

УДК 656.015

ОБОСНОВАНИЕ ПРОДОЛЖИТЕЛЬНОСТИ НАБЛЮДЕНИЙ ЗА ПОВЕДЕНИЕМ ПАССАЖИРОВ ПРИ ВЫБОРЕ ПУТИ ПЕРЕДВИЖЕНИЯ

П.Ф. Горбачев, профессор, д.т.н., А.В. Макаричев, доцент, к.ф.-м.н.,
О.В. Свичинская, аспирант, ХНАДУ

Аннотация. Приводится обоснование целесообразности рассмотрения ситуации равно-возможного выбора пассажиром пути передвижения при моделировании такого выбора с помощью соответствующих математических моделей. Представлен подход к определению количества дней обследования фактического выбора пассажиром пути передвижения, достаточного для построения моделей предпочтений пассажиров.

Ключевые слова: модели дискретного выбора, адекватность, пассажир, путь передвижения, альтернатива, вероятность выбора пути, равновозможный выбор.

ОБҐРУНТУВАННЯ ТРИВАЛОСТІ СПОСТЕРЕЖЕНЬ ЗА ПОВЕДІНКОЮ ПАСАЖИРІВ ПРИ ВИБОРІ ШЛЯХУ ПЕРЕСУВАННЯ

П.Ф. Горбачов, професор, д.т.н., О.В. Макаричев, доцент, к.ф.-м.н.,
О.В. Свичинська, аспірант, ХНАДУ

Анотація. Наводиться обґрунтування доцільності розгляду ситуації рівноможливого вибору пасажиром шляху пересування при його моделюванні за допомогою відповідних математичних моделей. Представлено підхід до визначення кількості днів обстеження фактичного вибору пасажиром шляху пересування, достатнього для побудови моделей переваг пасажирів.

Ключові слова: моделі дискретного вибору, адекватність, пасажир, шлях пересування, альтернатива, ймовірність вибору шляху, рівноможливий вибір.

JUSTIFICATION OF PERIOD OF PASSENGER ROUTE CHOICE SURVEY

P. Gorbachov, Professor, Doctor in Technical Science, A. Makarichev, Associate
Professor, Candidate in Physico-Mathematical Science,
O. Svichinskaya, postgraduate, KhNAHU

Abstract. This paper presents the justification of reasonability of taking into account the situation of equally possible passenger route choice when modeling such a choice with proper mathematical models. The approach to determine the number of days to survey the factual passenger route choice, which will be enough to develop the choice models, is given.

Key words: discrete choice models, adequacy, passenger, route, alternative, route choice probability, equally possible passenger route choice.

Введение

Изучение предпочтений пассажиром параметров того или иного пути передвижения является важным этапом в транспортном планировании и должно основываться на проведении специальных обследований. Существующие методы проведения таких обследований, описанные в научной литературе

[1, 2], создают для исследователя возможность выявить факторы, влияющие на выбор пассажиром альтернатив. Однако их сложность и ограниченные прогностические возможности привели к тому, что широкого практического применения они не получили. Одной из причин этого является то, что моделирование выбора пассажиром пути передвижения всегда связано с субъективностью

оценки имеющихся альтернатив респондентами. Поэтому главной задачей моделирования остается обеспечение максимальной близости расчетной вероятности выбора пассажиром пути передвижения к фактической. И здесь очень важным является вопрос о количестве наблюдений за выбором альтернатив, которое обеспечивает высокую достоверность результатов моделирования.

Анализ публикаций

Среди существующих математических моделей, применяемых для определения вероятности выбора пассажиром пути передвижения в транспортной системе города, в наибольшей степени распространены модели дискретного выбора, разработанные зарубежными учеными [3, 4].

Различают бинарные и мультиномиальные модели в зависимости от количества рассматриваемых альтернативных вариантов пути передвижения [5]. Особенностью этих моделей является то, что при проведении обследования для их получения выбор пассажиром альтернативного пути передвижения фиксируется однократно для каждого респондента, причем выбранному пути принято отдавать полное предпочтение. То есть вероятность выбора пути передвижения принимается равной единице, если альтернатива выбрана, и равной нулю – в противном случае.

Цель и постановка задачи

Связь между вероятностями бинарного выбора p_1 и p_2 определяется зависимостью $p_1 + p_2 = 1$ и может быть представлена в виде отрезка прямой (рис. 1).

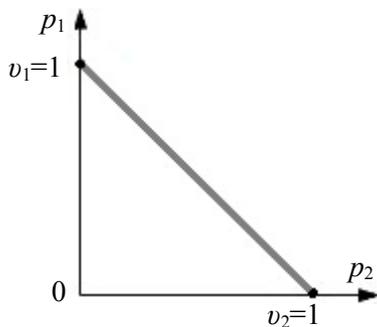


Рис. 1. Связь между вероятностями выбора при двух альтернативах

Точки данной прямой представляют собой множество возможных вероятностей бинар-

ного выбора $(p_1; p_2)$, то есть потенциальных результатов наблюдения за выбором альтернатив при количестве наблюдений $n \rightarrow \infty$.

Однако результаты однократного наблюдения за выбором могут дать только фактическую частоту выбора $v_1=1$ или $v_2=1$. То есть любая точка отрезка будет в результате обследования представлена одной из точек пересечения отрезка с осями координат. Единственным утверждением, которое достоверно можно сделать из результатов однократного наблюдения, является то, что при $v_i=1, p_i \neq 0$ и $p_{j \neq i} \neq 1$. В случае дискретного выбора при $v_i=1$ нельзя строго утверждать, что $p_i > p_{i \neq j}$, то есть результаты наблюдения вовсе не обязательно соответствуют ситуации более частого выбора. И это является одной из самых важных причин слабых прогностических способностей моделей дискретного выбора.

Для простейшего случая мультиномиального выбора из трех альтернативных путей передвижения связь между вероятностями их выбора p_1, p_2, p_3 , определяемую по зависимости $p_1 + p_2 + p_3 = 1$, можно представить в виде плоского треугольника (рис. 2).

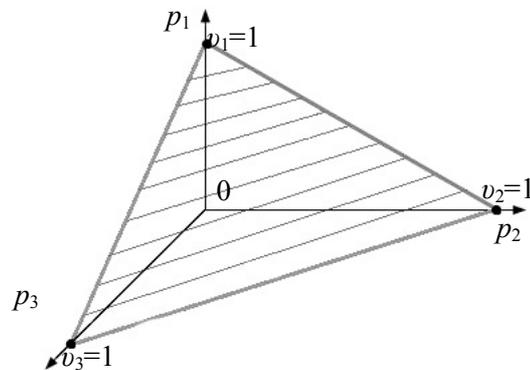


Рис. 2. Связь между вероятностями выбора из трех альтернатив

Как и в предыдущем примере, фактические вероятности выбора альтернатив могут являться любой точкой с координатами (p_1, p_2, p_3) в упомянутом треугольнике, что существенно увеличивает количество возможных исходов. Однако результатами фиксации дискретного выбора могут быть только три точки пересечения треугольника с осями координат, которые обладают такими же воз-

возможностями отображения фактических вероятностей, что и в бинарном случае. Таким образом, можно утверждать, что чем больше количество альтернатив, тем меньшей прогностической способностью будут обладать модели дискретного выбора.

Подобных недостатков лишены модели, относящиеся к классу нормировочных [6]. В этих моделях фактическая вероятность выбора пассажиром пути передвижения отображается его частотой, полученной в результате многократного наблюдения за выбором альтернативы одним респондентом. Однако здесь все еще открытым остается вопрос о количестве наблюдений, которое обеспечивает необходимую точность результатов прогнозирования.

Проверку моделей на адекватность целесообразно осуществлять на основе обоснованного в работе [7] критерия s_N^2 , который позволяет учесть отклонения расчетных значений вероятности выбора пассажиром пути передвижения от фактических.

Целью данной работы является обоснование количества наблюдений за поведением одного пассажира при выборе пути передвижения, с точки зрения критерия (статистики) s_N^2 .

Для достижения указанной цели необходимо обосновать решение поставленной задачи с позиции равновозможного выбора альтернативы пассажиром, определить условие, при котором гипотеза о пригодности модели равновозможного выбора по критерию s_N^2 не отвергается, определить границы целесообразного применения модели равновозможного выбора и рассчитать количество наблюдений, достаточное для построения адекватных моделей выбора пассажиром пути передвижения.

Обоснование области применения равновозможного выбора

Согласно методике проверки моделей выбора пассажиром пути передвижения на адекватность, представленной в работе [7], после этапа проведения обследования нужно провести предварительный расчет критерия s_N^2 для ситуации равновозможного выбора, для которого предполагается, что выбор пасса-

жиром альтернативы произведен без учета каких-либо факторов. Такой шаг необходим для проверки целесообразности моделирования выбора пассажиром пути передвижения с помощью различных математических моделей. Таким образом, способ проведения обследования должен сохранить равновозможный выбор в качестве альтернативы для более сложных моделей. А это будет справедливым только по отношению к критерию [7]

$$s_N^2 = \sum_{i=1}^N \frac{(v_1^{(i)} - P_1^{(i)})^2}{P_1^{(i)}(1 - P_1^{(i)})}, \quad (1)$$

где $v_1^{(i)}$ – фактическая (относительная) частота выбора i -м пассажиром первой альтернативы; $P_1^{(i)}$ – вероятность выбора i -м пассажиром первой альтернативы, рассчитанная по исследуемой модели; N – количество анкет обследования пассажира; n – количество принятых решений о выборе того или иного пути передвижения (количество дней обследования в анкете).

При использовании критерия s_N^2 необходимо определить некоторое пороговое значение a , при котором с ростом числа анкет обследования N до бесконечности гипотеза о равновозможности выбора пути передвижения не отвергается. Для начала требуется найти, какое наибольшее значение может принять a , чтобы выполнение неравенства

$$\frac{(v_1^{(i)} - P_1^{(i)})^2}{P_1^{(i)}(1 - P_1^{(i)})} \leq a, \quad i=1,2,3,\dots,N \quad (2)$$

при $N \rightarrow \infty$ не приводило бы к отвержению гипотезы о равновозможности. Данные неравенства можно назвать неравенствами о равномерной близости частот и вероятностей, позволяющих не отвергнуть гипотезу о равновозможности. Рассмотрим два случая: когда значение $a > 1$ и когда $a \leq 1$.

При $n \rightarrow \infty$ и $N \rightarrow \infty$ статистика s_N^2 асимптотически нормальна с параметрами $(N; 2N)$. Поэтому в первом случае, когда $a > 1$, максимальное значение статистики $s_N^2 = \max(s_N^2) = Na$, – с ростом числа анкет N к бесконечности превзойдет значение

$$N + \beta \sqrt{2N}, \quad (3)$$

где β – любое, но фиксированное число.

Действительно, их разность при $N \rightarrow \infty$ также стремится к бесконечности

$$Na - N + \beta\sqrt{2N} = N(a - 1) - \beta\sqrt{2N} \rightarrow \infty. \quad (4)$$

Это обусловлено тем, что выражение вида

$$N(a - 1) = Nb, \quad b = (a - 1), \quad b > 0, \quad (5)$$

отнесенное к произведению $\beta\sqrt{2N}$, с ростом количества анкет N к бесконечности также стремится к бесконечности

$$\frac{Nb}{\beta\sqrt{2N}} \rightarrow \infty \text{ при } N \rightarrow \infty. \quad (6)$$

Таким образом, при $a > 1$ с ростом количества анкет N к бесконечности величина s_N^2 , равная максимально возможному значению Na , уходит в критическую область, то есть в область отвержения ситуации равновозможного выбора при любом уровне значимости критерия.

В другом случае, когда каждое слагаемое статистики (2) не превосходит единицы

$$\frac{(v_1^{(i)} - P_1^{(i)})^2}{P_1^{(i)}(1 - P_1^{(i)})} \leq 1, \quad i=1,2,3,\dots,N, \quad (7)$$

суммарное значение этой статистики не превосходит значение N , равное математическому ожиданию случайной величины с распределением χ^2 с N степенями свободы. В такой ситуации гипотеза о равновозможности не отвергается при любом значении количества анкет N .

Определение необходимого количества наблюдений за выбором альтернативы

Используя вышеизложенный материал, рассмотрим ситуацию равновозможного выбора одной из двух альтернатив с вероятностями

$$P_1^{(i)} = \frac{1}{2}, \quad i=1, 2, \dots, N. \quad (8)$$

Тогда числитель нормировочного коэффициента из выражения (7) $P_1^{(i)}(1 - P_1^{(i)})$ для любого количества анкет $i = 1, 2, \dots, N$ будет равен

$$P_1^{(i)}(1 - P_1^{(i)}) = \frac{1}{2}\left(1 - \frac{1}{2}\right) = \frac{1}{4}. \quad (9)$$

После этого множество неравенств (3) о равномерной близости частот и вероятностей, позволяющих не отвергнуть гипотезу о равновозможности, можно записать в следующем виде

$$\frac{\left(v_1^{(i)} - \frac{1}{2}\right)^2}{\frac{1}{4n}} \leq 1, \quad i=1, 2, \dots, N. \quad (10)$$

Выражение (10) можно записать в более удобной форме

$$\left|v_1^{(i)} - \frac{1}{2}\right| \leq \frac{1}{2\sqrt{n}}, \quad i=1, 2, \dots, N. \quad (11)$$

Пользуясь неравенством (11), можно определить количество наблюдений, необходимое для построения модели выбора пассажиром пути передвижения. Для этого, последовательно изменяя количество принятых решений о выборе того или иного пути передвижения n в правой части неравенства, определяется значение максимального отклонения частоты от вероятности равновозможного выбора (рис. 3).

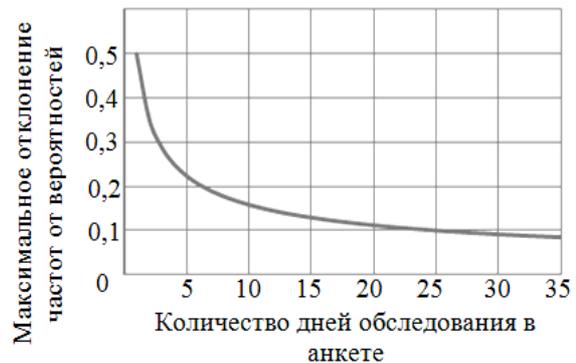


Рис. 3. Зависимость отклонения частот от количества дней обследования

При подстановке полученных отклонений в левую часть неравенства (11) будет получен диапазон частот выбора пути передвижения. Он определяется такими значениями n , при которых неравенство (11) выполняется и, соответственно, гипотеза о равновозможности не отвергается.

По графику видно, что оптимальным количеством наблюдений является 5 дней обследо-

вания. В этом случае, из всех возможных значений частот: 0, 0,2, 0,4, 0,6, 0,8, 1, выполнение неравенства (11) обеспечивают только значения 0,4 и 0,6. При этом обеспечиваются реальные максимальные отклонения от равновозможной вероятности выбора пути передвижения, равные 0,1 из заданного множества возможных частот.

Аналогичным образом получается и при количестве дней, равном $n = 10$. Значения частот v_i , обеспечивающие выполнение неравенства (12), равны 0,6; 0,5 и 0,4. Такой же результат дает количество дней, равное $n = 25$. В этом случае значения частот, наиболее отклоняющихся от вероятности равновозможного выбора, равной 0,5, снова равны 0,6 и 0,4. Представленные результаты показывают, что реальные значения частот, дающих реальные максимальные отклонения, равные 0,1, совпадают при количестве дней обследования 5, 10 и 25 и соответствуют значениям 0,6 и 0,4.

В дальнейшем, с увеличением количества дней обследования n , максимальные реальные отклонения уменьшаются. Например, при $n = 50$ максимальное отклонение частоты от вероятности равновозможного выбора равно 0,07, что соответствует реальным значениям частот 0,57 и 0,43, а при $n = 100$ максимальное отклонение равно 0,05, что соответствует реальным значениям частот 0,45 и 0,55. Все вышеизложенное говорит о том, что, с точки зрения критерия s_N^2 , увеличение продолжительности наблюдений более чем на 5 дней обследования нецелесообразно. Следовательно, при проведении обследования с целью фиксации фактического выбора пассажиром того или иного пути передвижения достаточно проводить пятидневное обследование для каждого пассажира.

Выводы

Наиболее распространенные на сегодняшний день модели дискретного выбора не обеспечивают адекватного описания реальных вероятностей использования пассажиром альтернативных вариантов передвижения из-за слишком малой продолжительности наблюдения. Уже незначительное, начиная с двух наблюдений, увеличение продолжительности обследования позволяет существенно уточнить результаты наблюдений, но при этом повышается вероятность отвержения гипоте-

зы о равновозможном выборе, что ухудшает условия оценки различных математических моделей выбора. Представленный подход к определению достаточного количества дней обследования фактического выбора пассажиром пути передвижения, с позиции критерия s_N^2 , позволил определить, что достаточным для построения пригодных для использования моделей выбора пути передвижения будет проведение пятидневного обследования.

Литература

1. Грановский Б.И. Моделирование пассажирских потоков в транспортных системах / Б.И. Грановский // Автомобильный и городской транспорт (Итоги науки и техники). – 1986. – Т. 11. – С. 67–105.
2. Рогова Г.Л. Моделирование выбора путей передвижения пассажиров в транспортных системах городов: автореф. дис. на соискание учен. степени канд. техн. наук: 05.22.02 «Транспортные системы городов и промышленных центров» / Г.Л. Рогова. – М., 1987. – 19 с.
3. Ben-Akiva Moshe E. Discrete Choice Methods and their Applications to Short Term Travel Decisions / Moshe E. Ben-Akiva, Michel Bierlaire // Handbook of Transportation Science. – 1999. – №23. – P. 5–33.
4. Ben-Akiva Moshe E. Discrete choice analysis. 1985, by The Massachusetts Institute of Technology. – P. 385.
5. Ortuzar J.D. Modelling Transport./J.D. Ortuzar, L.G. Willumsen.– Fourth Edition.– Chichester: John Wiley&Sons Ltd, 2011.– 586 p.
6. Заблоцкий Г.А. Методы расчета потоков пассажиров и транспорта в городах / Г.А. Заблоцкий; под ред. А.К. Старинкевич. – М. : ЦНТИ по гражд. строит. и архит., 1968. – 92 с.
7. Горбачев П.Ф. Подход к оценке адекватности моделей выбора пассажиром пути передвижения / П.Ф. Горбачев, О.В. Макаричев, О.В. Свичинская и др. // Вестник ХНАДУ: сб. науч. тр. – 2013. – Вып. 60. – С. 27–33.

Рецензент: Е.В. Нагорный, профессор, д.т.н., ХНАДУ.

Статья поступила в редакцию 21 января 2013 г.