

## СТАТИСТИЧЕСКИЙ АНАЛИЗ ПОТРЕБИТЕЛЬСКИХ ПРЕДПОЧТЕНИЙ В ЭЛЕКТРОННОЙ КОММЕРЦИИ\*

О.В. Никитина

*В статье рассматриваются проблемы статистического анализа потребительских предпочтений в электронной коммерции (на примере гостиничной отрасли). Автором проведен множественный регрессионный анализ для оценки влияния различных факторов на потребительский выбор и интернет-продажи в данной отрасли. С использованием ценностно-ориентированного маркетингового подхода в статье предложен метод расчета индекса воспринимаемой потребительской ценности (карты ценности) в электронной коммерции. Массив используемых в статье для статистического анализа данных составляет около 1,5 млн наблюдений (отзывов потребителей о качестве услуг отелей в 13 популярных туристических городах мира из рейтинга агентства Euromonitor International).*

*Наиболее важным результатом статистического исследования потребительских предпочтений в электронной коммерции (на примере гостиничного рынка) представляется подтверждение гипотезы о том, что доступность оценок других покупателей с помощью информационных технологий позволяет потребителям более точно прогнозировать для себя ценность данных услуг, чем ориентируясь на другие сигналы качества, такие, как количество звезд и бренд (принадлежность к сети). Таким образом, бронирование гостиниц через онлайн-системы будет вести к большему удовлетворению потребителя, чем с использованием других каналов сбыта гостиничных услуг.*

*Ключевые слова:* потребительские предпочтения, воспринимаемая потребительская ценность, методы измерения качества продукта, статистический анализ больших данных, электронная коммерция, онлайн-туризм.

*JEL:* C35, D46, L83, M31.

Стремительное распространение информационных технологий и существенный рост количества интернет-пользователей за последнее десятилетие стали причиной того, что все большее число потребителей использует электронную торговлю для приобретения товаров и услуг. По данным [4], объем электронных продаж в 2014 г. насчитывал 840 млн долларов США, что означает более чем 20%-ный рост по сравнению с предыдущим годом. В целом ряде отраслей интернет-торговля становится одним из основных каналов продаж. Среди лидеров в этом направлении - гостиничная индустрия, где онлайн-продажи, по оценке исследовательской компании Phocus Wright, занимают долю в 13% (по данным за 2013 г.).

С точки зрения возможностей проведения статистических исследований, важной особенностью развития этого рынка стало появление в открытом онлайн-доступе огромных массивов данных, характеризующих потребительский спрос и предпочтения для различных товаров и услуг. В маркетинге анализ многих параметров принятия решения о покупке, ранее доступный для исследования толь-

ко в рамках опросов небольших групп потребителей (выборки), теперь стало возможным проводить статистически на основе действительно больших данных - о миллионах интернет-пользователей.

*Большими данными* принято называть наборы данных, размер которых превосходит возможности типичных баз данных по занесению, хранению, управлению и анализу информации. Помимо огромных объемов, их основные характеристики - необходимость в быстрой обработке и высокая скорость накопления этих данных, а также их разнообразие. Это понятие было введено Клиффордом Линчем в 2008 г. в статье «Как могут повлиять на будущее науки технологии, открывающие возможности работы с большими объемами данных?» [5].

Доступность такого рода объемов открытых данных, активное развитие статистических методов работы с полными данными (в отличие от выборок), а также совершенствование компьютерной техники для проведения комплексных статистических расчетов с большими наборами данных стимулировали появление ряда статистических исследований в мар-

*Никитина Ольга Владимировна (o.nikitina@rospravо.ru) - руководитель направления маркетинга российской юридической фирмы «Муранов, Черняков и партнеры», аспирант кафедры стратегического маркетинга, Национальный исследовательский университет «Высшая школа экономики».*

\* На примере гостиничной отрасли.

кетинге, основанных на отзывах и рекомендациях покупателей в Интернете. В частности, К. Форман, Э. Гоз и Б. Визенфельд [2] изучили влияние отзывов и рейтингов потребителей на продажи книг. В. Дьюан, Б. Гу и Э. Винстон [1] исследовали эффект онлайн-отзывов в киноиндустрии. Изучению влияния онлайн-рейтингов потребителей на гостиничном рынке положили начало Х. Огют и Б. Тас [6].

Несмотря на значительное количество таких исследований, где предлагаются подходы к измерению потребительских предпочтений на рынке электронной коммерции с помощью отзывов и рейтингов потребителей, в большинстве случаев они адаптированы к специфике отдельных отраслей и не несут в себе единой методологии измерения.

В связи с этим данным исследованием сделана попытка статистического изучения предпочтений потребителей в электронной коммерции (на примере гостиничного рынка) с использованием универсальной общепринятой маркетинговой модели «карты ценности»: оценки восприятия потребителями соотношения между воспринимаемой ценой и воспринимаемым качеством конкурирующих предложений и метода регрессионного анализа большого объема данных.

**Теоретическая база исследования.** Одним из основных направлений развития современной теории маркетинга является ценностно-ориентированный подход, ключевая идея которого состоит в том, что компании продают потребителям не товары или услуги, а совокупность физических, эмоциональных и других характеристик продукции, имеющую для них ценность.

При этом ценности являются субъективными, поэтому ключевой категорией называется «воспринимаемая потребителем ценность». Существующие определения этого понятия сводятся к следующему: «воспринимаемая потребителем ценность» - это относительное, воспринимаемое потребителем качество товара или услуги, которое рассматривается в контексте воспринимаемой цены в сравнении с конкурирующими товарами/услугами.

Наиболее часто воспринимаемую ценность определяют с помощью оценок экспертов и

опросов потребителей. В качестве измерительного инструментария данная категория предлагает модель «карты ценности», которая впервые была определена в основополагающей работе Бредли Гэйла «Управление потребительской ценностью» [3]. Карта ценности является графическим отображением того, как потребители оценивают соотношение между воспринимаемой ценой и воспринимаемым качеством конкурирующих предложений.

Для построения карты ценности на оси абсцисс откладывается индекс воспринимаемого качества продукта, а на оси ординат - индекс его воспринимаемой цены. Ниже представлены формулы расчета данных показателей:

*Индекс воспринимаемого качества = (воспринимаемое качество / среднее воспринимаемое качество по рынку) × 100;*

*Индекс воспринимаемой цены = (воспринимаемая цена / средняя воспринимаемая цена по рынку) × 100;*

*Воспринимаемая ценность = индекс воспринимаемого качества - индекс воспринимаемой цены.*

Таким образом, индекс воспринимаемой ценности определяет покупательское поведение, отражая то, насколько чувствительными являются восприятие и предпочтения потребителя к изменениям в свойствах и цене продукта.

По диагонали карты ценности, построенной на основе этого индекса, располагается «линия эквивалентной ценности» (value equivalence line, VEL), каждая точка которой представляет собой баланс индексов воспринимаемой цены и воспринимаемого качества.

Построение карты ценности имеет большое практическое значение, так как визуализация распределения покупателей и конкурентов на карте ценности позволяет компаниям оценить позиционирование собственного предложения.

Как представляется, данная теоретическая модель несет в себе универсальные возможности для анализа потребительских предпочтений с использованием открытых массивов данных отзывов и рекомендаций потребителей на рынке электронной коммерции. Тем не менее важно отметить необходимость ча-

стого мониторинга и обновления данных: карта ценности рынка претерпевает постоянные изменения, конкуренты могут повышать и понижать цены, изменять характеристики своих товаров/услуг.

**Исходные данные.** Для проведения исследования с помощью автоматического скрипта, написанного на языке программирования, был собран массив данных из 1514150 отзывов потребителей с крупнейшего сайта онлайн-бронирования Booking.com, насчитывающего в целом более 47 млн отзывов, и характеристик потребительского спроса для каждого отзыва. Совокупность данных включала оценки потребителей, забронировавших отель через данный сайт в период с января 2013 г. по март 2015 г. в следующих 13 городах: Анталия, Бангкок, Барселона, Гонконг, Стамбул, Лондон, Макао, Москва, Париж, Санкт-Петербург, Пхукет, Прага, Сингапур. Таким образом,

в составе перечня городов пять крупнейших туристических городов в 2014 г. по данным агентства Euromonitor International (мирового лидера в области международных исследований потребительских рынков), крупнейшие российские города по входящему туристическому и деловому потоку, а также выборочно города из первых 50.

Собранные данные отзывов были агрегированы по отелям и оставлены только те отели, которые предоставляли информацию о цене номера (2124 отеля). Для каждого из них была отобрана следующая информация: средний рейтинг клиентов, общее число отзывов, количество звезд, стоимость одноместного размещения (переведенная в рубли по среднему курсу за рассматриваемый временной период) и другие показатели.

Ключевые описательные статистики для понимания характера получившейся совокупности данных представлены в таблице 1.

Таблица 1

Описательные статистики исследуемого набора данных

	Количество наблюдений (отелей)	Минимальное значение	Максимальное значение	Среднее значение	Стандартное отклонение
Количество отзывов об отеле	2124	5,00	6070,00	473,56	513,39
Рейтинг отеля (среднее арифметическое значение оценок потребителей)	2124	4,34	9,76	8,11	0,69
Количество звезд	1527	1,00	5,00	3,48	0,90
Цена одноместного номера в сутки, в рублях	2124	1050,00	82795,00	9871,99	7705,60
Количество номеров в отеле	2068	1,00	3896,00	93,43	173,93
Количество отзывов в расчете на один номер	2068	0,06	193,50	11,61	13,63
Количество валидных наблюдений (в которых заполнены все показатели и отсутствуют пропущенные значения)	1509				

**Индекс воспринимаемой потребительской ценности.** На основе показателя рейтинга (среднего арифметического оценок потребителей) для каждого отеля в данной совокупности была рассчитана переменная индекса воспринимаемого качества, а на основе цены одноместного размещения в сутки - индекс воспринимаемой цены относительно конкурентов в каждом конкретном городе. Полученная модель карты ценности (соотношения между воспринимаемой ценой и воспринимаемым качеством, см. рис. 1) позволяет говорить о группировке значений воспринима-

емой ценности преимущественно в рамках одного большого кластера отелей. Это свидетельствует об ограниченном представлении большинства потребителей о преимуществах гостиничных услуг или об отказе от их приобретения по определенной цене.

Для контроля качества модели был построен график распределения значений воспринимаемой ценности в рамках указанного набора данных (показывает частоту встречаемости каждого значения). Гипотеза о нормальности распределения была проверена с использованием критерия Колмогорова-Смирнова. Та-

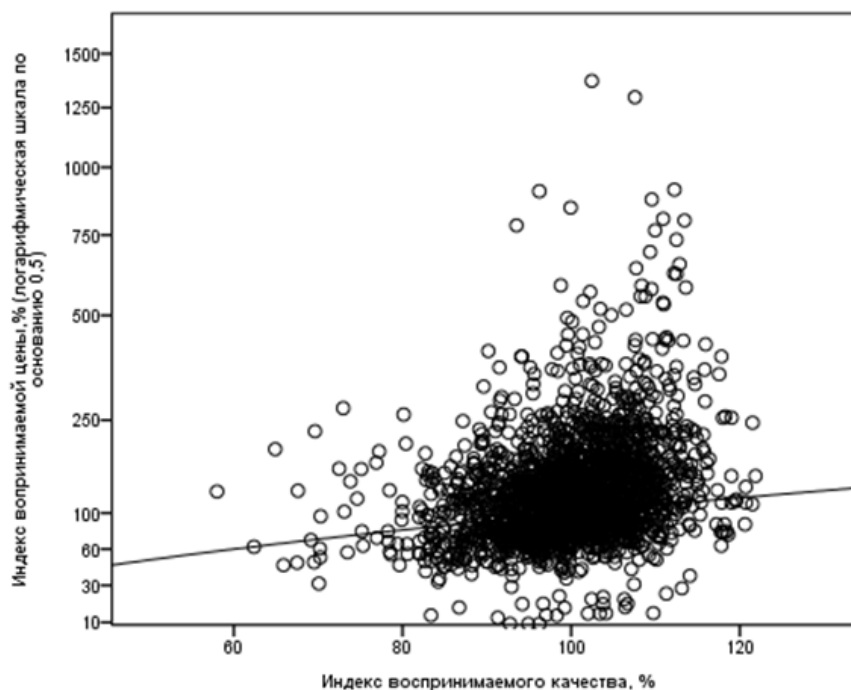


Рис. 1. Карта ценности

ким образом, можно сказать, что полученный в указанном наборе данных индекс воспринимаемой ценности является суммой многих случайных слабо взаимосвязанных величин.

Расчитав универсальные показатели индексов воспринимаемой цены и восприни-

маемого качества для каждого отеля, а также проверив жизнеспособность маркетинговой модели карты ценности на этом наборе данных, мы получили возможность использования этих факторов в регрессионном анализе потребительских предпочтений.

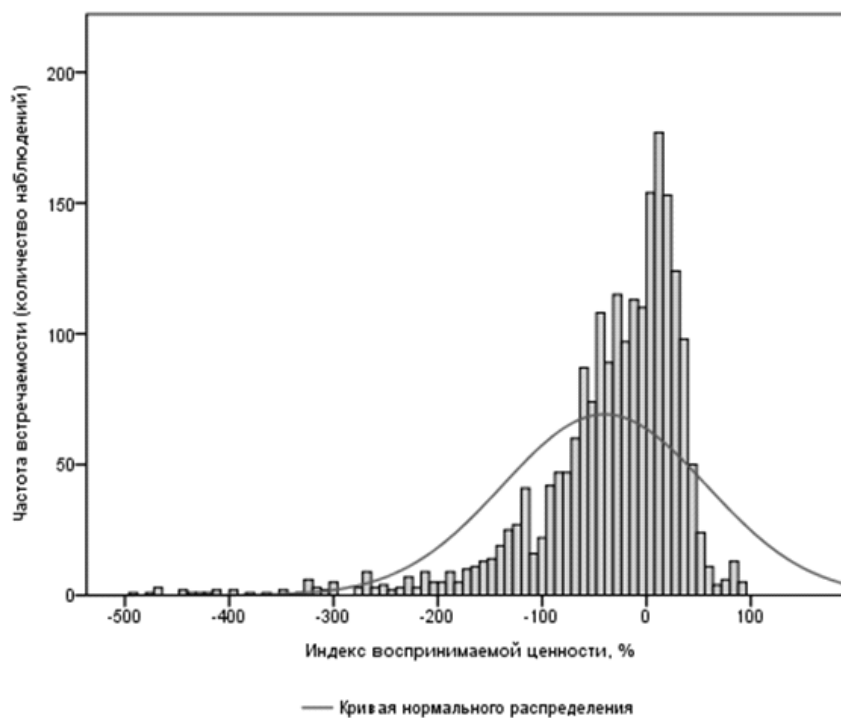


Рис. 2. График распределения значений воспринимаемой ценности

В рамках статистического исследования предпочтений потребителей была проверена гипотеза о влиянии на продажи гостиничных услуг через системы онлайн-бронирования следующих независимых переменных: количество звезд, бренд или принадлежность отеля к сети (принимает значение 1 в случае сетевого отеля, в остальных случаях показатель равен 0), индекс воспринимаемой цены, индекс воспринимаемого качества. Для целей исследования косвенным показателем продаж гостиничных услуг через данную систему онлайн-бронирования является рассчитанное нами отношение количества отзывов об отеле к количеству номеров в нем.

Для проведения множественной регрессии был использован пошаговый метод отбора переменных (пошаговое исключение из модели статистически незначимых переменных). Предварительно была осуществлена диагно-

стика коллинеарности с проверкой фактора инфляции дисперсии VIF (Variance Inflation Factor). Полученные значения этого фактора позволяют утверждать, что в модели отсутствует мультиколлинеарность (значения VIF меньше 10). Для контроля качества полученной модели была подтверждена гипотеза о нормальном распределении остатков, тест Дарбина-Ватсона дал удовлетворительное значение коэффициента, равное 1,940. Все это подтверждает, что выбранная регрессионная модель хорошо описывает истинную зависимость. Полученное уравнение регрессии для прогнозирования продаж на основе влияния потребительских предпочтений выглядит следующим образом:

$$\text{Продажи} = 14,03 - 4,62 \times (\text{кол-во звезд}) - 2,60 \times (\text{принадлежность к сети}) + 0,15 \times (\text{индекс воспринимаемого качества}) - 0,01 \times (\text{индекс воспринимаемой цены}).$$

Таблица 2

Коэффициенты модели, рассчитанные в результате выполнения регрессионного анализа<sup>1</sup>

Модель	Нестандартизованные коэффициенты		Стандартизованные коэффициенты	t-статистика <sup>4</sup>	Значимость <sup>5</sup>	Статистика коллинеарности		
	B (коэффициент регрессии <sup>2</sup> )	стандартная ошибка	бета-коэффициент <sup>3</sup>			толерантность (допуск) <sup>6</sup>	VIF (фактор инфляции дисперсии)	
1	(Константа)	28,74	1,12		25,70	0,00		
	Количество звезд	-5,22	0,31	-0,40	-16,81	0,00	1,00	1,00
2	(Константа)	27,71	1,14		24,27	0,00		
	Количество звезд	-4,70	0,34	-0,36	-14,00	0,00	0,85	1,18
	Принадлежность к сети	-2,65	0,66	-0,10	-4,05	0,00	0,85	1,18
3	(Константа)	16,68	3,60		4,63	0,00		
	Количество звезд	-5,06	0,35	-0,39	-14,34	0,00	0,76	1,31
	Принадлежность к сети	-2,60	0,65	-0,10	-3,97	0,00	0,85	1,18
	Индекс воспринимаемого качества	0,12	0,04	0,08	3,22	0,00	0,89	1,13
4	(Константа)	14,03	3,66		3,83	0,00		
	Количество звезд	-4,62	0,37	-0,35	-12,46	0,00	0,68	1,46
	Принадлежность к сети	-2,60	0,65	-0,10	-4,00	0,00	0,85	1,18
	Индекс воспринимаемого качества	0,15	0,04	0,10	3,90	0,00	0,85	1,17
	Индекс воспринимаемой цены	-0,01	0,00	-0,10	-3,67	0,00	0,80	1,25

<sup>1</sup> Зависимая переменная: количество продаж (отношение количества отзывов к количеству номеров).

<sup>2</sup> Коэффициент уравнения регрессии показывает, на сколько единиц изменится результат при изменении фактора на одну единицу.

<sup>3</sup> Бета-коэффициент показывает, на какую часть величины своего среднего квадратичного отклонения изменится в среднем значение резуль- тативного признака при изменении факторного признака на величину его среднеквадратического отклонения при фиксированном на постоянном уровне значении остальных независимых переменных.

<sup>4</sup> Проверка существования статистической значимости между переменными.

<sup>5</sup> Наблюдаемая значимость статистик равна нулю, поэтому гипотеза о равенстве коэффициентов нулю отвергается для каждого коэффициента.

<sup>6</sup> Множественный коэффициент корреляции.



**Результаты и выводы.** В рамках этого исследования с использованием больших данных получено статистическое подтверждение того, что потребители предпочитают отели с более высокими отзывами в системах онлайн-бронирования. Результаты регрессионного анализа показывают, что влияние индекса воспринимаемой цены, количества звезд и принадлежности отеля к сети является статистически значимым и имеет отрицательную связь с интернет-продажами, а индекс воспринимаемого качества (построенный на основе рейтинга потребителей) - положительную связь. Необходимо еще раз упомянуть о важном ограничении: в отсутствие открытых данных о продажах в качестве косвенного показателя онлайн-продаж гостиничных услуг нами было взято отношение количества отзывов об отеле к количеству номеров в нем.

В целом для клиентов, бронирующих отель через онлайн-агрегатор, предоставляющий выбор из большого количества отелей и демонстрирующий отзывы других путешественников, рейтинг других пользователей служит более важным сигналом ценности (качества услуг), чем количество звезд, бренд отеля или цена номера. В исследуемом наборе данных сетевые отели в целом предоставляют клиентам меньшую воспринимаемую ценность, нежели несетевые отели. При этом группа сетевых отелей в среднем демонстрирует более высокий средний рейтинг потребителей, нежели несетевые. Однако более высокий рейтинг не оправдывается существенно более высокой ценой.

Выявлено преимущественное группирование отелей на карте ценности в рамках одного кластера, что говорит об ограниченном представлении большинства потребителей о преимуществах гостиничных услуг или об отказе от их приобретения по определенной цене.

Наиболее важным результатом статистического исследования потребительских предпочтений в электронной коммерции (на примере гостиничного рынка) представляется статистическое подтверждение гипотезы о том, что доступность оценок других покупателей с помощью информационных техно-

логий позволяет потребителям более точно оценить и прогнозировать для себя ценность данного продукта, чем ориентируясь на другие сигналы качества, такие, как цена и бренд. Следовательно, бронирование гостиниц через онлайн-агрегаторы будет вести к большему удовлетворению потребителя, чем бронирование по другим каналам сбыта гостиничных услуг.

Исследование имеет практическое значение для маркетинга гостиничных услуг. Также мы предполагаем, что предложенные в данной модели универсальные показатели индекса воспринимаемого качества и индекса воспринимаемой цены могут быть использованы в регрессионных моделях оценки влияния отзывов на потребительский выбор в электронной коммерции для различных отраслей.

### Литература

1. **Duan W., Gu B., Whinston A.B.** Do online reviews matter? - An empirical investigation of panel data // *Decision Support Systems*. 2008. 45(4). 1007-1016.
2. **Forman C., Ghose A., Wiesenfeld B.** Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets // *Information Systems Research*. 2008. 19(3). 292-313.
3. **Gale B.** *Managing customer value: Creating quality and service that customers can see.* New York: Simon & Schuster, 1994.
4. 2015 Global Retail E-Commerce Index by A.T. Kearney. URL: [http://www.atkearney.com/consumer-products-retail/e-commerce-index/full-report/-/asset\\_publisher/87xbENNHPZ3D/content/global-retail-e-commerce-keeps-on-clicking/10192?\\_101\\_INSTANCE\\_87xbENNHPZ3D\\_redirect=%2Fconsumer-products-retail%2Fe-commerce-index](http://www.atkearney.com/consumer-products-retail/e-commerce-index/full-report/-/asset_publisher/87xbENNHPZ3D/content/global-retail-e-commerce-keeps-on-clicking/10192?_101_INSTANCE_87xbENNHPZ3D_redirect=%2Fconsumer-products-retail%2Fe-commerce-index).
5. **Lynch C.** Big data: How do your data grow? // *Nature*. 2008. Vol. 455, No. 7209.
6. **Ogut H., Onur Tas B.K.** The influence of internet customer reviews on the online sales and prices in hotel industry // *The Service Industries Journal*. 2012. Vol. 32. No. 1, 2.
7. **Волков Д.Л.** Управление ценностью: показатели и модели оценки // *Российский журнал менеджмента*. 2005. Т. 3. № 4. С. 67-76.
8. **Марн М., Регнер Э., Завада К.** *Ценовое преимущество.* - М.: Альбина Бизнес Букс, 2004. - 317 с.
9. **Третьяк О.А.** Ценность клиента в течение его жизненного цикла: развитие одной из ключевых идей маркетинга взаимоотношений // *Российский журнал менеджмента*. 2011. № 3. С. 55-68.

## STATISTICAL ANALYSIS OF CONSUMER PREFERENCES IN ELECTRONIC COMMERCE

Olga Nikitina

*Author affiliation:* National Research University Higher School of Economics (Moscow, Russia). E-mail: o.nikitina@rospravo.ru.

*This article addresses the issues concerning statistical analysis of consumer preferences in e-commerce (on the example of the hotel industry). The author conducted multiple regression analysis to assess the impact of various factors on consumer choice and the internet-sales in the industry. Using value-based marketing approach the article proposes the method for calculating the index of perceived customer value (value maps) in e-commerce. The data array used in an article for the statistical analysis composes of about 1.5 million observations (customer feedback on the quality of hotel services in 13 most popular cities for tourism by Euromonitor International ranking).*

*The most crucial result of studying consumer preferences in e-commerce (on the example of hotel market) appears to confirm the hypothesis that the availability of other customers' evaluations using information technology allows users to more accurately predict the value of these services for themselves rather than by focusing on the other quality signals, such as hotel's star rating and brand (being part of a hotel chain). Therefore, making hotel reservations through the online system will lead to greater customer satisfaction compared with other channels for selling hotel services.*

*Ключевые слова:* consumer preferences, perceived customer value, methods of product quality measurement, statistical analysis of large data arrays, electronic commerce (e-commerce), online tourism.

*JEL:* C35, D46, L83, M31.

### References

1. **Duan W., Gu B., Whinston A.B.** Do online reviews matter? - An empirical investigation of panel data. *Decision Support Systems*, 2008, 45(4), 1007-1016.
2. **Forman C., Ghose A., Wiesenfeld B.** Examining the relationship between reviews and sales: The role of reviewer identity discloser in electronic markets. *Information Systems Research*, 2008, 19(3), 292-313.
3. **Gale B.** *Managing customer value: Creating quality and service that customers can see.* New York: Simon & Schuster, 1994.
4. 2015 Global Retail E-Commerce Index by A.T. Kearney. Available at: [http://www.atkearney.com/consumer-products-retail/e-commerce-index/full-report/-/asset\\_publisher/87xbENNHPZ3D/content/global-retail-e-commerce-keeps-on-clicking/10192?\\_101\\_INSTANCE\\_87xbENNHPZ\\_3D\\_redirect=%2Fconsumer-products-retail%2Fe-commerce-index](http://www.atkearney.com/consumer-products-retail/e-commerce-index/full-report/-/asset_publisher/87xbENNHPZ3D/content/global-retail-e-commerce-keeps-on-clicking/10192?_101_INSTANCE_87xbENNHPZ_3D_redirect=%2Fconsumer-products-retail%2Fe-commerce-index).
5. **Lynch C.** Big data: How do your data grow? *Nature*, 2008, vol. 455, no. 7209.
6. **Ogut H., Onur Tas B.K.** The influence of internet customer reviews on the online sales and prices in hotel industry. *The Service Industries Journal*, 2012, vol. 32, no. 1, 2.
7. **Volkov D.L.** Upravleniye tsennost'yu: pokazateli i modeli otsenki [Value management: indicators and assessment model]. *Rossiyskiy zhurnal menedzhmenta*, 2005, vol. 3, no. 4, pp. 67-76. (In Russ.).
8. **Marn M., Regner E., Zavada K.** *Tsenovoye preimushchestvo* [Price advantage]. Moscow, Al'bina Biznes Buks Publ., 2004. 317 p. (In Russ.).
9. **Tretiak O.A.** Tsennost' klyienta v techeniye yego zhiznennogo tsikla: razvitiye odnoy iz klyuchevykh idey marketinga vzaimootnosheniy [Customer value throughout the life cycle: development of one of the key ideas of relationship marketing]. *Rossiyskiy zhurnal menedzhmenta*, 2011, no. 3, pp. 55-68. (In Russ.).