}$. На основании этого, следуя [8], можно записать, $F_0^n\left(\left(X_{(k)}^{(n)} - \mu\right)/\sigma\right) = p_k + \varepsilon_{(k)}$, где $\varepsilon_{(k)} = u_{(k)} - p_k$, $u_{(k)}$ – k-я порядковая статистика выборки из равномерного в $(0,1)$ закона. Отсюда

$$\frac{X_{(k)}^{(n)} - \mu}{\sigma} = F_0^{-1}\left(\left(p_k + \varepsilon_{(k)}\right)^{1/n}\right),\ k = \overline{1,N},\ \frac{X_{(k)}^{(n)} - \mu}{\sigma} = F_0^{-1}\left(p_k^{1/n}\right) + \varepsilon_{(k)},\ k = \overline{1,N}, \qquad (1)$$

$$X_{(k)}^{(n)} \approx \mu + \sigma \cdot Y_{(k)}^{(n)} + \varepsilon_{(k)},\ k = \overline{1,N}, \qquad (2)$$

где $Y_{(k)}^{(n)} = F_0^{-1}\left(p_k^{1/n}\right) = \left(-\ln\left(-n^{-1}\ln\left(p_k\right)\right)\right)$, $k = \overline{1,N}$, $F_0(x) = F(x;0,1)$ с $M\varepsilon_{(k)} = 0$ и матрицей ковариации $V = \operatorname{cov}\left(\varepsilon_{(k)}, \varepsilon_{(k')}\right) = \dfrac{k(N-k'+1)}{(N+1)^2(N+2)}$, $k, k' = \overline{1,N}$ для порядковых статистик из равномерного закона в $(0,1)$. Модель (2) отражает линейную регрессионную зависимость значений максимальных цен акций от квантилей закона распределения $F_0(x)$. Параметры μ и σ находятся по обобщенному методу наименьших квадратов (ОМНК) [8].

Рассмотрим уравнение (1), разложив функцию F_0^{-1} по формуле Тейлора:

$$\frac{X_{(k)}^{(n)} - \mu}{\sigma} = F_0^{-1}\left(p_k^{1/n}\right) + \left[F_0^{-1}\left(p_k^{1/n}\right)\right]' \varepsilon_{(k)} + \frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]'' \varepsilon_{(k)}^2 + O\left(\varepsilon_{(k)}^3\right),$$

$$X_{(k)}^{(n)} = \mu + \sigma\left(F_0^{-1}\left(p_k^{1/n}\right) + \left[F_0^{-1}\left(p_k^{1/n}\right)\right]' \varepsilon_{(k)} + \frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]'' \varepsilon_{(k)}^2 + O\left(\varepsilon_{(k)}^3\right)\right),$$

$$EX_{(k)}^{(n)} = \mu + \sigma\left(F_0^{-1}\left(p_k^{1/n}\right) + \frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]'' E\varepsilon_{(k)}^2 + O\left(E\varepsilon_{(k)}^3\right)\right) + \sigma\left[F_0^{-1}\left(p_k^{1/n}\right)\right]' E\varepsilon_{(k)}.$$

Так как $\varepsilon_{(k)}$ – это порядковые статистики из равномерного распределения, то согласно [8] при $k = k'$ получим $E\varepsilon_{(k)}^2 = \dfrac{k(N-k+1)}{(N+1)^2(N+2)}$.

$$EX_{(k)}^{(n)} = \mu + \sigma\left(F_0^{-1}\left(p_k^{1/n}\right) + \frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]'' \cdot \frac{k(N-k+1)}{(N+1)^2(N+2)} + O\left(E\varepsilon_{(k)}^3\right)\right) + \sigma\left[F_0^{-1}\left(p_k^{1/n}\right)\right]' E\varepsilon_{(k)}.$$

Если пренебречь величиной $O\left(E\varepsilon_{(k)}^3\right)$, то

$$X_{(k)}^{(n)} = \mu + \sigma\left(F_0^{-1}\left(p_k^{1/n}\right) + \frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]'' \cdot \frac{k(N-k+1)}{(N+1)^2(N+2)}\right) + \sigma\left[F_0^{-1}\left(p_k^{1/n}\right)\right]' E\varepsilon_{(k)},$$

$$EX_{(k)}^{(n)} = \mu + \sigma\left(F_0^{-1}\left(p_k^{1/n}\right) + \frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]'' \cdot \frac{k(N-k+1)}{(N+1)^2(N+2)}\right). \qquad (3)$$

Обозначим $\alpha_{(k)}^{(n)} = F_0^{-1}\left(p_k^{1/n}\right) + \frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]'' \cdot \frac{k\left(N-k+1\right)}{\left(N+1\right)^2\left(N+2\right)}$. Подставив $\alpha_{(k)}^{(n)}$ в (3), получим $EX_{(k)}^{(n)} = \mu + \sigma\alpha_{(k)}^{(n)}$. Таким образом, имеем уточненную, по сравнению с (2), модель:

$$X_{(k)}^{(n)} = \mu + \sigma \cdot \alpha_{(k)}^{(n)} + \varepsilon_{(k)}, \; k = \overline{1, N}. \tag{4}$$

2. Алгоритм прогноза

Задача заключается в прогнозировании тенденции (увеличении или уменьшении) максимальной цены акции в $N+1$ периоде по сравнению с предыдущим. Прогнозирование тенденции максимальной цены акции в $N+1$ периоде предлагается проводить следующим образом [1]. Пусть известны максимальные значения цен акций за N периодов, и вычислены квантили $Y_{(1)}^{(n)} < \ldots < Y_{(N+1)}^{(n)}$ за $N+1$ периодов. Поставим в соответствие N упорядоченным значениям цен $X_{(k)}^{(n)}$ N квантилей $Y_{(k)}^{(n)}$ из $N+1$ возможных. Наилучшим будем считать такое соответствие, при котором погрешность

$$SKO = \sqrt{\sum_{k=1}^{N}\left(X_{(k)}^{(n)} - \hat{X}_{(k)}^{(n)}\right)^2}, \tag{6}$$

где $\hat{X}_{(k)}^{(n)} \approx \hat{\mu} + \hat{\sigma}Y_{(k)}^{(n)} + \varepsilon_{(k)}$, $k = \overline{1, N}$, ($\hat{X}_{(k)}^{(n)} = \hat{\mu} + \hat{\sigma}\alpha_{(k)}^{(n)} + \varepsilon_{(k)}$, $k = \overline{1, N}$) будет наименьшей по всем возможным переборам соответствий с учетом порядка, а также $\hat{\mu}$ и $\hat{\sigma}$, вычисляемых по ОМНК для каждого соответствия. Когда такое соответствие найдено, то становится известен «свободный» квантиль $Y_{(k)}^{(n)}$, которому соответствия не нашлось. Подставляя этот квантиль в оцененное уравнение регрессии при наилучшем соответствии, находим прогнозное значение неизвестной цены акции на $N+1$ период и, соответственно, тенденцию цены.

3. Прогнозирование модельных данных

Возьмем в качестве $F(x)$ двойной экспоненциальный закон распределения $F(x;\mu,\sigma) = \exp\left\{-\exp\left\{-\left(x-\mu\right)/\sigma\right\}\right\}$. Сгенерируем 100 выборок для каждого набора значений параметров из этого закона с заданными μ, σ, N и n. Следовательно, можно увидеть растет или убывает значение $N+1$ периода. При этом качество прогноза будем отслеживать по количеству верных прогнозов. Для сравнения будем использовать две модели, определенными формулами (2) и (4).

В табл. 1 приведены результаты по прогнозированию значения на $N+1$ период при фиксированном μ и разных N, n и σ. Также показано, сколько раз тенденция этого значения была верно определена. По результатам прогноза оказалось, что обе модели (2) и (4) улавливают тенденции одинаково.

Таблица 1

Результаты прогноза на *N* + 1 период

Значения параметров при *n* = 1	Процент верных прогнозов тенденции, %	Значения параметров при *n* = 5	Процент верных прогнозов тенденции, %
μ = 2, σ = 5, *N* = 4	85	μ = 2, σ = 5, *N* = 4	65
μ = 2, σ = 5, *N* = 9	80	μ = 2, σ = 5, *N* = 9	60
μ = 2, σ = 0.1, *N* = 4	90	μ = 2, σ = 0.1, *N* = 4	70
μ = 2, σ = 0.1, *N* = 9	90	μ = 2, σ = 0.1, *N* = 9	70

Табл. 1 показывает, что наибольшее влияние на точность прогноза тенденции оказывает масштабный параметр σ: чем больше σ при фиксированном μ, тем хуже модель оценивает все параметры, и тем сложнее определить значение на $N+1$ период. При этом количество элементов N в выборке не так сильно влияет на качество прогноза.

В табл. 2 приведены результаты по прогнозированию значения на $N+1$ период при фиксированном σ и разных N, n и μ. Также показано, сколько раз тенденция этого значения была верно определена.

Таблица 2

Результаты прогноза на $N+1$ период

Значения параметров при $n = 1$	Процент верных прогнозов тенденции, %	Значения параметров при $n = 5$	Процент верных прогнозов тенденции, %
μ = 2, σ = 2, N = 4	80	μ = 2, σ = 2, N = 4	65
μ = 2, σ = 2, N = 9	76	μ = 2, σ = 2, N = 9	61
μ = 1000, σ = 2, N = 4	90	μ = 1000, σ = 2, N = 4	75
μ = 1000, σ = 2, N = 9	85	μ = 1000, σ = 2, N = 9	70

По результатам из табл. 2 можно сделать вывод о том, что разные значения параметра положения μ несильно влияют на качество прогноза.

Таким образом, наиболее точный результат прогнозирования тенденции для значения в $N+1$ период достигается при $n=1$ и любом объёме выборки N. При этом оказалось, что обе модели (2) и (4) улавливают тенденции одинаково. Следовательно, данные результаты целесообразно использовать для дальнейшего прогноза реальных цен акций.

4. Прогноз тенденции по реальным данным

Рассмотрим реальные данные: цены акций ПАО «Газпром» на 15 временных интервалах за 2006–2018 гг. Поскольку торги на бирже открыты с понедельника по пятницу, то за неделю возьмем пять дней, т.е. $N=4$, $N+1=5$. $n=1$. Точность подгонки будем определять по наименьшему среднеквадратическому отклонению из (6). Сравнивать полученные тенденции цен будем по моделям из (2) и (4). В табл. 3–4 знаком «+» показано, что данный способ отразил тенденцию цены на пятый день, т.е. если реальная цена росла, то и прогнозное значение этой цены также росло и наоборот, если реальная цена падала, то и прогноз падал. Соответственно, знак «–» отражает тот факт, что выбранный способ не показывает реальную ситуацию. Приведем результаты прогноза.

Таблица 3

Результаты прогнозирования

№ п/п	СКО для модели (2)	Прогнозируемая цена для модели (2), руб.	СКО для модели (4)	Прогнозируемая цена для модели (4), руб.
1	0.709	139.064 (+)	0.552	139.138 (+)
2	0.676	134.188 (+)	0.608	133.77 (+)
3	0.592	140.899 (+)	0.496	140.97 (+)
4	0.091	130.147 (+)	0.164	130.191 (+)
5	0.435	131.294 (+)	0.438	132.086 (+)
6	0.019	129.329 (+)	0.045	129.703 (+)
7	0.589	129.743 (+)	0.531	129.858 (+)
8	0.502	126.484 (–)	0.283	126.678 (–)
9	0.384	122.635 (+)	0.291	122.673 (+)
10	0.24	129.394 (–)	0.29	140.052 (+)
11	0.138	143.196 (+)	0.211	143.261 (+)
12	0.218	142.407 (+)	0.209	142.388 (+)
13	0.228	131.705 (+)	0.167	133.455 (–)
14	0.368	122.902 (–)	0.292	123.028 (–)
15	1.55	121.504 (+)	1.351	121.634 (+)

За представленный период реальная ситуация отражена в 12 случаях из 15, что соответствует 80%. При оценивании тенденции значения цены на пятый день разница между моделями (2) и (4) несущественна.

Рассмотрим следующую ситуацию. Предположим, что известны цены акций за четыре предыдущих периода (недели). Требуется спрогнозировать тенденцию цены акции на пятую неделю. Для данного случая рассмотрим 15 наблюдений в период с 2006 по 2018 год, т.е. $N = 4$, $N + 1 = 5$, $n = 5$. Приведём результаты прогноза.

Таблица 4

Результаты прогнозирования

№ п/п	СКО для модели (2)	Прогнозируемая цена для модели (2), руб.	СКО для модели (4)	Прогнозируемая цена для модели (4), руб.
1	1.824	146.331 (+)	1.669	146.003 (+)
2	0.856	134.892 (+)	1.114	133.921 (+)
3	3.379	130.143 (–)	4.684	128.644 (–)
4	1.019	125.703 (+)	0.888	125.759 (+)
5	0.248	122.157 (+)	0.183	122.098 (+)
6	0.846	118.526 (+)	0.611	118.797 (+)
7	0.679	122.686 (+)	0.484	122.447 (+)
8	2.331	128.79 (–)	2.796	129.615 (–)
9	1.697	144.605 (–)	1.491	144.63 (–)
10	3.474	131.205 (+)	3.409	131.495 (+)
11	0.49	146.015 (–)	0.624	145.863 (–)
12	2.244	143.408 (–)	3.179	142.419 (–)
13	1.209	157.72 (+)	0.961	157.345 (+)
14	1.503	145.883 (–)	2.153	145.255 (–)
15	0.614	138.679 (+)	0.549	138.629 (+)

За рассмотренный период реальная ситуация проявилась в 9 случаях из 15, что соответствует 60%. При оценивании значения тенденции цены на пятую неделю разница между моделями (2) и (4) несущественна.

Заключение

В работе рассмотрено применение математических моделей растущих систем для анализа максимальных цен акций с помощью линейной регрессии, основанной на квантилях равноотстоящих уровней двойного экспоненциального закона распределения. На примерах показаны прогностические свойства моделей (2) и (4), как на модельных, так и на реальных данных. Выявлено, что разница в прогностических свойствах между этими моделями несущественна. Полученные результаты целесообразно использовать для дальнейшего изучения прогнозных тенденций и реальных цен акций.

ЛИТЕРАТУРА

1. *Дмитриев Ю.Г., Устинов Ю.К.* Математические модели растущих систем // Журнал «Вычислительные технологии». – 2007. – Т. 12. – №1. – С. 68–75.

2. *Dmitriev Yu.G., Tarasenko F.P.* Towards a Modeling of Natural Systems // Cybernetics and Systems: An International Journal. – 2009. – № 40. – P. 236–248.

3. *Дмитриев Ю.Г., Тарасенко П.Ф.* Интерпретация геологических данных с использованием компьютерных технологий // Журнал «Известия ТПУ». – 2008. – Т. 313. – № 5. – С. 26–32.

4. *Дмитриев Ю.Г., Тарасенко П.Ф.* Автоматизированная система «Октава» для геологического прогнозирования СО РАН // Новосибирск: Вычислительные технологии, 2003. – Вып. 8.

5. *Дмитриев Ю.Г., Тарасенко П.Ф.* Об одной природной закономерности // Материалы Всероссийской конференции с международным участием (ЗОНТ-09), Новосибирск, 22–24 октября 2009. – Т.1. – С. 4–13.

6. *Талейко А.В.* Закономерности формирования численности населения по территориям // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2009. – № 3. – С. 80–86.

7. *Смирнов Н.В., Дунин-Барковский И.В.* Курс теории вероятностей и математической статистики для технических приложений. – М.: Наука, 1969. – 512 с.

8. Введение в теорию порядковых статистик: пер. с англ. / под редакцией А.Я. Боярского – М.: Статистика, 1970. – 413 с.