

наличии коинтеграции между преступлениями, среднедушевыми доходами и безработицей.

СПИСОК ЛИТЕРАТУРЫ

1. Краткая характеристика состояния преступности в Российской Федерации за январь-май 2018 г. 2018: Стат. сб. МВД. [Электронный ресурс]. URL: <https://мвд.рф/reports> (дата обращения: 20.08.2018).
2. Pyle D. J., Deadman D. F. Crime and the Business Cycle in Post-War Britain // The British Journal of Criminology. 1994. Vol. 34. Vol. 37.
3. Mohsin H. M., Jabbar S.M. Economics of Property Crime Rate in Punjab // The Pakistan Development Review. Vol. 52, No. 3. 2017.
4. Андриенко Ю. В. В поисках объяснения роста преступности в России в переходный период: криминометрический подход // Экономический журнал НИУ ВШЭ. 2001. Т. 5. № 2. С. 194-220.

ИССЛЕДОВАНИЕ ВЗАИМОСВЯЗИ МЕЖДУ ИНТЕНСИВНОСТЬЮ НОВОСТНОГО ПОТОКА И ОБЪЕМОМ БИРЖЕВЫХ ТОРГОВ

А. З. Сарсенова

Саратовский государственный университет, Россия
E-mail: aizhan.xd@gmail.com

Константное состояние цен на биржевом рынке – довольно редкое явление. На биржевом рынке постоянно происходят колебания цен, которые называют волатильностью. Различные явления и факторы могут оказывать влияние на данную характеристику. В статье рассматривается возможность существования взаимосвязи между интенсивностью новостной аналитики и объемом биржевых торгов. Для этого будет построена модель линейной регрессии, на основе оценки которой можно будет судить о влиянии новостей на волатильность торгов.

RESEARCH OF INTERRELATIONSHIP BETWEEN THE INTENSITY OF NEWS ANALYTICS AND THE VOLUME OF STOCK TRADING

A. Z. Sarsenova

The constant state of prices on the stock market is quite rare. The stock market is constantly fluctuations of prices. These fluctuations are call volatility. Various phenomena and factors can influence this characteristic. The article discusses the possibility of the interrelationship between the intensity of news analytics and the volume of stock trading. For this, a linear regression model will be construct. Based on the evaluation of this model, it will be possible to judge the impact of news on the volatility of trading

Волатильность – это статистический показатель, характеризующий изменение цены за некоторый промежуток времени. Это очень важный показатель, так как представляет собой меру риска использования финансового инструмен-

та за определенный интервал времени.

Для расчета волатильности обычно применяется выборочное стандартное отклонение, что позволяет инвесторам с некоторой точностью определить риск приобретения финансового инструмента [1]. Тем не менее использование выборочного стандартного отклонения не всегда дает верный результат, данная выборочная статистика обладает рядом недостатков. Полученная таким образом оценка волатильности будет состоятельной, то есть будет сходиться к истинному значению, только в случае асимптотики, что требует бесконечно большой объем данных. Также этот показатель не позволяет идентифицировать изменение волатильности во времени. Одним из популярных примеров отображения волатильности в финансово-математических моделях является уравнение Блэка-Шоулза, однако в этом случае волатильность подразумевают как постоянную величину, что на практике бывает не всегда. Сама оценка волатильности складывается из исследований исторических данных, которые позволяют проследить за динамикой цен в прошлом, и, исходя из этих данных, строится гипотеза о потенциальной изменчивости.

Известно, что различные виды активов имеют периоды низкой и высокой волатильности, то есть в некоторые моменты времени цена может меняться быстро, тогда как в другое время она практически не изменяется [1]. Безусловно, такие колебания скорее всего не возникают на пустом месте, должно быть существуют некоторые компоненты, оказывающие влияние на волатильность. Такими факторами фундаментального характера могут быть релизы важных экономических отчетов, решения в отношении монетарной политики, принимаемые Центральными банками стран, какие-либо события, происходящие в политической сфере государства – одним словом новости.

Для исследования возможности существования взаимосвязи между интенсивностью новостной аналитики и объемом биржевых торгов будет построена линейная регрессионная модель. Для этого необходимы данные новостной аналитики и объемы торгов по различным компаниям.

Новостные данные были предоставлены одним из ведущих провайдеров Big Data-аналитики для финансовых учреждений – RavenPack. RavenPack преобразовывает массивы неструктурированных больших данных, состоящих из новостей масс-медиа и информации из социальных сетей, в структурированные данные, в состав которых входят даже индикаторы настроения и внимания общественности к СМИ. Платформа, предоставляющая новостную аналитику по более чем 28 тыс. компаниям со всего мира, называется Raven Pack News Analytics (RPNA). RPNA анализирует каждую новость, публикуемую профессиональными поставщиками новостей (такими как Dow Jones или Reuters), а также сотни финансовых сайтов, онлайн-газет и даже блогов [2].

Полученные данные – это новости, вышедшие в период с 1 января 2015 года по 31 сентября 2015 года. Сами данные представлены в обычном текстовом формате с расширением «.txt». Каждую новость RPNA представляет как массив, состоящий из следующих атрибутов: время выхода, id компании, релевантность новости, взвешенное настроение новости, тип новости, категория события и название компании. При этом для анализа необходимы только назва-

ние компании и дата выхода новости, связанной с этой фирмой.

Информация об объеме биржевых торгов компаний была получена из базы данных сервиса Yahoo!Finance в соответствующий новостям период, то есть с 1 января 2015 года по 31 сентября 2015 года, по каждой компании. Вид биржевой информации об Apple Inc. представлен следующими полями: дата торгов, цена открытия, максимум цены, минимум цены, цена закрытия, скорректированная цена закрытия, объем торгов. Для анализа нас интересуют только дата и объем торгов.

Для реализации регрессионного анализа был использован язык программирования R, предназначенный для интеллектуального анализа и визуализации данных. В 2012 году язык R занял 1 место (30,7%) в опросе от KD Nuggets (лидирующий сайт в области бизнес-аналитики, больших данных, интеллектуального анализа данных, науки о данных и машинного обучения) как программное обеспечение, которое широко применяется для обработки данных, то есть данный язык программирования зарекомендовал себя как надежный инструмент для статистической обработки данных [3].

Математическая модель линейной регрессии выражается через формулу (1):

$$y = a + bx, \quad (1)$$

где x – это исходные данные, а y – предсказываемая величина. Формальное описание задачи звучит так: найти коэффициенты a и b , минимизирующие величину ошибки. То есть для нашего исследования уравнение (1) можно представить как формулу (2). Формула (2) описывает предположение о том, что существует связь между объемом торгов некоторой компании в определенный день и количеством новостей, вышедших в этот день по данной компании.

$$\text{Объем торгов} = K * \text{Количество новостей} + \text{Некая константа} \quad (2)$$

Для реализации линейной регрессии в языке R используется функция *lm*, а для вывода полученных данных – функция *summary*. На рис. представлен вывод функции для компании Apple Inc. Для доказательства или опровержения гипотезы нам необходимо значение величины «p-value». Эта величина показывает, насколько объем торгов зависит от количества новостей, вышедших в этот день. Для этого выдвигается нулевая гипотеза, что предсказываемая величина вообще не зависит от предикторов. В нашем случае значение «p-value» равно $9.133e-12$, то есть на 99,9999990867% предсказываемая величина (объем торгов) зависит от предикторов (количество новостей).

```

Call:
lm(formula = total$Volume ~ total$count)

Residuals:
    Min       1Q   Median       3Q      Max
-34140019 -12568066  -3480645   7545155  92723606

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 42246499   2251224  18.766 < 2e-16 ***
total$count  495204     67848    7.299 9.13e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 19870000 on 179 degrees of freedom
Multiple R-squared:  0.2294,    Adjusted R-squared:  0.225
F-statistic: 53.27 on 1 and 179 DF,  p-value: 9.133e-12

```

Результат работы функции lm для Apple Inc.

В соответствии с таблицей «Значения показателя p-value» были получены следующие величины по 10 компаниям: Apple, Google, Microsoft, Amazon, Tesla, Facebook, Intel, General Electric, IBM и eBay. Столбец «Зависимость» отражает процентное соотношение связи объема торгов и количества новостей.

Таблица

Значения показателя p-value

№	Компания	p-value	Зависимость
1	Apple	9.133e-12	99,9999990867%
2	Google	9.154e-14	99,99999990846%
3	Microsoft	3.386e-11	99,999996614%
4	Amazon	2.2e-16	99,99999999999978%
5	Tesla	3.849e-11	99,999996151%
6	Facebook	1.224e-10	99,99998776%
7	Intel	4.73e-07	99,0000473%
8	General Electric	2.2e-16	99,9999999999978%
9	IBM	3.084e-08	99,996916%
10	eBay	0.0007676	99,92324%

Исходя из полученных данных можно сделать вывод, что связь между интенсивностью новостного потока и объемом биржевых торгов существует. Но помимо линейной регрессии есть множество других математических моделей, с помощью которых можно не только получать сведения о зависимости между предсказываемой величиной и предикторами, но также прогнозировать будущие значения.

Работа выполнена при финансовой поддержке гранта Президента РФ (проект НШ-2781.2012.2) и РФФИ (проект 12-07-00057).

СПИСОК ЛИТЕРАТУРЫ

1. *Ананьев М. А., Митин Н. А.* Сравнение линейных и нелинейных авторегрессионных моделей условной гетероскедастичности на примере индекса доходности РТС // Препринты

ИПМ им. М. В. Келдыша. 2013. № 24. С. 222-229.

2. Сидоров С. П., Дате П., Балаш В. А. Использование данных новостной аналитики в GARCH моделях // Прикладная эконометрика. 2013. Т. 1. № 29. С. 82-96.

3. Chen Min. Big Data Related Technologies, Challenges and Future Prospects // Pearson Education International. 2014. С. 89.

МОДЕЛИ БИНАРНОГО ОТКЛИКА И ПОРЯДКОВОЙ РЕГРЕССИИ В ИССЛЕДОВАНИИ СТРАХОВОГО ПОВЕДЕНИЯ

Т. Г. Синявская, А. А. Трегубова

Ростовский государственный экономический университет, Россия

Email: sin-ta@yandex.ru, alexandra_a_t@mail.ru

В статье приводятся результаты моделирования страхового поведения населения России с использованием инструментария бинарной и порядковой логистической регрессии. Используются данные 25 волны опроса «Российский мониторинг экономического положения и здоровья населения НИУ-ВШЭ (RLMS-HSE)» за 2016 год. Показано, что оба типа моделей дают адекватный результат и могут быть использованы для анализа. На страховое поведение оказывают значимое влияние факторы дохода, возраста, пола, уровня образования, состояния в браке, типа поселения, удовлетворенности жизнью респондента, исповедание ислама, а также демонстрации иных, помимо страхового, форм финансового поведения. Полученные результаты свидетельствуют, что страховое поведение является одним из видов финансовой активности, которую склонны проявлять индивиды, достигшие определенного уровня благосостояния, и не относится, таким образом, к поведенчески-обусловленным инструментам управления финансовыми рисками.

BINARY AND ORDINAL LOGISTIC REGRESSION MODELS IN THE STUDY OF INSURANCE BEHAVIOR

T. G. Sinyavskaya, A. A. Tregubova

The paper presents the results of modeling the population's insurance behavior using binary and ordinal logistic regression. Individual Russia Longitudinal Monitoring Survey – Higher School of Economics (RLMS-HSE) Round 25 data is used for model estimation. It is showed that both models give an adequate result and can be used for analysis. The factors of income, age, gender, education level, marital status, settlement type, respondents' satisfaction with their lives, Islamic religion professed, and different financial strategies used, affect significantly insurance behavior of Russians. These results suggest that insurance behavior is one of the types of financial activity that is manifested by individuals who have reached a certain level of welfare. Thus, insurance behavior as financial risk management instrument is not guided by the behavioral characteristics of individuals.

Страховое поведение населения России является одной из самых малоизученных форм финансового поведения. Как показано в ряде зарубежных исследований (например, в работах Грина, 1963 [1]; Хоканссона, 1969 [2]; Кэмпбелла, 1980 [3]), страховое поведение зависит от общей финансовой стратегии