

# АНАЛИЗ ПРЕДПОЧТЕНИЙ ПОТРЕБИТЕЛЕЙ НА РЫНКЕ АВТОМОБИЛЬНОГО БЕНЗИНА

УДК 339.13

**Татьяна Александровна Сафина**,  
доцент каф. Математики, информатики  
и информационной безопасности Авто-  
номной некоммерческой организации  
высшего профессионального образо-  
вания Межрегиональный открытый со-  
циальный институт (АНО ВПО МОСИ)  
Тел.: 8 (8362) 63-56-70  
Эл. почта: t-safina@list.ru

В статье предложен статистический анализ, основанный на результатах выборочного обследования потребителей на рынке автомобильного бензина, который позволил выявить предпочтения респондентов в выборе марки бензина с учетом их социально-экономического статуса.

**Ключевые слова:** анонимное анкетирование, логистическая регрессионная модель, рынок автомобильного бензина, предпочтения респондентов.

**Tatyana A. Safina**,  
Interregional Open Social Institute As-  
sociate Professor of the Department of  
mathematics, Informatics and Information  
Security  
Tel.: 8 (8362) 63-56-70  
E-mail: t-safina@list.ru

## ANALYSIS OF CONSUMER PREFERENCES ON THE MARKET AUTOMOBILE GASOLINE

In this paper the statistical analysis based on results of selective investigation of customers in the market of automobile gasoline is offered. The analysis makes the possibility for finding preferences of respondents in choices of gasoline brand depending of social and economics status of customers.

**Keywords:** anonymous survey, logistical regression model, market of automobile gasoline, customer preferences.

### 1. Введение

В России выпускается четыре марки автомобильного бензина: «Нормаль-80», «Регуляр-92», «Премиум-95» и «Супер-98». Основным объемом производимого в стране автомобильного бензина приходится на марку «Регуляр-92» – 67,3%. Бензин марки «Нормаль-80» производится в объеме 10,6%. Практически вся оставшаяся часть – это производство бензина марки «Премиум-95» (21,3%). Бензин марки «Супер-98» в РФ выпускается в крайне малых объемах – всего 0,8% от общего объема производства автомобильного бензина.

Целью данного исследования является, выявить предпочтения потребителей на рынке автомобильного бензина в Республике Марий-Эл.

### 2. Статистическое исследование выявления предпочтений в выборе потребителей марки автомобильного бензина Республики Марий Эл на основе анонимного анкетирования.

В Республике Марий Эл действует 72 крупных АЗС, в т. ч. в Йошкар-Оле – 39. При этом 48,6% из них принадлежат трём компаниям: Лукойлу принадлежит 17 АЗС; Татнефти – 9 и АЗС» – 9.

С целью выявления предпочтений потребителей автозаправочных станций и определения предпочтительной марки топлива проводилось анкетирование на основе опросов.

В анкетировании приняли участие 500 жителей г. Йошкар-Олы. Общее количество жителей города – 252935 человек, из которых 160107 человек трудоспособного возраста. По данным Росстата, на каждую 1000 жителей республики приходится 163 автомобиля.

Исследование проводилось в местах сосредоточенности автомобилистов: на автостоянках крупных магазинов, на круглосуточных стоянках. Чтобы избежать искажения мнений респондентов относительно предпочтений, опрос намеренно не проводился на АЗС.

С предполагаемыми респондентами было проведено анонимное анкетирование. Для определения предпочтений респондентов относительно марки бензина с учетом их социально-экономического статуса, были выбраны следующие показатели:

1. Пол
2. Возраст
3. Социальный статус
  - Образование
  - Занятость
4. Благосостояние
  - Доход на одного члена семьи в месяц
  - Марка автомобиля
5. Личные предпочтения
  - Предпочтительный бренд АЗС
  - Используемое топливо
  - Марка используемого бензина
  - Предпочтение при выборе АЗС
  - Объем потребления топлива по сезонам
  - Предпочтения в дополнительных услугах АЗС
  - Получение информации о предпочтенной АЗС.

Первичный анализ обработки данных показал, что по половому признаку данные распределились следующим образом: из общего числа респондентов (500) мужчин – 403 человека (80,6%) и 97 – женщин (19,4%).

Далее респонденты были разделены по возрастным группам: «молодежь» – возраст респондентов до 30 лет – 115 (23%), «средний возраст» –

30–50 лет – 320 респондентов (64%) и «старший возраст» – старше 51 года – 65 человек (13%).

Наибольшее число респондентов имеет среднее специальное образование – 320 (64%), высшее образование – 180 респондентов (36%).

По занятости респондентов группа «безработные» – 59 респондентов (11,8%), группа «пенсионеры (неработающие)» – 24 (4,8%) и группа «работающие» – 417 (83,4%).

По уровню дохода на одного члена семьи в месяц большинство опрошенных – 332 человека (66,4%) – указали доход на 1-го члена семьи до 10 тысяч рублей.

Большинство участников опроса имеют автомобиль стоимостью свыше 300 тыс. руб. – 284 респондента (56,8%), соответственно, автомобили стоимостью ниже 300 тыс. руб. – 216 человек (43,2%).

По предпочтению бренда АЗС предпочтения разделились следующим образом: «Лукойл» – 151 респондент (30,2%), «Татнефть» – 141 человек (28%), «Марийский бензин» – 86 опрошенных (17,2%), «МНК» – 36 человек (7,2%), «МТК» – 32 респондента (6,4%), «Мар-Ойл» – 54 опрошенных (11%).

По предпочтению выбора АЗС критерий «месторасположение» является ключевым, его выбрали 472 респондента (94,6%).

Критерий предпочтения «качество топлива» у АЗС отметили 2% опрошенных, критерий «приемлемая цена» – 3% респондентов и остальные 0,4% приходятся на «проведение розыгрышей» и «дис-

контные карты», однако данные респонденты отметили и критерий «месторасположение».

Покупательские предпочтения показали, что больше половины респондентов (56,3%) заправляют автомобиль от одного до двух раз в неделю, ежедневно это делают 19,0% респондентов и по потребности – 24,7%.

Заправка автомобиля ГСМ изменяется в зависимости от сезона, о чем свидетельствуют данные, приведенные в таблице 1.

На АЗС чаще пользуются следующими видами дополнительных услуг: услугами автомойки – 44,9% респондентов, шиномонтажа – 35,7%, оставшиеся 19,4% опрошенных посещают магазин при АЗС, пользуются автопылесосом и компрессором.

«Собственный опыт» в качестве основного источника получения информации при выборе АЗС отметило колоссальное большинство опрошенных – 90%, ответ «на работе» отметили 10% респондентов, ответ «реклама» не выбрал ни один респондент.

Бензин марки АИ-92 пользуется наибольшим спросом у большинства респондентов – 323 (64,4%).

В итоге можно сделать вывод что, потребителем АЗС в г. Йошкар-Оле является мужчина в возрасте от 30 до 50 лет, который имеет среднее специальное образование, владеет новым автомобилем экономкласса отечественной сборки и заправляет свой автомобиль бензином марки АИ-92 один-два раза в неделю. Доход респондента в месяц составляет

на 1-го члена семьи меньше 10 тысяч рублей, и при выборе АЗС для него главным является месторасположение заправочной станции.

### 3. Выявление предпочтений респондентов в выборе марки бензина при помощи инструмента статистической обработки данных – логистической регрессии.

Из предварительно проведенного исследования был выбран ряд показателей, которые могут помочь определить предпочтения респондентов:  $y$  – используемый вид топлива;  $x_1$  – пол;  $x_2$  – возраст;  $x_3$  – образование;  $x_4$  – занятость;  $x_5$  – доход на одного члена семьи в месяц;  $x_6$  – марка автомобиля;  $x_7$  – предпочтительный бренд АЗС;  $x_8$  – месторасположение АЗС.

При анализе влияния факторов на предпочтения респондентов в выборе вида топлива использована логистическая регрессия, где в качестве результирующей рассматривалась бинарная переменная:

$$y = \begin{cases} 1 - \text{бензин марки АИ-92} \\ 0 - \text{другие виды топлива} \end{cases}$$

В качестве объясняющих рассматривалось шесть качественных переменных: показатели  $x_1$  – пол;  $x_3$  – образование;  $x_4$  – занятость;  $x_6$  – марка автомобиля;  $x_7$  – предпочтительный бренд АЗС,  $x_8$  – месторасположение АЗС, а также две количественные переменные: показатель  $x_2$  – возраст респондента,  $x_5$  – доход на 1-го члена семьи в месяц.

$$x_1 = \begin{cases} 1 - \text{мужчина} \\ 0 - \text{женщина} \end{cases}$$

$$x_3 = \begin{cases} 0 - \text{другое} \\ 1 - \text{высшее} \end{cases}$$

$$x_4 = \begin{cases} 1 - \text{работающий} \\ 0 - \text{неработающий} \end{cases}$$

$$x_6 = \begin{cases} 0 - \text{автомобиль до 300 тыс. руб.} \\ 1 - \text{свыше 300 тыс. руб.} \end{cases}$$

$$x_7 = \begin{cases} 1 - \text{Лукойл} \\ 2 - \text{Татнефть} \\ 3 - \text{Марийский бензин} \\ (\text{в т.ч. МТК, МНК, М - Ойл}) \end{cases}$$

Таблица 1  
Объем покупаемого ГСМ в зависимости от сезона

В теплый сезон (литров)	Доля респондентов, %	В холодный сезон (литров)	Доля респондентов, %
5	5,2	5	0
10	21	10	6,4
15	19	15	14,2
20	23	20	23,5
25	11	25	22,4
30	11	30	17
35	2,5	35	7,3
40	0,3	40	2,2
более 45	7	более 45	7
	100		100

$$x_8 = \begin{cases} 0 - \text{далеко от дома или} \\ \text{работы.} \\ 1 - \text{близко к дому или к} \\ \text{работе} \end{cases}$$

В проводимом исследовании выдвинуто предположение о том, что все выбранные переменные являются значимыми факторами, оказывающими влияние на предпочтение автомобилистов в выборе используемого вида топлива. Построенная модель основана на последовательном разложении исходных вероятностей выбора того или иного вида топлива, рассчитанных на основе статистических данных отдельных компонент, которые различаются в зависимости от индивида (пола, возраста, уровня дохода, наличия автомобиля и предпочтительного бренда АЗС). С этой целью использованы данные анонимного анкетирования жителей республики. В качестве инструмента статистической обработки данных выбрана логистическая регрессия [1,2], в которой определяют вероятность события как функцию специального вида  $F(Z)$ , в котором в роли аргумента  $Z_i$  выступает линейная функция от объясняющих переменных:

$$Z_i = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_px_p \quad (1)$$

Оценка параметров  $b_0, b_1, b_p, \dots, b_p$  в таких моделях связана с необходимостью решения специального класса нелинейных задач оптимизации [2]. Модель бинарного вида можно представить следующим образом:

$$\begin{cases} \hat{y} = \frac{e^z}{1 + e^z} \\ \hat{z} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_px_p \end{cases} \quad (2)$$

На первом этапе проверяется качество приближенной модели. Вероятность получить наблюдаемые результаты при заданных оценках параметров называют правдоподобием оценок. Поскольку оценки максимального правдоподобия не имеют аналога  $R^2$ , используемого в методе МНК, применяется несколько коэффициентов: псевдо –  $R^2$  и  $R^2$  – Макфаддена, называемые индексом отношения правдоподобия.

$$R^2_{psedo} = 1 - \frac{1}{1 + 2(\ln L - \ln L_0)/n}, \quad (3)$$

$$R^2_{Mc-Fadden} = 1 - \frac{\ln L}{\ln L_0}, \quad (4)$$

Индекс отношения правдоподобия варьируется от 0 до 1, подобно коэффициенту  $R^2$ , применяемому для оценки качества линейной регрессии. Чем больше различие, тем лучше модель. Кроме рассмотренных коэффициентов для анализа оценки качества приближенной модели существуют и другие, например псевдо-коэффициент детерминации  $R^2_{psedo}$ . В качестве критической статистики можно использовать разницу максимумов логарифмических функций правдоподобия, что является тестом отношения правдоподобия ( $LR$ ) и аналогом  $F$  – критерия [1,3]:

$$LR = -2(\ln L - \ln L_0) \quad (5)$$

Для оценки параметров модели бинарного вида проверяется значимость коэффициентов – статистики. Для большого количества ограниченный проверка гипотез о значимости может проводиться при помощи тестов Вальда, отношения правдоподобия, множителей Лагранжа. Так как  $y$  в логит-модели представлен двумя переменными 1 и 0, то все предсказанные значения будут находиться внутри отрезка от 0 до 1, поэтому полученные значения можно рассматривать как вероятностные. По критерию Вальда проверяется гипотеза: существенно ли отличен от нуля коэффициент регрессии  $b_j$  при некотором заданном уровне значимости, который пока-

зывает вероятность отвергнуть правильную гипотезу [1,4]. Для построения логистической модели использован регрессионный анализ, который основан на исключении несущественных факторов из модели, (тест Вальда), т.е. из модели удаляется переменная с наибольшим  $p$  – значением, при условии, что он превышает пороговое значение 0,1. Для оценивания панели был выбран метод Квази-Ньютона и Резенброка. В полученной логит-регрессии  $p < 5\%$ . Значение статистики  $\chi^2 = 60,98 > \chi^2_{табл} = 1,89$ , значение высоко значимо. На основании этого сделан вывод, вошедшие в модель переменные влияют на предпочтение потребителей на рынке автомобильного бензина в РМЭ. Значимость регрессионных коэффициентов проверена при помощи статистики Вальда и –статистики, которые показали значимость пяти коэффициентов при переменных:  $x_1$  – пол;  $x_3$  – образование;  $x_4$  – занятость;  $x_6$  – наличие бюджетного автомобиля; переменные  $x_2$  – возраст и  $x_5$  – доход на одного члена семьи в месяц;  $x_7$  – предпочтительный бренд АЗС не оказали влияние на  $y$  – используемый вид топлива.

В результате оценки коэффициентов была получена модель:

$$\begin{cases} \hat{y} = \frac{e^z}{1 + e^z} \\ \hat{z} = 8,84 + 1,23x_1 - 0,06x_2 + 1,03x_3 + \\ + 1,07x_4 - 0,0006x_5 - 3,28x_6 \end{cases} \quad (6)$$

По параметрам  $b$  логит-модель является нелинейной. Разъяснение параметра  $b$  логит-модели отличается от объяснения коэффици-

Таблица 2

Итоговые значения логит-регрессии

$n = 500$	$b_0$	$x_1$ пол	$x_2$ возраст	$x_3$ образ-е	$x_4$ занятость	$x_5$ доход	$x_6$ марка авто
Коэффициент $b$	8,84	1,23	-0,06	1,03	1,07	-0,0006	-3,28
Ошибка стандартная ( $\hat{S}$ )	1,71	0,29	0,02	0,35	0,41	0,0001	1,05
$t$ статистика (288)	5,16	4,13	-2,49	2,98	2,63	-4,57	-3,12
( $p$ – уровень)	0,000	0,000	0,013	0,003	0,008	0,000	0,001
Статистика Вальда	26,59	17,02	6,21	8,87	6,94	20,88	9,72
( $p$ – уровень)	0,000	0,000	0,012	0,002	0,008	0,000	0,001
Шансы	6914	3,41	0,93	2,80	2,91	0,99	0,03

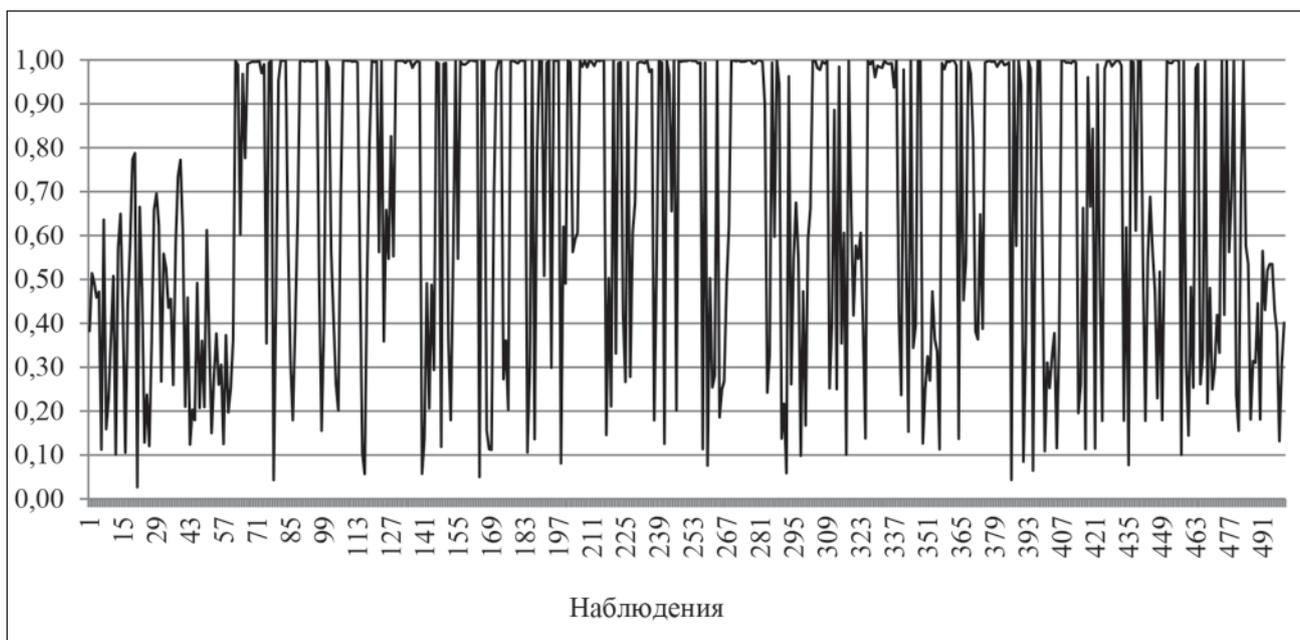


Рис. 1. Рассчитанные по модели вероятности предпочтения марки бензина АИ-92

ентов линейных регрессионных моделей. Их интерпретацию можно рассматривать как вероятность наступления какого-либо события или отношение вероятности того, что событие никогда не наступит, к вероятности, что наступит. Можно использовать отношение шансов, а не воздействие на логарифм такого отношения.

$$e^{b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p} = e^{b_0} \cdot e^{b_1 x_1} \cdot \dots \cdot e^{b_p x_p}, \quad (7)$$

Из этого следует, что  $e^{b_i}$  показывает, насколько меняются шансы при возрастании на единицу значения  $i$ -й переменной [5].

Положительные коэффициенты показывают, что шансы в выборе марки бензина АИ-92 увеличиваются, а отрицательные уменьшаются.

Полученные признаки в модели проявляют себя вполне предсказуемо. Чем ниже уровень образования, тем ниже оплачиваемая работа, тем ниже доход; чем дешевле автомобиль, тем дешевле выбираемая марка бензина.

Наличие бюджетного автомобиля напрямую повышает шансы на выбор той или иной марки топлива.

На графическом изображении модели (рис.1) видно, что, действительно, большинство респондентов

с точностью приблизительно до 1,0 выберут бензин марки АИ-92.

При помощи аналога  $R^2$  для линейной регрессии  $Mc - Fadden$  который называется индексом отношения правдоподобия, проводилась оценка качества полученной модели:

$$R^2_{Mc-Fadden} = 1 - \frac{LL_{model}}{LL_0}, \quad (8)$$

где  $LL_{model} = 380,68$ ;  $LL_0 = 541,66$ ; значение  $R^2_{Mc-Fadden}$ , это означает, что 71% решений зависит от рассмотренных факторов. Оценка правильности предсказаний логистической регрессии представлена в Таблице 3.

Модель правильно прогнозирует респондента, который желает заправить бензин марки АИ-92, для 276 из 323 анкетированных (85,45%).

Модель правильно предсказала намерение респондента заправить другой вид топлива у 35 из 177 опрошенных (80,23%). Всего было правильно классифицировано 418 респондентов из 500, т.е. 83,60%. Оценка достоверности и надежности показывает, что для определения намерения заправить бензин определенной марки достаточно учитывать 6 факторов, характеризующих пол, образование, занятость, наличие автомобиля.

#### 4. Заключение

Проведенный статистический анализ на основе выборочного обследования предпочтений потребителей на рынке автомобильного бензина позволил выявить предпочтения респондентов в выборе

Таблица 3

Корректно классифицируемые наблюдения

Предпочтение респондента	Предсказание, что респонденты выберут марку бензина АИ-92	Предсказание, что респонденты выберут другой вид топлива	Процент корректности предсказаний
Намерение респондента заправить бензин марки АИ-92	276	47	85,45
Намерение респондента заправить другой вид топлива	35	142	80,23
Общий процент			83,60

марки бензина в зависимости от их социально-экономического статуса, показал, что среднестатистическим потребителем продукции АЗС в г. Йошкар-Оле является мужчина в возрасте от 30 до 50 лет, который имеет среднее специальное образование, владеет новым автомобилем экономкласса отечественной сборки и заправляет свой автомобиль бензином марки АИ-92 один-два раза в неделю. Доход респондента в месяц составляет на 1-го члена семьи меньше 10 тысяч рублей, при выборе АЗС для него главным является месторасположение заправочной станции. Построенная модель логит-регрессии позволила доказать, что первичная обработка анкет была проведена достаточно точно. Модель правильно прогнозирует респондента, который желает заправить бензин марки АИ-92, для 276 из 323 анкетированных (85,45%). Модель правильно предсказала намерение респондента заправить другой вид топлива у 35 из 177 опрошенных респондентов (80,23%). Оценка достоверности и надежности показывает, что для опреде-

ления намерения заправить бензин определенной марки достаточно учитывать 6 факторов, характеризующих пол; возраст, образование; занятость; доход на 1 члена семьи, наличие бюджетного автомобиля и месторасположение АЗС. Модель правильно относит намерение заправить автомобиль бензином марки АИ-92 (85,45%) к желающим заправить другой вид топлива (80,23%).

#### Литература

1. Бакуменко Л.П. Статистические аспекты исследования качества жизни населения [Текст]: моногр. / Бакуменко Л.П.; Мар. гос. тех. ун-т, – Йошкар-Ола, 2011.–152 с.
2. Мхитарян В.С. Нелинейный регрессионный анализ в системе Statistica и SPSS [Текст]: учеб. пособие. / Мхитарян В.С., Архипова М.Ю., Архипов В.Ю. – М.: МЭСИ, 2006.–91 с.
3. Айвазян С.А. Прикладная статистика. Основы эконометрики: учебник для вузов [Текст]: в 2 т. – т. 2.// Айвазян С.А. Основы эконометрики. –2-е изд., пере-

раб. и доп. – М.: ЮНИТИ–ДАНА, 2001. – 432 с.

4. Елисеева И.И. Эконометрика [Текст]: учеб. / под ред. Елисеевой И. И. – М.: Финансы и статистика, 2004. – 342 с.

5. Гусынин А.Б., Минашкин В.Г. Теория выборочных обследований [Текст]: учеб. пособие. / Гусынин А.Б., Минашкин В.Г. – М.: МЭСИ, 2003. – 67 с.

#### References

1. Bakumenko L.P. Statistical aspects of investigation of population life quality. – Mari State technical University publishing, Yoshkar-Ola, 2011, 152 p.
2. Mhitaryan V.S. The nonlinear regression analysis in Statistica and SPSS system //Mhitaryan V.S., Arhipova M.Yu., Arhipov V.Yu. – М.: MESI, 2006.–91 s.
3. Aivasyan S.A. Applied statistics. Base of econometrics. – М.: YUNITI DANA, 2001, 432 p.
4. Eliseeva I.I. Econometrics. – М.: Finansy i statistika, 2004. – 342 s.
5. Gusynin A.B., Minashkin V.G. Theory of selective investigation. – М.: MESI, 2003, 67 p.