



# Некоторые возможности использования эконометрических методов в исследованиях сферы туризма

*Лебедева С.А.*<sup>1</sup>, *Шершуков Д.В.*<sup>2</sup>

<sup>1</sup> Российская академия народного хозяйства и государственной службы при Президенте РФ, Москва, Россия

<sup>2</sup> Национальный исследовательский университет «Высшая школа экономики», Москва, Россия

## АННОТАЦИЯ:

В данной статье с помощью метода бинарной логистической регрессии были выявлены основные характеристики российских гостротуристов. Гастрономический туризм – новое модное направление, динамично развивающееся во всем мире, в том числе на территории Российской Федерации. Однако для создания конкурентоспособного предложения на данном рынке необходимо определить характеристики потенциальной аудитории. Именно по этой причине данная работа представляет особую ценность. Определение основных демографических и психографических характеристик российских гостротуристов будет способствовать созданию конкурентного туристского продукта на российском рынке, а также развитию внутреннего туризма на территории страны, что крайне актуально в условиях сложившейся политико-экономической ситуации. Кроме того, в работе проведен корреляционный анализ различных видов туризма, что также важно для развития туристской индустрии на территории России.

**КЛЮЧЕВЫЕ СЛОВА:** регрессионный анализ, бинарная логистическая регрессия, корреляционный анализ, туризм, гастрономический туризм, характеристики российских гостротуристов.

## Some possibilities of using econometric methods in tourism research

*Lebedeva S.A.*<sup>1</sup>, *Shershukov D.V.*<sup>2</sup>

<sup>1</sup> The Russian Presidential Academy of National Economy and Public Administration, Russia

<sup>2</sup> The National Research University Higher School of Economics (HSE), Russia

## Введение

Большая часть современных методов анализа различных областей экономики опирается на эконометрические модели и концепции. Без глубоких знаний в области эконометрики невозможно научиться их использовать. Для практически значимых исследований сферы туризма также важно иметь хорошую эконометрическую подготовку.

Гастрономический туризм как динамично развивающийся сегмент туристической индустрии и его развитие на российском рынке представляет интерес с разных позиций и, прежде всего, с точки зрения возможных положительных эффектов для развития национальной экономики. Для

изучения одного из аспектов нового вида туризма использовался регрессионный анализ. Помимо этого в работе представлено использование корреляционного анализа в туристской отрасли. Оба анализа выполнены на основании проведенного авторами опроса с помощью специализированной программы для статистических вычислений и графического анализа данных «R» [4] (Kabakov, 2014).

Регрессионный анализ – это метод установления формы и изучения связей между метрически зависимой переменной и одной или несколькими независимыми переменными [3, с. 779] (Malkhotra, 2016). Данный метод дает возможность для прогнозирования значения зависимой переменной, отталкиваясь от значения независимой переменной. Бинарная логистическая регрессия, использованная в данной работе, в свою очередь, позволяет исследовать зависимость дихотомических переменных от независимых переменных, имеющих любой вид шкалы. Таким образом, с помощью метода бинарной логистической регрессии можно рассчитать вероятность наступления события в зависимости от значений независимых переменных. Значение вероятности ( $P(Y)$ ) может принимать значения между 0 и 1. Если для  $P(Y)$  получится значение меньше 0,5, то можно предположить, что событие не наступит. В противном случае предполагается наступление события.

Коэффициенты корреляции используются для описания связей между переменными. Знак коэффициента свидетельствует о направлении связи, а величина коэффициента показывает силу этой связи. Соответственно, чем ближе модуль коэффициента к 1, тем сильнее связь между исследуемыми объектами.

#### ABSTRACT:

In this article, using the method of binary logistic regression, authors identified the main characteristics of the Russian gastrotourism. Gastronomic tourism is a new fashionable trend, dynamically developing all over the world, including in the territory of the Russian Federation. However, in order to create a competitive offer in this market, it is necessary to determine the characteristics of the potential audience. The definition of the main demographic and psychographic characteristics of Russian gastrotourists will contribute to the creation of a competitive tourist product in the Russian market, as well as the development of domestic tourism in the country, which is extremely important in the current political and economic situation. In addition, the paper conducted a correlation analysis of different types of tourism, which is also important for the development of the tourism industry in Russia.

**KEYWORDS:** regression analysis, binary logistic regression, correlation analysis, tourism, gastronomic tourism, characteristics of the Russian gastrotourism

JEL Classification: L83, C15, C20

Received: 20.06.2018 / Published: 31.08.2018

© Author(s) / Publication: CREATIVE ECONOMY Publishers

For correspondence: Lebedeva S.A. (lebedeva.svetlana1994@gmail.com)

#### CITATION:

Lebedeva S.A., Shershukov D.V. [2018] Nekotorye vozmozhnosti ispolzovaniya ekonometricheskikh metodov v issledovaniyakh sfery turizma [Some possibilities of using econometric methods in tourism research]. Kreativnaya ekonomika. 12. (8). – 1221-1238. doi: 10.18334/ce.12.8.39289

Логистическая регрессия с бинарными переменными использовалась для создания портрета типичного российского гастрономического туриста на базе иностранных исследований. В качестве зависимой переменной принимался ответ на вопрос анкеты, поехал ли бы респондент в гастрономическое путешествие или нет. Соответственно, если ответ на данный вопрос был положительным, то зависимая переменная принимала значение 1, если ответ был отрицательным – то 0.

В качестве независимых переменных проверялись такие характеристики, как половая принадлежность, возраст, уровень образования и дохода, семейное положение респондентов, наличие детей, принадлежность туристов к одной из категорий, описывающих стиль жизни, и некоторые психографические характеристики – психоцентрики или аллоцентрики, а также отношение респондентов к гастрономии в целом. Данные характеристики взяты из международных исследований, определяющих их как значимые или незначимые. Так, исследование Всемирной ассоциации гастрономического туризма, проведенное в 2010 году, говорит о неважности таких факторов, как пол, возраст и уровень дохода [10]. В то же время исследователи из Гвельфского университета в Канаде выяснили, что незамужняя женщина в случае увеличения доходов с большей вероятностью увеличит свои расходы на рестораны, нежели холостой мужчина или женатый/замужний человек [1, с. 308] (*Kotler, Bouen, Meykenz, 2013*).

Американская исследовательница Елена Уильямс (Helena A. Williams) из Техасского технического университета со своими коллегами Робертом Уильямсом (Robert L. Williams Jr.) из университета Сускуэханна и Мактобой Омар (Maktoba Omar) из университета Конвентри также приводят классификацию гастрономических туристов, основываясь именно на демографических характеристиках. Согласно этому же исследованию, а также работе румынских ученых, гастротуристы – это обычно люди с высоким уровнем образования [7, с. 6; 8, с. 399] (*Williams et al., 2014; Boita, Anghelina, Costi, 2014*).

Психографические характеристики гастротуристов были приведены словацкими учеными, которые отмечают, что обычно данная категория туристов отвергает единообразия и отличается насыщенным стилем жизни [6, с. 16] (*Gheorghe, Tudorache, Nistoreanu, 2014*). Кроме всего прочего, гастротуристы имеют особое отношение к еде, напиткам и в целом к гастрономии и кулинарии. Они рассматривают гастрономию как средство коммуникации, социализации и платформы для обмена опытом. Этот

#### **ОБ АВТОРАХ:**

*Лебедева Светлана Алексеевна*, магистрант факультета менеджмента спортивной и туристской индустрии института отраслевого менеджмента (lebedeva.svetlana1994@gmail.com)

*Шершуков Данила Вадимович*, магистрант факультета экономических наук (danila.shershukov@yandex.ru)

#### **ЦИТИРОВАТЬ СТАТЬЮ:**

Лебедева С.А., Шершуков Д.В. Некоторые возможности использования эконометрических методов в исследованиях сферы туризма // Креативная экономика. – 2018. – Том 12. – № 8. – С. 1221-1238. doi: 10.18334/ce.12.8.39289

тип туристов тщательно выбирает продукты питания: их волнует все, начиная с места происхождения продуктов до способа их приготовления. Они ориентированы на открытия, постоянно стремятся открыть что-то новое, активно интересуются музеями, магазинами, различными шоу, мюзиклами, кинофестивалями и культурными достопримечательностями [5, с. 6] (Irigüler, Güler, 2017).

Каждая из описанных выше характеристик является качественной. Для трансформации данных переменных в количественные характеристики была использована процедура переопределения. Так, для обозначения лиц женского пола, респондентов, не состоящих в браке, а также опрашиваемых, не имеющих детей, использовался 0, для лиц мужского пола, интервьюируемых, состоящих в браке, и респондентов, имеющих детей, – 1.

Переменные возраста, образования и дохода состоят из нескольких категорий, где за 1 принимается группа, наилучшим образом описывающая ту или иную характеристику участника опроса, а за 0 все остальные категории. Переменная возраста включает в себя категории: младше 16, от 16 до 20, от 21 до 25, от 26 до 30, от 31 до 35, от 36 до 40, от 41 до 45, от 46 до 50, от 51 до 55, от 56 до 60 и старше 61. Переменная образования включает следующие группы: *неполное среднее и полное среднее образование, среднее профессиональное образование, неоконченная степень бакалавра/специалиста, бакалавр/специалист, неоконченная степень магистра, магистр, два и более высших образования, ученая степень*.

Для конструирования вопроса об уровне дохода в исследовании использовался оценочный подход. Переменная дохода состоит из 6 категорий: *очень низкий уровень дохода, низкий, ниже среднего, средний, выше среднего, высокий уровень дохода*. Для каждой категории в анкете были добавлены комментарии, содержащие оценочные суждения об уровне благосостояния и позволяющие респондентам более точно определиться с группой, к которой они относятся.

Психометрическая характеристика «психоцентрик – аллоцентрик», предложенная американским социологом Стенли Плогом (Stanley Plog), – комплексная переменная, имеющая в своей основе пять прочих категорий: *склонность к переменам и степень консерватизма, отношение к риску, готовность к путешествиям, требовательность к комфорту во время путешествий и важность интеллектуальной составляющей во время поездки* [2, с. 325; 9] (Lebedeva, 2017; Stanley, 2001). Каждая из подкатегорий также является бинарной переменной, принимающей значения 0 или 1. Так, респонденты-неофобы получают 0 баллов, а респонденты-неофилы – 1; рискофобы – 0, рискофилы – 1; домоседы – 0, а любители путешествовать – 1; респонденты, требовательные и привыкшие к комфорту во время своих путешествий, – 0, в то время как лица, не нуждающиеся в комфортных условиях для отдыха – 1. Таким образом, психоцентрикам, набравшим от 0 до 2 баллов, присваивается 0, а аллоцентрикам (от 3 до 5 баллов) – 1 (рис. 1).



**Рисунок 1.** Классификация туристов по С. Плогу

Источник: дополнено авторами на основании [2, с. 325] (Lebedeva, 2017)

Отношение респондентов к еде рассматривалось в исследовании с нескольких ракурсов: с точки зрения потребления продуктов питания, с точки зрения приготовления блюд, а также получения новых знаний о продуктах питания и посещения всевозможных мероприятий, связанных с кулинарией и гастрономией. Опрашиваемым, благосклонно относящимся к гастрономии, присваивается 1, а негативно относящимся – 0.

Таким образом, опрос прошли 52 мужчины (37,1%) и 88 женщин (62,9%). Большинство респондентов относятся к возрастной категории от 21 до 25 (35,7%), следующая группа находится в диапазоне от 16 до 20 (25,7%). Большинство респондентов не состоят в браке (74,3%) и не имеют детей (74,3%). Большинство опрашиваемых имеют степень бакалавра или специалиста (28,6%). Чуть меньший процент респондентов еще не закончил обучение по бакалаврской программе (27,1%). Что касается уровня дохода, большая часть участников опроса отнесла свои домохозяйства к категориям со средним уровнем дохода (36,4%), ниже среднего (30,0%) и выше среднего (25,0%). 80 из 140 респондентов оказались аллоцентриками. Опрос также показал, что для 78,6% опрашиваемых гастрономия является важным и интересным фактором не только во время путешествий, но и в повседневной жизни.

После первичной обработки данных был получен набор 0 и 1. Даем команду для расчета описательной статистики исходных данных для исключения переменных с пропущенными значениями. Функция `summary()` позволяет вычислить основные параметры описательной статистики в программе R (рис. 2).

<b>gender</b>	<b>ageb16</b>	<b>age16t20</b>	<b>age21t25</b>	
Min. :0.0000	Min. :0.00000	Min. :0.0000	Min. :0.0000	
1st Qu.:0.0000	1st Qu.:0.00000	1st Qu.:0.0000	1st Qu.:0.0000	
Median :0.0000	Median :0.00000	Median :0.0000	Median :0.0000	
Mean :0.3714	Mean :0.02143	Mean :0.2571	Mean :0.3571	
3rd Qu.:1.0000	3rd Qu.:0.00000	3rd Qu.:1.0000	3rd Qu.:1.0000	
Max. :1.0000	Max. :1.00000	Max. :1.0000	Max. :1.0000	
<b>age26t30</b>	<b>age31t35</b>	<b>age36t40</b>	<b>age41t45</b>	
Min. :0.00000	Min. :0.00000	Min. :0.00000	Min. :0.00000	
1st Qu.:0.00000	1st Qu.:0.00000	1st Qu.:0.00000	1st Qu.:0.00000	
Median :0.00000	Median :0.00000	Median :0.00000	Median :0.00000	
Mean :0.07857	Mean :0.07857	Mean :0.04286	Mean :0.05714	
3rd Qu.:0.00000	3rd Qu.:0.00000	3rd Qu.:0.00000	3rd Qu.:0.00000	
Max. :1.00000	Max. :1.00000	Max. :1.00000	Max. :1.00000	
<b>age46t50</b>	<b>age51t55</b>	<b>age56t60</b>	<b>age61h</b>	
Min. :0.00000	Min. :0.00000	Min. :0.00000	Min. :0.00000	
1st Qu.:0.00000	1st Qu.:0.00000	1st Qu.:0.00000	1st Qu.:0.00000	
Median :0.00000	Median :0.00000	Median :0.00000	Median :0.00000	
Mean :0.01429	Mean :0.02857	Mean :0.02143	Mean :0.04286	
3rd Qu.:0.00000	3rd Qu.:0.00000	3rd Qu.:0.00000	3rd Qu.:0.00000	
Max. :1.00000	Max. :1.00000	Max. :1.00000	Max. :1.00000	
<b>school_i</b>	<b>school</b>	<b>college</b>	<b>bchlr_i</b>	
Min. :0.00000	Min. :0.00000	Min. :0.00000	Min. :0.0000	
1st Qu.:0.00000	1st Qu.:0.00000	1st Qu.:0.00000	1st Qu.:0.0000	
Median :0.00000	Median :0.00000	Median :0.00000	Median :0.0000	
Mean :0.02857	Mean :0.05714	Mean :0.05714	Mean :0.2786	
3rd Qu.:0.00000	3rd Qu.:0.00000	3rd Qu.:0.00000	3rd Qu.:1.0000	
Max. :1.00000	Max. :1.00000	Max. :1.00000	Max. :1.0000	
<b>bchlr</b>	<b>mstr_i</b>	<b>mstr</b>	<b>hedu2</b>	
Min. :0.0000	Min. :0.0000	Min. :0.00000	Min. :0.00000	
1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.00000	1st Qu.:0.00000	
Median :0.0000	Median :0.0000	Median :0.00000	Median :0.00000	
Mean :0.2857	Mean :0.1143	Mean :0.08571	Mean :0.05714	
3rd Qu.:1.0000	3rd Qu.:0.0000	3rd Qu.:0.00000	3rd Qu.:0.00000	
Max. :1.0000	Max. :1.0000	Max. :1.00000	Max. :1.00000	
<b>degree</b>	<b>incomel</b>	<b>income2</b>	<b>income3</b>	<b>income4</b>
Min. :0.00000	Min. :0	Min. :0.00	Min. :0.0000	Min. :0.0000
1st Qu.:0.00000	1st Qu.:0	1st Qu.:0.00	1st Qu.:0.0000	1st Qu.:0.00000
Median :0.00000	Median :0	Median :0.00	Median :0.0000	Median :0.0000
Mean :0.03571	Mean :0	Mean :0.05	Mean :0.3214	Mean :0.3357
3rd Qu.:0.00000	3rd Qu.:0	3rd Qu.:0.00	3rd Qu.:1.00000	3rd Qu.:1.0000
Max. :1.00000	Max. :0	Max. :1.00	Max. :1.0000	Max. :1.0000
<b>income5</b>	<b>income6</b>	<b>family</b>	<b>children</b>	
Min. :0.0000	Min. :0.00000	Min. :0.0000	Min. :0.0000	
1st Qu.:0.0000	1st Qu.:0.00000	1st Qu.:0.0000	1st Qu.:0.0000	
Median :0.0000	Median :0.00000	Median :0.0000	Median :0.0000	
Mean :0.2357	Mean :0.05714	Mean :0.2571	Mean :0.2571	
3rd Qu.:0.0000	3rd Qu.:0.00000	3rd Qu.:1.00000	3rd Qu.:1.0000	
Max. :1.0000	Max. :1.00000	Max. :1.0000	Max. :1.0000	
<b>ps_alo</b>	<b>food</b>	<b>questionY</b>		
Min. :0.0000	Min. :0.0000	Min. :0.00		
1st Qu.:0.0000	1st Qu.:1.0000	1st Qu.:0.75		
Median :1.0000	Median :1.0000	Median :1.00		
Mean :0.5714	Mean :0.7857	Mean :0.75		
3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:1.00		
Max. :1.0000	Max. :1.0000	Max. :1.00		

Рисунок 2. Описательные статистики для оцениваемых моделей

Источник: составлено авторами с помощью программы R

Расчет логистической регрессии производится командой *glm()*, а с помощью функции *summary()*, как говорилось ранее, вычисляем описательные статистики. Вычисленные описательные статистики логистической регрессии представлены на рисунке 3.

```

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.29181 -0.00002  0.00018  0.58305  1.70091

Coefficients: (4 not defined because of singularities)
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  4.674e+01  5.490e+03  0.009  0.9932
gender       -7.717e-01  5.792e-01 -1.332  0.1828
ageb16       1.916e+01  1.075e+04  0.002  0.9986
age16t20     4.814e-01  2.043e+00  0.236  0.8137
age21t25     1.423e+00  1.871e+00  0.761  0.4469
age26t30    -1.353e-01  1.957e+00 -0.069  0.9449
age31t35     2.225e+00  2.189e+00  1.016  0.3095
age36t40    -5.009e-02  1.907e+00 -0.026  0.9790
age41t45     1.821e+01  3.496e+03  0.005  0.9958
age46t50    -1.664e+01  2.677e+03 -0.006  0.9950
age51t55    -3.235e-01  2.056e+00 -0.157  0.8750
age56t60     2.057e+01  5.668e+03  0.004  0.9971
age61h              NA              NA      NA      NA
school_1     -3.500e+01  1.147e+04 -0.003  0.9976
school       -1.703e+01  3.976e+03 -0.004  0.9966
college      -1.538e+01  3.976e+03 -0.004  0.9969
bchlr_1     -1.640e+01  3.976e+03 -0.004  0.9967
bchlr       -1.710e+01  3.976e+03 -0.004  0.9966
mstr_1       1.452e+00  4.677e+03  0.000  0.9998
mstr        -1.788e+01  3.976e+03 -0.004  0.9964
hedu2       -1.663e+01  3.976e+03 -0.004  0.9967
degree              NA              NA      NA      NA
income1              NA              NA      NA      NA
income2     -3.133e+01  3.786e+03 -0.008  0.9934
income3     -3.160e+01  3.786e+03 -0.008  0.9933
income4     -3.178e+01  3.786e+03 -0.008  0.9933
income5     -3.128e+01  3.786e+03 -0.008  0.9934
income6              NA              NA      NA      NA
family      -6.356e-01  1.268e+00 -0.501  0.6162
children    1.951e+00  1.425e+00  1.369  0.1711
ps_alo      8.161e-01  6.234e-01  1.309  0.1905
food        1.651e+00  6.579e-01  2.509  0.0121 *

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 157.454  on 139  degrees of freedom
Residual deviance: 96.941  on 112  degrees of freedom
AIC: 152.94

Number of Fisher Scoring iterations: 18

```

Рисунок 3. Описание бинарной логистической регрессии  
 Источник: составлено авторами с помощью программы R

Как видно из анализа, около некоторых переменных в категориях возраст, уровень образования и уровень дохода домохозяйств стоит код *NA*, говорящий об отсутствии информации. Данные отсутствуют только для переменной *income1*, так как в выборке нет респондентов, оценивших доход своих домохозяйств как очень низкий.

Это также видно из описательной статистики исходных данных (рис. 2), представленной выше. Другие переменные исключены программой из-за наличия базовых переменных. Базовыми переменными в данном случае являются: переменная возраста – *age21t25*, переменная образования – *bchlr* и переменная дохода – *income4*. Таким образом, исключаем указанные переменные. С ними будут сравниваться все остальные. Получили следующие параметры описательной статистики бинарной логит-регрессии для модели без указанных выше переменных (рис. 4).

Deviance Residuals:				
Min	1Q	Median	3Q	Max
-2.29181	-0.00002	0.00018	0.58305	1.70091
Coefficients:				
	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-0.7179	0.9378	-0.765	0.4440
gender	-0.7717	0.5792	-1.332	0.1828
age16	17.7384	10754.0129	0.002	0.9987
age16t20	-0.9413	0.9387	-1.003	0.3160
age26t30	-1.5579	1.1016	-1.414	0.1573
age31t35	0.8020	1.4034	0.572	0.5677
age36t40	-1.4728	1.3459	-1.094	0.2739
age41t45	16.7858	3495.9610	0.005	0.9962
age46t50	-18.0620	2677.0175	-0.007	0.9946
age51t55	-1.7462	1.6946	-1.030	0.3028
age56t60	19.1465	5667.9866	0.003	0.9973
age61h	-1.4226	1.8706	-0.761	0.4469
school_1	-17.9069	10754.0128	-0.002	0.9987
school	0.0675	1.4212	0.047	0.9621
college	1.7153	1.2372	1.386	0.1656
bchlr_1	0.6973	1.0363	0.673	0.5010
mstr_1	18.5477	2462.6687	0.008	0.9940
mstr	-0.7823	1.2767	-0.613	0.5401
hedu2	0.4652	1.4804	0.314	0.7533
degree	17.0957	3975.6387	0.004	0.9966
income2	0.4494	1.4400	0.312	0.7550
income3	0.1847	0.6856	0.269	0.7877
income5	0.4992	0.7322	0.682	0.4954
income6	31.7829	3785.8372	0.008	0.9933
family	-0.6356	1.2680	-0.501	0.6162
children	1.9505	1.4250	1.369	0.1711
ps_alo	0.8161	0.6234	1.309	0.1905
food	1.6508	0.6579	2.509	0.0121 *
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
(Dispersion parameter for binomial family taken to be 1)				
Null deviance: 157.454 on 139 degrees of freedom				
Residual deviance: 96.941 on 112 degrees of freedom				
AIC: 152.94				
Number of Fisher Scoring iterations: 18				

Рисунок 4. Описание бинарной логистической регрессии (Модель № 1)

Источник: составлено авторами с помощью программы R

Цифры рядом с независимыми переменными – это соответствующие коэффициенты  $b_i$  (*Estimate*) и их стандартное отклонение (*Std. Error*). Коэффициент *Intercept* – свободный член, обозначающий пересечение с осью ординат. При экспоненцировании данный коэффициент дает отношение рисков наступления события в том случае, когда отсутствуют какие-либо факторы. Значения  $Pr(>|z|)$  характеризуют вероятность того, что коэффициенты  $b_i$  равны 0. Справа находится колонка, содержащая астериксы (\*), указывающие на значимость переменных. Три астериска (\*\*\*) означают, что величина p-value находится в промежутке между 0 и 0,001, два (\*\*) – в промежутке между 0,001 и 0,01, один (\*) – в промежутке между 0,01 и 0,05, точка (.) говорит о том, что p-value находится в диапазоне от 0,05 до 0,1, а пробел ( ) – от 0,1 до 1,0.

В Модели № 1 только одна значимая переменная – *food*, показывающая отношение респондентов к гастрономии, для которой p-value находится в диапазоне от 0,01 до 0,05. Так как коэффициент перед данной переменной положительный, то можно сделать вывод, что турист с большей вероятностью поедет в гастрономическое путешествие, если он/она положительно относится к еде в глобальном понимании.

Центр полевых исследований Барселоны выявил, что развитию гастротуризма также способствует появление огромного количества кулинарных каналов, программ и шоу [11]. В этой связи добавим в разобранную Модель № 1 еще две переменные: просмотр респондентами телепередач о путешествиях (*tv\_trip*) и программ о гастрономии, еде и ресторанах (*tv\_food*), где 0 говорит о том, что респондент не смотрит данные передачи, а 1 – наоборот. Соответственно, получаем Модель № 2, представленную на *рисунке 5*.

В Модели № 2 программа показала две объясняющие переменные: психотип гастротуриста (*ps\_alo*) и переменную, отвечающую за просмотр телепередач и программ о кулинарии (*tv\_food*). Переменная *tv\_food* более значима, чем *ps\_alo*, а положительный коэффициент перед данной переменной свидетельствует о том, что человек, который смотрит телепередачи о еде, с большей вероятностью поедет в гастротур, нежели тот, который такие программы не смотрит. Возвращаясь к переменной *ps\_alo*, видим, что коэффициент перед переменной также является положительным. Это говорит о том, что туристы-аллоцентрики с большей вероятностью отдадут предпочтение гастрономическому туризму, чем психоцентрики. Кроме того, значимой получилась константа, но обычно она не интерпретируется.

Возникает вопрос: «Какая из двух моделей лучше?» Для того чтобы сравнить Модель № 1 и Модель № 2 между собой, обратим внимание на предпоследнюю строчку, AIC (Akaike's Information Criterion) – информационный критерий Акаике, оценивающий оптимальность модели, меньшее значение которого указывает на хорошее соответствие данным при использовании меньшего числа параметров. Иначе говоря, информационный критерий Акаике представляет собой «штраф» за увеличение сложности модели. Несмотря на то, что в Модели № 2 были добавлены две переменные, то есть модель стала сложнее, показатель AIC ощутимо уменьшился. Следовательно, Модель № 2 с меньшим значением критерия AIC лучше.

```

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.72726 -0.00002  0.00022  0.52049  1.77595

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.998e+00  1.338e+00  -2.240  0.0251 *
gender       -9.031e-01  6.588e-01  -1.371  0.1704
ageb16       1.614e+01  1.075e+04   0.002  0.9988
age16t20    -1.493e+00  1.099e+00  -1.358  0.1743
age26t30    -1.771e+00  1.250e+00  -1.417  0.1566
age31t35    -1.586e-02  1.537e+00  -0.010  0.9918
age36t40    -2.320e+00  1.530e+00  -1.516  0.1294
age41t45    1.579e+01  3.429e+03   0.005  0.9963
age46t50    -1.919e+01  2.962e+03  -0.006  0.9948
age51t55    -1.522e+00  1.815e+00  -0.838  0.4018
age56t60    1.872e+01  5.823e+03   0.003  0.9974
age61h      -1.015e+00  1.908e+00  -0.532  0.5950
school_i    -1.636e+01  1.075e+04  -0.002  0.9988
school      3.685e-01  1.526e+00   0.241  0.8092
college     1.651e+00  1.355e+00   1.218  0.2233
bchlr_i     1.952e+00  1.279e+00   1.526  0.1269
mstr_i      1.967e+01  2.242e+03   0.009  0.9930
mstr        -4.960e-02  1.393e+00  -0.036  0.9716
hedu2       1.634e+00  1.633e+00   1.001  0.3170
degree      1.605e+01  3.933e+03   0.004  0.9967
income2     -5.491e-02  1.412e+00  -0.039  0.9690
income3     2.546e-01  7.489e-01   0.340  0.7339
income5     8.328e-01  8.482e-01   0.982  0.3262
income6     3.173e+01  4.189e+03   0.008  0.9940
family      2.008e-01  1.456e+00   0.138  0.8903
children    1.845e+00  1.545e+00   1.194  0.2323
ps_alo      1.365e+00  7.520e-01   1.815  0.0696 .
tv_trip     1.280e+00  8.052e-01   1.590  0.1118
tv_food     1.611e+00  7.177e-01   2.245  0.0248 *
food        1.144e+00  7.062e-01   1.620  0.1052

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 157.454  on 139  degrees of freedom
Residual deviance:  85.216  on 110  degrees of freedom
AIC: 145.22

Number of Fisher Scoring iterations: 18

```

Рисунок 5. Описание бинарной логистической регрессии (Модель № 2)

Источник: составлено авторами с помощью программы R

Теперь рассчитаем вероятность совершения гастрономической поездки туристами для Модели № 2. Задаем код, указанный на рисунке 6, и получаем, что вероятность того, что респондент с заданными в Модели № 2 характеристиками поедет в гастропутешествие, равна 85%.

```

> p<-predict(res,newdata=df)
> p<-ifelse(p>0.5,1,0)
> mean(p==df$question)
[1] 0.85

```

**Рисунок 6.** Расчет вероятности отправления респондентами в гастрономическое путешествие для Модели № 2

Источник: составлено авторами с помощью программы R

Coefficients:											
	Estimate	Std. Error	z value	Pr(> z )							
(Intercept)	1.8153	0.4076	4.454	8.43e-06 ***							
age16t20	-1.1221	1.2908	-0.869	0.3847							
age26t30	-1.1221	0.5395	-2.080	0.0375 *							
age31t35	-1.9976	0.7299	-2.737	0.0062 **							
age36t40	-0.3112	0.8816	-0.353	0.7241							
age41t45	-1.8153	0.9126	-1.989	0.0467 *							
age46t50	15.7508	1398.7210	0.011	0.9910							
age51t55	-1.8153	1.4718	-1.233	0.2174							
age56t60	-1.8153	1.0799	-1.681	0.0928 .							
age61h	15.7508	2284.1018	0.007	0.9945							
---	-0.2059	1.1688	-0.176	0.8602							
Signif. codes:	0	****	0.001	***	0.01	**	0.05	.	0.1	'	1

**Рисунок 7.** Определение степени влияния переменных из группы «Возраст» на модель

Источник: составлено авторами с помощью программы R

Рассмотрим возможность упрощения модели, а следовательно, ее улучшения посредством уменьшения АИС. Для этого с помощью той же функции сравним влияние *dummy*-переменных на модель в каждой группе (возраст, образование, доход) для их возможного последующего исключения. Для определения значимости переменных группы возраста вводим программный код, содержащий только переменные группы «Возраст» без базовой переменной *age21t25* (рис. 7).

Значимыми получились следующие переменные: *age26t30*, *age16t20*, *age36t40* и *age51t55* (переменные расставлены в порядке убывания значимости). Коэффициенты перед всеми переменными отрицательные, что можно интерпретировать следующим образом: данные возрастные категории (от 26 до 30, от 16 до 20, от 36 до 40 и от 51 до 55) поедут в гастрономическое путешествие с меньшей вероятностью, чем базовая группа респондентов в возрасте от 21 до 25. При этом туристы, относящиеся к возрастной категории 26-30 лет, поедут в гастротур с меньшей вероятностью, чем туристы в возрасте от 16 до 20 и от 36 до 40 лет. А туристы 16-20 лет и 36-40 лет – с меньшей вероятностью, чем возрастная группа от 51 до 55. Переменные, коэффициент перед которым не значим, не отличаются от базовой категории.

Что касается групп «Образование» и «Доход», представленных на рисунках 8 и 9, соответственно, значимых переменных, кроме свободного члена, в них нет, следовательно, их можно исключить из Модели № 2.

Coefficients:				
	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.7309	0.3376	2.165	0.0304 *
school_i	-0.7309	1.0554	-0.692	0.4886
school	-0.2201	0.8045	-0.274	0.7845
college	0.3677	0.8835	0.416	0.6773
bchlr_i	0.3338	0.4984	0.670	0.5030
mstr_i	17.8352	1630.6597	0.011	0.9913
mstr	0.3677	0.7473	0.492	0.6227
hedu2	0.3677	0.8835	0.416	0.6773
degree	17.8352	2917.0127	0.006	0.9951
---				
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

**Рисунок 8.** Определение степени влияния переменных из группы «Образование» на модель  
 Источник: составлено авторами с помощью программы R

Coefficients:				
	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.7577	0.3129	2.421	0.0155 *
income2	0.1586	0.8933	0.178	0.8591
income3	0.3708	0.4672	0.794	0.4274
income5	0.5545	0.5284	1.049	0.2940
income6	16.8084	1398.7210	0.012	0.9904
---				
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

**Рисунок 9.** Определение степени влияния переменных из группы «Доход» на модель  
 Источник: составлено авторами с помощью программы R

Таким образом, получаем Модель № 3, описательные статистики которой представлены на *рисунке 10*.

Информационный критерий Акаике в Модели № 3 незначительно меньше, чем в Модели № 2, следовательно, можно сделать предположение, что Модель № 3 лучше. Теперь рассчитаем вероятность совершения гастропоездки для данной модели (*рис. 11*).

Вероятность отправления респондентами в гастротур в Модели № 3 меньше, чем в предыдущей разобранный модели. В связи с тем, что из полученных результатов нельзя сделать однозначный вывод о том, какая из двух моделей (№ 2 или № 3) лучше, воспользуемся матрицей неточностей (Confusion Matrix) или, иначе говоря, четырехпольной таблицей сопряженности, а также показателем AUC (Area Under Curve) ROC (Receiver Operator Characteristic) – площадь под кривой ошибок, рассчитанных для обеих моделей.

Матрица неточностей показывает, как часто модель делает правильные предсказания (True Negatives, True Positives) и как часто ошибается (False Positives, False Negatives). TN – верно классифицированные отрицательные примеры; TP – истинно положительные случаи; FN показывает число «ложных пропусков», то есть положительных примеров, классифицированных как отрицательные; FP – ложное обнаружение, когда отрицательные примеры классифицируются как положительные (*табл. 1*).

```

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.68404 -0.07247  0.34510  0.64590  1.55559

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -0.5846    0.7714  -0.758  0.4486
gender        -0.8115    0.5271  -1.540  0.1236
agebl6        -1.2273    1.6185  -0.758  0.4483
age16t20      -1.1461    0.6374  -1.798  0.0721 .
age26t30      -2.2346    0.9577  -2.333  0.0196 *
age31t35      -1.0457    1.2084  -0.865  0.3869
age36t40      -3.7058    1.4479  -2.560  0.0105 *
age41t45      13.1622   1318.9136  0.010  0.9920
age46t50      -4.3225    2.2070  -1.959  0.0502 .
age51t55      -2.9796    1.4218  -2.096  0.0361 *
age56t60      14.8326   2218.6627  0.007  0.9947
age61h        -1.7052    1.5249  -1.118  0.2635
family         0.7765    1.0089  0.770  0.4415
children       1.3535    1.2721  1.064  0.2873
ps_alo         1.0277    0.5590  1.839  0.0660 .
tv_trip        0.9994    0.6703  1.491  0.1360
tv_food        0.9561    0.5602  1.707  0.0879 .
food           1.1758    0.5598  2.100  0.0357 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 157.45  on 139  degrees of freedom
Residual deviance: 109.03  on 122  degrees of freedom
AIC: 145.03

Number of Fisher Scoring iterations: 16

```

**Рисунок 10.** Описание бинарной логистической регрессии (Модель № 3)  
 Источник: составлено авторами с помощью программы R

```

> p<-predict(res,newdata=df)
> p<-ifelse(p>0.5,1,0)
> mean(p==df$question)
[1] 0.8071429

```

**Рисунок 11.** Расчет вероятности отправления респондентами в гастрономическое путешествие для Модели № 3

Источник: составлено авторами с помощью программы R

На основе этих данных можно посчитать такие показатели, как чувствительность (Sensitivity, True Positives Rate), показывающая долю истинно положительных значений; специфичность (Specificity, True Negative Rate) – долю истинно отрицательных случаев, которые были правильно идентифицированы моделью; долю ложно положи-

Таблица 1

## Матрица неточностей для Модели № 2 и Модели № 3

	Модель №2		Модель №3	
	False	True	False	True
0	24 (TN)	11 (FP)	18 (TN)	17 (FP)
1	7 (FN)	98 (TP)	7 (FN)	98 (TP)

Источник: составлено авторами на основе данных из программы R

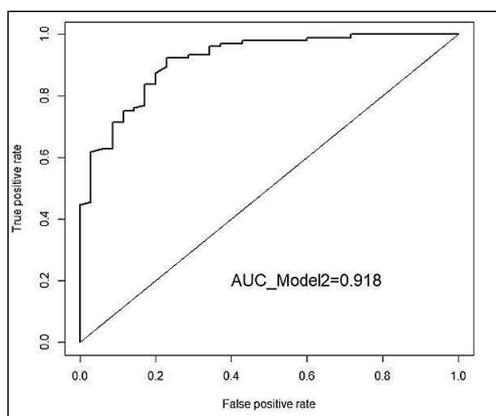


Рисунок 12. ROC-кривая и AUC для Модели № 2

Источник: составлено авторами с помощью программы R

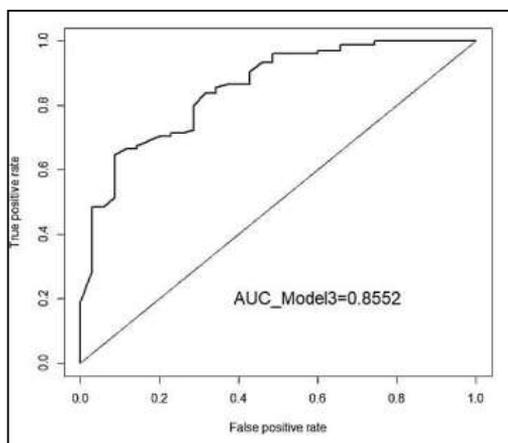


Рисунок 13. ROC-кривая и AUC для Модели № 3

Источник: составлено авторами с помощью программы R

тельных примеров (Fall-Out, False Positives Rate); долю ложно отрицательных значений (Miss Rate, False Negative Rate) и точность (Accuracy).

Построим ROC-кривую, часто называемую «кривой ошибок», которая представляет собой зависимость доли истинно положительных значений (TPR) от доли ложно положительных примеров (FPR), и вычислим показатель AUC – площадь под ROC-кривой для каждой из двух моделей. Чем выше данный показатель, тем выше качество модели (рис. 12 и 13).

Для характеристики полученных результатов и, соответственно, качества Модели № 2 и Модели № 3 воспользуемся экспертной шкалой для оценки значений показателя AUC, где интервал AUC от 0,5 до 0,6 говорит о неудовлетворительном качестве модели; 0,6–0,7 – среднем качестве модели; 0,7–0,8 – хорошем качестве; 0,8–0,9 – очень хорошем и интервал от 0,9 до 1,0 демонстрирует отличное качество модели [12].

Из рисунков 12 и 13 видно, что показатель AUC для Модели № 2 немного выше и говорит об отличном качестве модели. Следовательно, по этому критерию Модель № 2 лучше.

Составим *таблицу 2*, в которой представлены результаты всех использованных методов и критериев для сравнения моделей, а также отмечены лучшие показатели для Модели № 2 и Модели № 3.

Исходя из данных, представленных в таблице 2, можно сделать вывод, что Модель № 2 гораздо лучше подходит для описания данного исследования.

Еще один вариант использования эконометрических методов для исследования отрасли туризма – корреляционный анализ видов туризма. Данные для этого вида анализа были собраны с помощью метода попарного сравнения [3, с. 397] (*Malkhotra, 2016*).

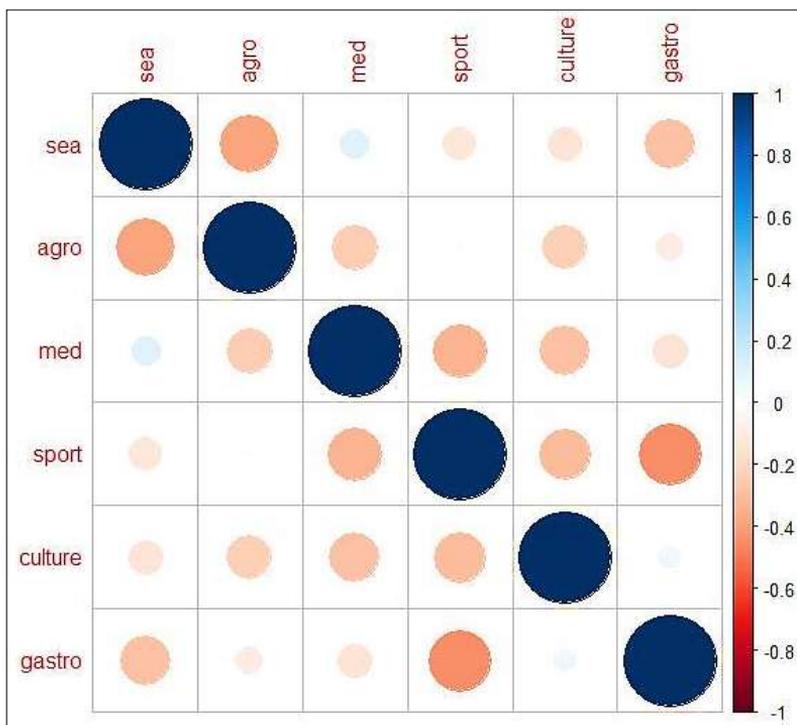
Попарно сравнивались между собой шесть видов туризма: агротуризм, гастрономический, культурно-познавательный, лечебно-оздоровительный, спортивный туризм и морской. После первичной обработки данные были приведены к порядковому виду.

Таблица 2

### Результаты сравнения Модели № 2 и Модели № 3

Критерий сравнения	Модель №2	Модель №3
AIC	145,22	145,03
P(Y)	85,00%	80,71%
Sensitivity	93,33%	93,33%
Specificity	68,57%	51,43%
FPR	31,43%	48,57%
FNR	6,67%	6,67%
Accuracy	87,14%	82,86%
AUC	0,981	0,8552

Источник: составлено авторами



**Рисунок 14.** Визуализация корреляционной матрицы  
 Источник: составлено авторами в программе R

Соответственно, для проведения корреляционного анализа порядковых данных необходимо использовать показатель неметрической корреляции – коэффициент ранговой корреляции Спирмена (рис. 14).

Таким образом, наблюдается отрицательная корреляция у всех пар видов туризма, кроме двух: морской – лечебно-оздоровительный и культурно-познавательный – гастрономический, у которых видна слабая положительная корреляция.

### Заключение

1. В современном мире туризм становится все более востребованным видом отдыха. С увеличением доли туристских услуг происходит и их дифференциация.

2. Представленные анализы являются только малой демонстрацией возможностей использования эконометрических моделей для исследования сферы туризма.

3. Проведение аналогичных, но более масштабных исследований данной отрасли представляет интерес, прежде всего, с точки зрения возможных положительных эффектов для развития национальной экономики через создание персонифициро-

ванного конкурентоспособного туристского предложения на территории Российской Федерации.

### ИСТОЧНИКИ:

1. Котлер Ф., Боуэн Дж., Мейкенз Дж. Маркетинг. Гостеприимство. Туризм. / Учебник для студентов вузов. 4-е изд. – М.: ЮНИТИ-ДАНА, 2013. – 1071 с.
2. Лебедева С.А. Современная классификация туристских сегментов // Молодой ученый. – 2017. – № 12. – с. 323-326.
3. Нэреш К. Малхотра Маркетинговые исследования. / Практическое руководство: Пер. с англ. – М.: ООО «И.Д. Вильямс», 2016. – 1184 с.
4. Роберт И. Кабаков R в действии. Анализ и визуализация данных на языке R. / (пер. с англ. Волковой П.А.). – М.: ДМК Пресс, 2014. – 588 с.
5. Feray Irigüler, Mehmet Emre Güler Tourist Guides' Perceptions on Gastronomic Tourism in Turkey // Journal of Tourism and Gastronomy Studies. – 2017. – № 5(1). – p. 3-19.
6. Georgică Gheorghe, Petronela Tudorache, Puiu Nistoreanu Gastronomic tourism, a new trend for contemporary tourism? // Cactus Tourism Journal. – 2014. – № 9(1). – p. 12-21.
7. Helena A. Williams, Robert L. Williams Jr., Maktoba Omar Gastro-tourism as destination branding in emerging markets // Leisure and Tourism Marketing. – 2014. – № 4(1). – p. 1-18.
8. Marius Boita, Andrei Anghelina, Bobby Costi Gastronomic tourism management revitalizer factor of the tourism offer // Annals of the “Constantin Brancusi” University of Targu Jiu, Economy Series. – 2014. – p. 398-401.
9. Stanley C. Why destination areas rise and fall in popularity // Cornell Hotel and Restaurant Administration Quarterly. – 2001. – № 3. – p. 13-24. – doi: 10.1177/0010880401423001.
10. Официальный сайт Всемирной ассоциации гастрономического туризма. [Электронный ресурс]. URL: <https://www.worldfoodtravel.org/cpages/home> ( дата обращения: 04.06.2018 ).
11. Официальный сайт Центра полевых исследований Барселоны. [Электронный ресурс]. URL: <https://geographyfieldwork.com/FoodTourism.htm> ( дата обращения: 04.06.2018 ).
12. Логистическая регрессия и ROC-анализ – математический аппарат. Сайт Технологий анализа данных. [Электронный ресурс]. URL: <https://basegroup.ru/community/articles/logistic> ( дата обращения: 04.06.2018 ).

### REFERENCES:

- Feray Irigüler, Mehmet Emre Güler (2017). Tourist Guides' Perceptions on Gastronomic Tourism in Turkey *Journal of Tourism and Gastronomy Studies*. (5(1)). 3-19.
- Georgică Gheorghe, Petronela Tudorache, Puiu Nistoreanu (2014). Gastronomic tourism, a new trend for contemporary tourism? *Cactus Tourism Journal*. (9(1)). 12-21.

- Helena A. Williams, Robert L. Williams Jr., Maktoba Omar (2014). Gastro-tourism as destination branding in emerging markets *Leisure and Tourism Marketing*. (4(1)). 1-18.
- Kotler F., Bouen Dzh., Meykenz Dzh. (2013). *Marketing. Gostepriimstvo. Turizm* [Marketing. Hospitality. Tourism] M.: YuNITI-DANA. (in Russian).
- Lebedeva S.A. (2017). Sovremennaya klassifikatsiya turistskikh segmentov [Modern classification of tourist segments]. *The young scientist*. (12). 323-326. (in Russian).
- Marius Boita, Andrei Anghelina, Boby Costi (2014). Gastronomic tourism management revitalizer factor of the tourism offer *Annals of the "Constantin Brancusi" University of Targu Jiu, Economy Series*. 398-401.
- Neresh K. Malkhotra (2016). *Marketingovye issledovaniya* [Marketing research] M.: OOO «I.D. Vilyams». (in Russian).
- Robert I. Kabakov (2014). *R v deystvii. Analiz i vizualizatsiya dannyh na yazyke R* [R in action. Data analysis and visualization in R language] M.: DMK Press. (in Russian).
- Stanley C. (2001). Why destination areas rise and fall in popularity *Cornell Hotel and Restaurant Administration Quarterly*. 42 (3). 13-24. doi: 10.1177/0010880401423001.