

$$\begin{aligned} \dot{\psi}(t) &= -\frac{\partial H_f(x^0(t), y^0(t), \psi(t), u^0(t))}{\partial x} - \\ & - \frac{\partial H_f(x^0(s), y^0(s), \psi(s), u^0(s))}{\partial y} \Big|_{s=t+h}, t \in T, \\ \psi(t^* + 0) &= -\frac{\partial \varphi(x^0(t^*))}{\partial x}, \psi(t) \equiv 0, t > t^*, \\ \psi(\tau - 0) &= \psi(\tau + 0) - \frac{\partial H_g(x^0(\tau - 0), \psi(\tau + 0), v^0(\tau))}{\partial x}, \tau \in T_h, \end{aligned}$$

удовлетворяет условиям максимума:

$$\begin{aligned} & \int_t^{t+h} H_f(x^0(s), y^0(s), \psi^0(s), u^0(s)) ds = \\ & = \max_{u \in U} \int_t^{t+h} H_f(x^0(s), y^0(s), \psi^0(s), u) ds, t \in T_h, \\ & H_g(x^0(t-0), \psi^0(t+0), v^0(t), h) = \\ & = \max_{v \in V} H_g(x^0(t-0), \psi^0(t+0), v(t), h), t \in T_h. \end{aligned}$$

### Литература

1. *Альсевич В.В.* Оптимизация динамических систем с запаздыванием. Мн.: БГУ, 2000.
2. *Габасов Р.Ф., Кириллова Ф.М., Мансимов К.Б.* Необходимые условия оптимальности второго порядка (обзор). Препринт/ Ин-т математики АН БССР: №30 (155). Мн.: 1982.
3. *Понтрягин Л.С. и др.* Математическая теория оптимальных процессов. М.: Наука, 1976.
4. *Габасов Р.Ф., Кириллова Ф.М.* Принцип максимума в теории оптимального управления. Мн.: Наука и техника, 1974.
5. *Габасов Р.Ф. и др.* Методы оптимизации. Мн.: Четыре четверти, 2011.

## ПРОГНОЗИРОВАНИЕ КЛАССА РЕСПОНДЕНТА ПО РЕЗУЛЬТАТАМ ВЫБОРОЧНЫХ ОБСЛЕДОВАНИЙ

**И. С. Сугако, И. М. Коляго**

### МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И ПОСТАНОВКА ЗАДАЧИ

Одной из развивающихся областей применения методов эконометрического прогнозирования является социология. С их помощью часто

анализируют предпочтения потребителей, на основе которых решают задачи маркетинга. Это необходимо для увеличения долгосрочной прибыли, управления корпоративными функциями, снижения издержек на привлечение новых и уменьшения потери постоянных потребителей. Подобные исследования проводятся на основе анализа статистических данных, целью которого является оценка предпочтений респондентов.

В данной статье предоставляются результаты решения задачи классификации респондентов по категориям, характеризующим их предпочтения к магазинам некоторых сетей, на основе реальных данных с использованием логит-моделей множественного выбора и алгоритма «дерево решений».

Согласно [1], логит-модель множественного выбора для рассматриваемой задачи описывает зависимость вероятности отношения респондента к некоторому классу предпочтений  $p_i$  от включенных в модель факторов, задаваемых вектором  $x_i = (x_{i1}, \dots, x_{ik})^T$ , и определяется соотношением

$$p_i = P(y_i = 1) = F(x_i^T \beta). \quad (1)$$

Предполагается, что функция  $F(\cdot)$  известна и принимает значения на интервале  $(0,1)$ . Для логит-модели  $F(\cdot)$  – функция логистического распределения вероятностей:

$$F(u) = \frac{e^u}{1 + e^u}. \quad (2)$$

Интерпретация модели множественного выбора (1), (2) основана на использовании так называемой *латентной переменной*  $y_i^*$ , которая связана с вектором факторов  $x_i$  моделью множественной линейной регрессии:

$$y_i^* = x_i^T \beta + \xi_i, \quad i=1,2,\dots, n, \quad (3)$$

где  $\beta = (\beta_0, \beta_1, \dots, \beta_k)^T$  –  $(k+1)$ -мерный вектор неизвестных параметров,  $\xi_i$  – случайная ошибка наблюдения в  $i$ -ом эксперименте. Ошибки  $\{\xi_i\}$  ( $i=1,2,\dots, n$ ) являются независимыми в совокупности и одинаково распределенными случайными величинами с нулевым средним и постоянной дисперсией.

Заданы три класса магазинов, таким образом,  $y_i \in \{1, 2, 3\}$ . Номер класса зависит от некоторой латентной (скрытой) переменной  $y^*$  следующим образом:

$$y = 1, \text{ если } y^* \leq c_1, \\ y = 2, \text{ если } c_1 \leq y^* \leq c_2,$$

$$y = 3, \text{ если } y^* > c_2,$$

где  $c_1, c_2$  – некоторые фиксированные пороговые значения. Предполагая, что переменная  $y^*$  удовлетворяет уравнению (1), получаем:

$$\begin{aligned} P(y_i = 1) &= F(c_1 - x_i^T \beta), \\ P(y_i = 2) &= F(c_2 - x_i^T \beta) - F(c_1 - x_i^T \beta), \\ P(y_i = 3) &= 1 - F(c_2 - x_i^T \beta). \end{aligned} \quad (4)$$

Для произвольного числа альтернатив  $m$  логарифмическая функция правдоподобия для параметров  $\beta = (\beta_0, \beta_1, \dots, \beta_k)^T$  и пороговых значений  $c = (c_1, \dots, c_{m-1})^T$  по эмпирическим данным  $\{y_i, x_i\}$  ( $i=1, 2, \dots, n$ ) принимает вид:

$$l(\beta, c) = \sum_{j=1}^m \sum_{\{i: y_i=j\}} \ln(P(y_i = j)) = \sum_{j=1}^m \sum_{\{i: y_i=j\}} \ln(F(c_j - x_i^T \beta) - F(c_{j-1} - x_i^T \beta)).$$

Из условия максимума логарифмической функции правдоподобия  $l(\beta, c)$  с помощью численных методов находятся оценки параметров модели, а также пороговых значений.

Таким образом, упорядоченная логит-модель позволяет «отранжировать» классы потребительских предпочтений. Модель с неупорядоченными альтернативными вариантами рассчитывает дерево последовательного выбора альтернативы и процентные показатели вероятностей выбора каждого из магазинов для отдельных категорий респондентов.

## ПОСТРОЕНИЕ ПРОГНОЗНЫХ МОДЕЛЕЙ КЛАССИФИКАЦИИ

Для построения и тестирования прогнозных моделей классификации респондентов была использована выборка, содержащая реальные статистические данные по 976 респондентам.

При описании результатов моделирования используются следующие обозначения: Q1 – район проживания в г. Минске, Q2 – оценка работы универсамов в районе, Q3 – перспективы изменения ситуации с очередями, Q7\_3 – чистота, хороший интерьер, Q7\_8 – наличие акций, Q7\_9 – наличие кулинарии, Q7\_10 – культура персонала, приветливость, Q9 – соотношение суммы, которую тратит респондент на продукты питания, со средней суммой, Q10 – возраст, Q11 – образование, Q12 – социальное положение.

## ОПИСАНИЕ МОДЕЛЕЙ

Оцененная по имеющимся данным эконометрическая модель вида (1) описывается следующими соотношениями

$$y' = -0.079590q_1 - 0.173484q_2 + 0.358454q_3 - 0.240889q_7 - 0.125895q_9 - 0.278191q_{10} + 0.259773q_9 - 0.026764q_{10} + 0.250716q_{11} - 0.044762q_{12}$$

$$y_i = 1, \text{ если } y_i^* \leq -4.5988, y_i = 2, \text{ если } y_i^* \leq -2.166, y_i = 3, \text{ если } y_i^* > -2.166$$

Все коэффициенты статистически значимы на уровне 0.05. Гипотеза об адекватности модели в целом не отклоняется на основании критерия отношения правдоподобия,  $P$ -значение для  $LR$ -статистики меньше  $1 \cdot 10^{-7}$ . Значение статистики  $R_{MF}^2$  равно 0.42, для реальных разнотипных данных это хороший показатель, что свидетельствует в пользу построенной модели.

Наибольшее влияние на результат оказывают переменные Q3 – перспективы изменения ситуации с очередями, Q7\_10 – культура персонала, приветливость, Q9 – сумма, которую тратит респондент на продукты питания, Q7\_3 – интерьер. На основании полученных результатов можно утверждать, что магазина классов А, Б и В ранжируются следующим образом:

- магазин сети А – Эконом-класс;
- магазин сети Б – Средний класс;
- магазин сети В – Дорогостоящий.

Построенная методами многомерного статистического анализа неупорядоченная логит-модель имеет следующие параметры:

Табл 1

Универсам какой сети Вы хотите видеть возле дома?	Q3	Q7_3	Q7_10	Q9	Q10
А	-1,472	0,912	1,315	-0,202	0,927
Б	-,440	0,760	0,476	-0,107	0,684

Заданная эталонная категория: В, её коэффициенты равны 0

Выбор категории В в качестве эталонной обусловлен её меньшими размерами и тем, что она позволяет лучше улавливать влияние переменной Q9. Так же данная модель позволяет построить дерево последовательного выбора альтернативного варианта и описать для конкретной категории респондентов процентные показатели вероятностей. Наибольшее влияние на результат оказывают переменные Q3, Q7\_10, Q9

Все параметры моделей значимы на уровне 0.07, адекватность модели подтверждает статистика псевдо R-квадрат, равная 0, 42.

## ОЦЕНКА ПРОГНОСТИЧЕСКИХ СПОСОБНОСТЕЙ МОДЕЛЕЙ

Исследуем возможность прогнозирования классов респондентов на основе полученных моделей. Все они были проверены на обучающей и контрольной выборке. Упорядоченная логит-модель имеет 61,4 % правильных решений. Неупорядоченная логит-модель имеет 63,4 % правильных решений. Так же для сравнения с результатами неупорядоченной модели было построено дерево решений. Выделенные основные параметры, влияющие на выбор респондентом магазина одной из сетей, совпали с предыдущими результатами (Q7\_10, Q7\_3). Совпадает и общая классификация самих магазинов. Для деревьев решения общий процент корректных решений составил лишь 55,9 %.

## ЗАКЛЮЧЕНИЕ

Из анализа полученных прогнозных моделей классификации можно сделать выводы, что наиболее важными объясняющими переменными являются доход респондента, его отношение к очередям и две характеристики самого магазина: вежливость персонала и хороший интерьер. Интерпретация выделенных классов потребительских предпочтений:

- респондент, ориентированный на эконом-класс (сеть А);
- респондент, ориентированный на магазин выше среднего (сеть Б);
- респондент, ориентированный на дорогостоящий магазин (сеть В).

## Литература

1. *Малюгин, В. И.* Оценка устойчивости коммерческих банков на основе эконометрических моделей с дискретными зависимыми переменными / В. И. Малюгин, Е. В. Пытляк // *Банковский Вестник*. – 2007. – № 4(369). – С. 30–36.
2. *Малюгин В. И.* Методы построения и анализа скоринговых карт. Минск, БГУ 2011
3. *Айвазян С. А. и др.* Прикладная статистика: Основы моделирования и первичная обработка данных. Справочное изд. / С. А. Айвазян, И. С. Енюков, Л. Д. Мешалкин. – М.: Финансы и статистика, 1983.

## ИГРА СУДОКУ ДЛЯ ОС ANDROID

**В. Е. Федотов**

## ВВЕДЕНИЕ

На сегодняшний день мобильные приложения находятся на пике своей популярности. Растет количество разработчиков мобильных приложений, увеличивается количество доступных приложений, а также число