

УДК 519.2

**Ж.Н. Зенкова, И.В. Краковецкая****НЕПАРАМЕТРИЧЕСКАЯ ОЦЕНКА ТЁРНБУЛЛА  
ДЛЯ ИНТЕРВАЛЬНО-ЦЕНЗУРИРОВАННЫХ ДАННЫХ  
В МАРКЕТИНГОВОМ ИССЛЕДОВАНИИ СПРОСА  
НА БИОЭНЕРГЕТИЧЕСКИЕ НАПИТКИ<sup>1</sup>**

С помощью непараметрического алгоритма Тёрнбулла по интервально-цензурированным данным оценивается ожидаемая цена спроса на биоэнергетический напиток, выводимый на рынок томским предприятием НПК «САВА». Диапазоны цен для анализа получены в результате масштабного маркетингового исследования, включающего в себя несколько фокус-групп и опросов. Работа носит междисциплинарный прикладной характер.

**Ключевые слова:** *интервальное цензурирование, непараметрический алгоритм Тёрнбулла, маркетинговые исследования, ценообразование.*

В процессе вывода нового продукта на рынок очень важно определить ту цену, с которой производитель готов выйти на массового потребителя. При этом возникает ряд проблем, связанных с тем, что он не обладает достаточной информацией о возможном уровне цен, так как предприятие не имеет накопленной статистики по продажам, что вынуждает маркетологов пользоваться так называемым историческим методом, когда в основу анализа ложатся данные о продажах схожего по свойству товара, уже выпускаемого фирмой. Нередко предприятие пытается организовать производство принципиально нового для него продукта, при этом основным, а зачастую и единственным ориентиром, является себестоимость конечного изделия. Однако этой информации явно недостаточно для определения рыночной цены. Тогда производитель вынужден прибегать к дорогостоящим инструментам маркетинговых исследований, таких, как фокус-группы, опросы, анкетирование потенциальной целевой группы потребителей и пр.

Полученные в итоге данные носят стохастический характер и требуют качественной статистической обработки. Особенно в случаях, когда наблюдения являются цензурированными, при этом использование классических методов приводит к неадекватным некорректным результатам.

В данной работе рассматривается вопрос определения цены потребительского спроса на товар-новинку, в основу исследования положены интервальные данные об ощущаемой ценности товара для потребителя, полученные в процессе многоэтапного маркетингового исследования. По результатам анализа фирме даны рекомендации о возможном уровне цены.

**1. Непараметрическая оценка Тёрнбулла  
для интервально-цензурированных данных**

Неполные данные часто встречаются в медицинских, биологических, социальных и экономических исследованиях, в страховании, а также в инженерных задачах,

<sup>1</sup> Работа выполнена в рамках государственного задания Минобрнауки РФ на проведение научных исследований в Томском государственном университете на 2012-2014 годы, задание 8.4055.2011.

например при анализе надежности и времени отказа технических устройств [1–6]. Здесь рассматриваются интервально-цензурированные данные, возникающие при проведении маркетинговых исследований, в частности в процессе работы фокус-групп, при опросах и анкетировании, когда опрашиваемый не может или не хочет указывать конкретное значение исследуемого показателя, например своего дохода, а оперирует лишь некоторым промежутком. При этом нередко маркетологи обрабатывают полученные цензурированные выборки некорректно, игнорируя цензурированные значения или ориентируясь на одну из границ интервала, что в итоге приводит к неадекватным результатам, которые используются как основа при принятии управленческих решений, чреватых финансовыми потерями.

Рассмотрим случайную величину (с.в.)  $\tau > 0$  с функцией распределения (ф.р.)  $F(x) = P(\tau \leq x)$ . Исходная выборка объема  $N$  состоит из интервалов вида  $(L_i, U_i]$ ,  $i = \overline{1, N}$ , т.е. известно, что  $i$ -е наблюдение попало в промежуток  $(L_i, U_i]$ , но его точное значение неизвестно. Для оценивания ф.р.  $F(x)$  рассмотрим разбиение  $0 < \tau_0 < \tau_1 < \dots < \tau_m$ , состоящее из всех неповторяющихся упорядоченных границ интервалов  $L_i$  и  $U_i$ ,  $i = \overline{1, N}$ . Заметим, что  $m \leq 2N - 1$ . При этом  $m = 2N - 1$ , если все левые и правые границы интервалов не совпадают друг с другом.

Для  $j = \overline{1, m}$  и  $i = \overline{1, N}$  определяем вес

$$\alpha_{ij} = \begin{cases} 1, & \text{если } (\tau_{j-1}, \tau_j] \subseteq (L_i, U_i], \\ 0, & \text{если } (\tau_{j-1}, \tau_j] \not\subseteq (L_i, U_i]. \end{cases}$$

Алгоритм Тёрнбулла заключается в следующем [1]: для произвольного шага алгоритма  $k \geq 0$

1. Вычисляем вероятность попадания с.в.  $\tau$  в интервал  $(\tau_{j-1}, \tau_j]$  по формуле

$$p_j^k = S^{(k)}(\tau_{j-1}) - S^{(k)}(\tau_j), \quad j = \overline{1, m}.$$

2. Находим значение

$$d_j^k = \sum_{i=1}^N \left( \alpha_{ij} p_j^k / \sum_{s=1}^m \alpha_{is} p_s^k \right), \quad j = \overline{1, m}.$$

3. Для  $j = \overline{1, m}$  определяем

$$Y_j^k = \sum_{s=j}^m d_s^k.$$

4. Производим перерасчет значения вероятности  $S^{(k+1)}(\tau_j)$  по формуле Каплана – Мейера [2]

$$S^{k+1}(\tau_j) = \prod_{i=1}^j \frac{Y_i^k - d_i^k}{Y_i^k}, \quad j = \overline{1, m},$$

при этом  $S^{k+1}(\tau_0) = 1$ .

Алгоритм повторяем до тех пор, пока для всех  $j = \overline{1, m}$  не выполнится условие

$$|S^{(k+1)}(\tau_j) - S^{(k)}(\tau_j)| \leq 10^{-7}.$$

Начальное значение  $p_j^0$ ,  $j = \overline{1, m}$ , определим как

$$p_j^0 = \frac{1}{m}.$$

В итоге значение ф.р.  $F(x)$  в точке  $\tau_j$ ,  $j = \overline{0, m}$ , можно оценить как

$$F_N(\tau_j) = 1 - S^{(k+1)}(\tau_j).$$

Таким образом, для  $x \in (\tau_{j-1}, \tau_j]$ ,  $j = \overline{1, m}$ ,

$$F_N(x) = F_N(\tau_j), \quad (1)$$

для  $x \in [0, \tau_0]$   $F_N(x) = 0$ .

Полученная оценка ф.р. является непараметрической и равномерно строго состоятельной [7], т.е.

$$P\left\{\sup_{x \in (0, +\infty)} |F_N(x) - F(x)| \xrightarrow{N \rightarrow +\infty} 0\right\} = 1.$$

Найдем оценку среднего значения исследуемого показателя

$$E\tau = \int_0^{+\infty} x dF(x) \quad (2)$$

методом подстановки в интеграл (2) оценки (1). Получим

$$\hat{E}\tau = \int_0^{+\infty} x dF_N(x) = \sum_{i=1}^m \tau_i (F_N(\tau_i) - F_N(\tau_{i-1})). \quad (3)$$

Оценка (3) является асимптотически несмещенной.

## 2. Исследование рынка биоэнергетических напитков и проблема коммерциализации нового продукта НПК «САВА»

В 2010 г. Сибирский ботанический сад Томского государственного университета (ТГУ) и компания ТПК «САВА» создали малое инновационное предприятие «Научно-производственная компания (НПК) "САВА"». Основной целью компании является вывод на рынок нового натурального биоэнергетического напитка на основе фитоадаптогенов, которые относятся к конкурентоспособным продуктам питания с высокой пищевой и биологической ценностью [8].

В ходе маркетинговых исследований выявлен портрет потенциальных потребителей – это молодые люди, в основном мужчины (80 %) в возрасте 16–35 лет, жители средних и больших городов, ведущие активный образ жизни, занятые в различных сферах деятельности: менеджеры, специалисты отраслей, требующих повышенной концентрации внимания, спортсмены, автомобилисты и студенты.

Несмотря на рост рынка энергетиков в целом [9], рынок натуральных энергетических напитков в России в настоящее время находится в стадии становления, поэтому на нем можно наблюдать такой большой разброс цен – от 33 до 1600 руб. за литр.

Заметим, при установлении цены на новые товары существует определенный диапазон цен, внутри которого может действовать производитель [6]. Нижняя его граница определяется себестоимостью товара, а верхняя – платежеспособным

спросом. Один из простейших методов ценообразования на новые товары – предложение товара потребителям «на пробу» с последующим опросом их в отношении приемлемой цены.

### 3. Применение алгоритма Тёрнбулла для исследования цены потребительского спроса на биоэнергетический напиток НПК «САВА»

В процессе фокус-группы было проведено анкетирование потенциальных потребителей продукции НПК «САВА». Участникам предлагалось указать диапазон для цены, которую они готовы заплатить за биоэнергетический напиток. В результате были получены данные по ценовым предпочтениям. Исходные интервалы представлены в табл. 1. С помощью алгоритма Тёрнбулла были вычислены вероятности  $p_j$ . Полученные значения  $p_j$  и соответствующие им интервалы  $(\tau_{j-1}, \tau_j]$  представлены в табл. 2.

Таблица 1

Исходные интервальные данные по ценам

№	Левая граница цены $L_i$ , руб./шт.	Правая граница цены $U_i$ , руб./шт.	№	Левая граница цены $L_i$ , руб./шт.	Правая граница цены $U_i$ , руб./шт.
1	18	22	14	48	52
2	25	29	15	48	52
3	48	52	16	58	62
4	28	32	17	48	52
5	23	27	18	43	47
6	29	33	19	28	32
7	28	32	20	48	52
8	18	22	21	58	62
9	28	32	22	48	52
10	25	29	23	38	42
11	48	52	24	48	52
12	16	20	25	38	42
13	28	32	26	46	50

Таблица 2

Значения вероятностей и соответствующие им временные интервалы

$\tau_{j-1}$	$\tau_j$	$p_j$	$F_N(\tau_j)$
18	20	0,085	0,085
20	22	0,021	0,106
23	25	0,01	0,116
25	27	0,052	0,168
28	29	0,173	0,341
29	32	0,089	0,430
38	42	0,077	0,507
43	46	0,012	0,519
46	47	0,036	0,555
48	50	0,251	0,806
50	52	0,067	0,873
58	62	0,077	0,950

График оценки ф.р.  $F(x)$  приведен на рис. 1.

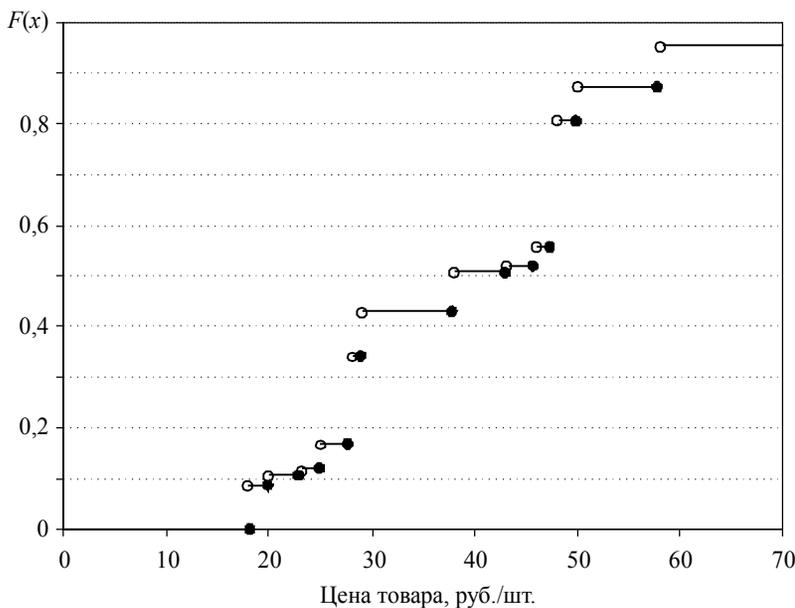


Рис. 1. График оценки ф.р.  $F(x)$

Используя полученную оценку ф.р., по формуле (3) найдем выборочную среднюю цену напитка  $\hat{E}\tau = 36,87$  руб./шт.

### Заключение

В результате проделанной работы получена оценка ощущаемой ценности биоэнергетического напитка НПК «САВА» потенциальными потребителями с помощью непараметрического алгоритма Тёрнбулла на основе интервально-цензурированных данных. Зная, что в процессе подобных исследований цена неизбежно занижается, производитель должен ориентироваться не столько на себестоимость товара, сколько на полученное значение потребительской цены, скорректированной с учетом рыночного коэффициента  $k > 1$ , при этом значение коэффициента выявляется в ходе дальнейшего исследования рынка с помощью пробного маркетинга [11]. Таким образом, в работе была получена минимальная рекомендованная рыночная цена нового биоэнергетического напитка НПК «САВА».

Заметим, что российские маркетологи зачастую ошибочно применяют статистический инструментарий, в итоге это приводит к искаженному пониманию многих экономических и социальных процессов и, следовательно, к финансовым потерям предприятия за счет «плохих» управленческих решений. Рассмотренный подход может устранить этот пробел в отношении оценивания многих показателей, базирующихся на интервально-цензурированных данных, и найти широкое применение в бизнес-практике, позволяя адекватно оценить не только предпочитаемую потребителями цену, но и другие важные количественные характеристики, используемые в маркетинге.

## ЛИТЕРАТУРА

1. *Giolo S.R., Turnbull's N.* Estimator for interval-censored data // Technical Report, August, 2004.
2. *Klein J.P., Moeschberger M.L.* Survival Analysis: Techniques for Censored and Truncated Data. Springer, 2010. 551 p.
3. Анализ надежности технических систем по цензурированным выборкам / В.М. Скрипник, А.Е. Назин, Ю.Г. Приходько, Ю.Н. Благовещенский. М.: Радио и связь, 1988. 184 с.
4. *Zhao, Guolin M.A.* Nonparametric and parametric survival analysis of censored data with possible violation of method assumptions // Thesis Submitted to the Faculty of the Graduate School at the University of North Carolina at Greensboro in Partial Fulfillment of the Requirements for the Degree Master of Arts. Greensboro, 2008. 55 p.
5. *Franses Ph.H., Paap R.* Quantitative Models in Marketing Research. Cambridge University Press, 2001. 206 p.
6. *Blattberg R.C., Kim P., Kim B., Neslin S.A.* Database Marketing: Analyzing and Managing Customers. Springer, 2008. 871 p.
7. *Chang M.N., Yang G.L.* Strong consistency of a nonparametric estimator of the survival function with doubly censored data // The Annals of Statistics. 1987. V. 15. No. 4. P. 1536–1547.
8. НПК «Сава» готовится к производству натуральных биоэнергетических напитков. <http://incubator.tsu.ru/taxonomy/term/1?page=1> [режим доступа – свободный]
9. Анализ рынка энергетических напитков в России в 2005–2010 гг., прогноз на 2011–2014 гг. Отчет о результатах маркетингового исследования компании BusinesStat. <http://marketing.rbc.ru> [режим доступа – свободный]
10. *Маренков Н.* Особенности установления цен на новые товары. <http://www.inventech.ru/lib/cost/cost-0014/> [режим доступа – свободный]
11. *Панкрухин А.П.* Маркетинг: учебник. М.: ИКФ Омега-Л, 2002. С. 272–273.

*Зенкова Жанна Николаевна*

*Краковецкая Инна Валентиновна*

Томский государственный университет,

Томский политехнический университет

E-mail: [thankoff@fpmk.tsu.ru](mailto:thankoff@fpmk.tsu.ru); [inna\\_krakov@mail.ru](mailto:inna_krakov@mail.ru)

Поступила в редакцию 5 мая 2012 г.

*Zenkova Zh.anna N., Krakovetckaia Inna V.* (Tomsk State University, Tomsk Polytechnic University). **Nonparametric Turnbull estimator for interval-censored data in the marketing research of the demand of bio-energy drinks.**

Keywords: interval-censored data, the nonparametric Turnbull algorithm, marketing research, price formation

Using the nonparametric Turnbull algorithm for interval-censored data there is estimated the expected price of the demand for bio-energy drink that will be commercialized by Tomsk's firm «SAVA». The price ranges for the analysis are the result of large-scale marketing research, which includes several focus groups and surveys. The lower bound of recommended price of the new product has been obtained.