

Горбачев Петр Федорович, Макаричев Александр Владимирович,  
Свичинская Ольга Владимировна

### **СПОСОБЫ ФОРМИРОВАНИЯ МОДЕЛИ ВЫБОРА ВАРИАНТА ТРУДОВОГО ПЕРЕДВИЖЕНИЯ МАРШРУТНЫМ ТРАНСПОРТОМ**

Рассматриваются вопросы использования моделей дискретного выбора для получения функции привлекательности пути следования в маршрутной системе городского пассажирского транспорта при интервальном способе обслуживания пассажиров. Предлагаются два варианта модели для обработки результатов обследования фактического выбора пассажирами пути следования. Показана перспективность использования регрессионного анализа для оценки коэффициентов модели привлекательности пути, полученной после преобразования частоты использования альтернатив.

Адрес статьи: [www.gramota.net/materials/1/2013/11/13.html](http://www.gramota.net/materials/1/2013/11/13.html)

**Статья опубликована в авторской редакции и отражает точку зрения автора(ов) по рассматриваемому вопросу.**

Источник

**Альманах современной науки и образования**

Тамбов: Грамота, 2013. № 11 (78). С. 47-58. ISSN 1993-5552.

Адрес журнала: [www.gramota.net/editions/1.html](http://www.gramota.net/editions/1.html)

Содержание данного номера журнала: [www.gramota.net/materials/1/2013/11/](http://www.gramota.net/materials/1/2013/11/)

**© Издательство "Грамота"**

Информация о возможности публикации статей в журнале размещена на Интернет сайте издательства: [www.gramota.net](http://www.gramota.net)  
Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: [almanac@gramota.net](mailto:almanac@gramota.net)

В результате исследований были определены рабочие составы керамзитобетона на мелком заполнителе из природного песка и различных зол для монолитного домостроения. Керамзитобетон готовится в условиях строительных площадок, где имеются турбулентные смесители роторного типа. Данный материал можно использовать в процессе изготовления ограждающих конструкций жилых, гражданских и промышленных зданий в г. Кемерово. Эффективность предлагаемой технологии состоит в применении доступных и относительно дешевых местных материалов.

*Список литературы*

1. ГОСТ 25820-2000. Бетоны легкие. Технические условия. М.: Стандартинформ, 2001.
2. СНиП 5.01.23-83. Типовые нормы расхода цемента для приготовления бетонов сборных и монолитных бетонных, железобетонных изделий и конструкций. М.: Стройиздат, 1985.

**STUDY OF CERAMSITE CONCRETE COMPOSITIONS FOR MONOLITHIC HOUSE BUILDING**

Gilyazidinova Natal'ya Vladimirovna, Ph. D. in Technical Sciences  
Santalova Tat'yana Nikolaevna  
Rudkovskaya Nadezhda Yur'evna  
Kuzbas State Technical University  
spien2009@rambler.ru

The problem of industrial ash wastes utilization, which cover great ground areas and damage environment is an urgent scientific task. The purpose of this research is to study the selection of the composition of structural and structural and heat-insulating ceramsite concrete with fine aggregate of natural sand, ash-carryover of dry selection and hydro-disposal of B3,5 – B7,5 sort with F35 – F50 freeze-thaw durability grade for solid house building.

*Key words and phrases:* compositions of light concretes; strength properties; structural and heat-insulating ceramsite concrete; ash-carryover of dry selection and hydro-disposal; solid house building.

УДК 656.072

**Технические науки**

*Рассматриваются вопросы использования моделей дискретного выбора для получения функции привлекательности пути следования в маршрутной системе городского пассажирского транспорта при интервальном способе обслуживания пассажиров. Предлагаются два варианта модели для обработки результатов обследования фактического выбора пассажирами пути следования. Показана перспективность использования регрессионного анализа для оценки коэффициентов модели привлекательности пути, полученной после преобразования частоты использования альтернатив.*

*Ключевые слова и фразы:* пассажир; модели дискретного выбора пути следования; альтернативный вариант пути; привлекательность передвижения; параметры пути следования; метод максимального правдоподобия; регрессионный анализ.

**Горбачев Петр Федорович**, д.т.н., профессор

**Макаричев Александр Владимирович**, к. ф.-м. н., доцент

**Свичинская Ольга Владимировна**

Харьковский национальный автомобильно-дорожный университет, Украина  
gorbachev\_pf@mail.ru; amakarichev@mail.ru; svichinskaya@ukr.net

**СПОСОБЫ ФОРМИРОВАНИЯ МОДЕЛИ ВЫБОРА ВАРИАНТА  
ТРУДОВОГО ПЕРЕДВИЖЕНИЯ МАРШРУТНЫМ ТРАНСПОРТОМ<sup>©</sup>**

Моделирование поведения пассажиров при выборе пути следования является одной из наиболее наукоемких задач в сфере транспортного планирования. Методика ее решения с помощью моделей дискретного выбора [10] стала основным инструментом определения вида функции полезности не только на транспорте. Она используется и в других сферах человеческой деятельности, в которых субъекты сталкиваются с необходимостью выбора лучшего варианта из имеющегося набора альтернатив [1].

Однако считать, что существующие методики транспортного моделирования с помощью специальных программ обеспечивают высокую точность прогнозирования транспортных процессов еще нельзя. Несмотря на наличие инструментов выполнения каждого этапа транспортного моделирования, любую модель необходимо подгонять (калибровать) под существующие транспортные или пассажирские потоки.

Конечно же, все проблемы начинаются с моделей спроса на передвижение, так как проверить их точность практически невозможно из-за отсутствия фактических матриц корреспонденций. Но и модели распределения пассажиропотоков вносят немалый вклад в погрешности прогнозирования. С учетом того, что количество переменных в задачах транспортного планирования значительно больше количества наблюдаемых потоков, любое из полученных с помощью калибровки решений не может считаться единственным и правильным. А использование таких решений в транспортном планировании не позволяет гарантировать объективности оценки результатов реализации мероприятий по развитию транспортных систем.

Спасают в этой ситуации отсутствие общепринятых критериев оценки точности транспортных моделей и чрезвычайно высокая устойчивость транспортных систем городов и регионов, обусловленная их социальным характером. Да и затратный характер большинства мероприятий также обеспечивает позитивные результаты их реализации уже в постановке задачи, ведь нужно серьезно постараться, чтобы капитальные вложения в транспорт дали бы отрицательные результаты.

Но эти особенности региональных систем пассажирского транспорта не отрицают необходимости объективной оценки эффективности их функционирования, что делает актуальным исследование возможностей современных моделей прогнозирования поведения субъектов транспортного рынка с помощью моделей дискретного выбора и разработку новых методов моделирования.

### 1. Анализ публикаций

Постановка задачи в подходе «моделирования поведенческого спроса» построена на гипотезе, согласно которой: а) каждый пассажир, выбирая  $k$ -й путь, максимизирует свою индивидуальную субъективную привлекательность  $U_l$ ; б) существуют однородные группы пассажиров с одинаковым отношением к измеримым параметрам пути [4].

$$k = l, \text{ при } U_l = V^h(\bar{X}_l) + \varepsilon_l \rightarrow \max, \quad (1)$$

где  $U_l$  – индивидуальная субъективная привлекательность  $l$ -го варианта пути;  $V^h(\bar{X}_l) = V_l$  – привлекательность  $l$ -го варианта пути следования для группы  $h$ , которая является функцией измеримых параметров  $\bar{X}_l$  (времени пеших подходов, общего времени и денежной стоимости передвижения);  $\varepsilon_l$  – случайный фактор, на величину которого влияют неправильные представления пассажиров о параметрах передвижения и параметры пути следования, которые нельзя измерить.

Следует отметить, что упомянутый подход изначально применялся в области биологических исследований [3; 9] и был заимствован из микроэкономики, поэтому в большей степени соответствует поведению покупателя товара, чем пассажира в системе городского пассажирского транспорта (ГПТ), что приводит к неполному соответствию выделенной случайной величины  $\varepsilon_l$  реальностям транспортного процесса. Кроме перечисленных составляющих  $\varepsilon_l$ , в транспортных процессах случайными также являются многие другие факторы, например, время ожидания и степень заполнения маршрутных транспортных средств, время поездки и расход топлива на индивидуальном транспорте. Эти факторы поддаются статистической оценке, для них может быть определен вид закона распределения. Однако в модели они все равно не выделяются из общей неопределенности, по-видимому, из-за сложностей в поиске коэффициентов значимости каждого фактора.

В соответствии с (1) можно записать уравнение, определяющее вероятность выбора  $k$ -го пути  $P_k$ :

$$P_k = P\{U_k \geq U_l\} = P\{\varepsilon_k \geq \varepsilon_l + V_l - V_k\}, l = 1, \dots, J, \quad (2)$$

где  $J$  – количество альтернативных вариантов передвижения.

С учетом независимости случайных составляющих, по формуле полной вероятности для непрерывных случайных величин, далее записывается [4]:

$$P_k = \int_{-\infty}^{\infty} f(\varepsilon_k) \cdot \prod_{l \neq k}^J F(\varepsilon_l + V_l - V_k) d\varepsilon_k, \quad (3)$$

где  $f(\varepsilon_k)$  – плотность распределения случайной составляющей полезности альтернативы  $k$ ;

$F(\varepsilon_l + V_l - V_k)$  – функция распределения случайной величины  $\varepsilon_l + V^h(\bar{X}_l) - V^h(\bar{X}_k)$ .

Затем делается предположение о виде закона распределения случайной составляющей полезности альтернатив. Для logit-модели принимается, что  $\varepsilon_l$  распределена в соответствии с законом Гумбеля, который описывает распределение экстремальных значений 1-го типа. В probit-модели случайная составляющая распределена по нормальному распределению Гаусса.

В первом случае интеграл (3) может быть решен в явном виде, что в случае бинарного выбора дает:

$$P_k = \frac{1}{1 + e^{V_l - V_k}}, J = 2. \quad (4)$$

Далее, для выделения в зависимости привлекательностей отдельных альтернатив он записывается в преобразованном варианте

$$P_k = \frac{e^{V_k}}{e^{V_k} + e^{V_l}}, \quad J = 2. \quad (5)$$

Обобщение logit-модели до множественного выбора достигается достаточно просто:

$$P_k = \frac{e^{V_k}}{\sum_{l=1}^J e^{V_l}}, \quad (6)$$

а для поиска коэффициентов значимости параметров в функциях  $V_l = f(\bar{X}_l)$  и  $V_k = f(\bar{X}_k)$  используется метод максимального правдоподобия [7].

В probit-модели интеграл (3) не имеет решения в явном виде, поэтому коэффициенты при параметрах пути передвижения отыскиваются с помощью метода Монте-Карло [4].

Logit-модели получили не только чрезвычайно широкое распространение в практических приложениях, но также и обширное наследие в форме различных усложненных вариантов модели, полученных для учета в ней дополнительных факторов или условий. Хороший обзор потребительских качеств большинства вариантов модели дискретного выбора содержится в работе [8], там же отмечается высокая чувствительность logit-модели к значениям параметров альтернатив. Из-за незначительных, с практической точки зрения, колебаний даже одного из параметров расчетная вероятность выбора альтернатив может изменяться очень существенно.

Это объясняется видом расчетной формулы для вероятности выбора альтернативы (4) или (5), в которой измеримая привлекательность является аргументом показательной функции, т.е. показателем степени. Из этого и других источников, например [3], также следует, что logit-модель обладает весьма невысокими прогностическими способностями. Это обусловлено заданной и жесткой структурой модели, которая получена из допущения об экстремальном распределении случайной составляющей.

Но так как в преобразовании от (3) к (4) подразумевается, что все случайные факторы выбора уже описаны, очевидно, что связь между результирующим признаком и независимыми переменными должна носить функциональный характер. Но результатов исследований, соответствующих этому очевидному факту, в известных источниках не приводится. Это означает, что связь между частотой использования альтернативы и ее параметрами как минимум может быть обусловлена и другими случайными факторами, не учтенными в  $\varepsilon_l$ . Это, безусловно, нужно принимать во внимание при формировании модели дискретного выбора, однако в постановке задачи моделирования дискретного выбора упоминания о наличии неучтенных факторов не встречаются. Аналитическая зависимость в моделях дискретного выбора не может быть построена, а прогностические способности logit-моделей не имеют количественной оценки и зачастую оставляют желать лучшего.

Существует еще одно ограничение на использование моделей дискретного выбора, обусловленное необходимостью рассмотрения упорядоченного набора альтернатив для всего набора передвижений. Эта особенность приводит к тому, что в них может быть рассмотрена только ситуация модального выбора между различными способами реализации потребности в передвижении, например, пешком, на автомобиле или общественным транспортом. Для решения задачи выбора рационального варианта пути следования в маршрутной системе с разным количеством и содержанием реальных альтернатив модели дискретного выбора в стандартном их понимании использоваться не могут.

Перечисленные выше умозрительные соображения заставили нас в процессе научной деятельности по разработке методик транспортного планирования в условиях Украины глубже разобраться в описательных и прогностических возможностях моделей дискретного выбора и предложить свой вариант моделирования поведения пассажиров в маршрутной системе города.

## 2. Сравнительная характеристика вариантов модели

В качестве основы для сравнения принципов моделирования выбора пассажиром пути следования были приняты существовавшие до появления моделей дискретного выбора способы выравнивания стандартного набора данных. Наиболее распространенным из них является регрессионный анализ, который уже давно доказал свою состоятельность как в прогностическом, так и в аналитическом смысле. В обычном виде он, безусловно, не может быть использован для оценки параметров модели выбора альтернатив, построенной на основе известных частот их использования. Однако можно сравнить цели и принципы построения регрессионных моделей и моделей дискретного выбора для оценки перспектив развития последних.

Регрессионные модели предназначены для внешнего описания связей между зависимым и независимыми векторами. Они используются тогда, когда реальные причинно-следственные связи между результирующим и факторными признаками не могут быть описаны аналитически. Такая ситуация может быть обусловлена слишком сложной структурой связей или большим количеством влияющих на результирующий признак факторов. На основе ряда наблюдений за всеми признаками строится регрессионная зависимость, описывающая корреляционное (стохастическое) влияние независимых факторов на результирующий признак. Коэффициенты уравнения регрессии оцениваются с помощью метода наименьших квадратов, в котором решение системы нормальных уравнений минимизирует длину вектора невязки расчетных и фактических значений результирующего признака. Вид регрессионного уравнения выбирается исследователем и ничем не ограничен, нелинейность искомых связей может быть учтена с помощью обычных преобразований исходных векторов.

При наличии одного или нескольких измеримых факторных признаков, оказывающих достаточно весомое влияние на результирующий признак, может быть получена регрессионная зависимость, которая устраивает исследователя по уровню точности. Требуемый уровень точности определяется задачами исследования. При прогнозировании результата с помощью регрессионной зависимости получают его точечные оценки, имеющие некоторую дисперсию. Влияние неучтенных факторов или неописанных связей оценивается остаточной дисперсией результирующего признака. Для количественной оценки точности построенной зависимости делается предположение о нормальности распределения остатков регрессии, которое выглядит весьма правдоподобным с учетом метода оценки коэффициентов регрессии.

Случай с выбором варианта пути следования (или выбора товаров) полностью соответствует условиям использования регрессионного анализа, за одним лишь исключением. Здесь результирующий признак, который представляет собой полезность (или привлекательность) альтернативы, не может быть зафиксирован в процессе наблюдений. Вместо него фиксируется частота использования альтернативы, и целью моделирования является адекватное описание исходных частот расчетными вероятностями использования альтернатив. То есть привлекательность передвижения является промежуточным, а не окончательным результатом моделирования. При этом ее значение является очень важным для анализа изучаемых процессов, и именно она становится основным предметом глубокого изучения в большинстве работ, посвященных моделям дискретного выбора [5].

После однократного наблюдения за выбором большого количества субъектов, в результате которого определяется частота выбора альтернативы пути, равная 0 либо 1, устанавливается связь между этой частотой и результирующей функцией привлекательности на основе сравнения случайных составляющих привлекательности альтернатив. Это сильно отличает моделирование дискретного выбора от регрессионного анализа, в котором случайной считается неописанная моделью дисперсия результата.

Формирование вектора результата через уравнение (3) делает случайную составляющую  $\varepsilon_i$  основой для определения коэффициентов создаваемой модели, а результирующий вектор привлекательности при этом, в принципе, может вообще не выделяться. Такое использование  $\varepsilon_i$  предъявляет очень серьезные требования к степени соответствия закона распределения, используемого для ее описания, реальности. К сожалению, проверить это весьма проблематично из-за отсутствия наблюдаемых значений и очень широкого толкования физического смысла этой случайной величины.

Для поиска коэффициентов в logit-модели используется метод максимального правдоподобия, который разработан исключительно для определения параметров специально выбранной из определенного класса теоретической функции распределения случайной величины на основе известной выборки [7]. Он дает состоятельные оценки искомых параметров на основе максимизации произведения вероятностей в точках, где зафиксированы наблюдения. Однако пригодность этого метода для оценки параметров logit-модели требует дополнительной проверки, поскольку условия его использования здесь существенно отличаются от тех, для которых он разрабатывался.

Но метод получения оценок, в принципе, может быть любым. Важно только, чтобы он обеспечивал достаточную точность оценки параметра. И тут кроется еще одно из отличий между рассматриваемыми подходами к моделированию корреляционных связей. В моделях дискретного выбора для оценки точности полученных результатов используется та же функция максимального правдоподобия, что и в методе оценки параметров (*Likelihood Ratio Index*, *Pseudo R<sup>2</sup>* и др.). Но она далеко не полностью описывает степень близости искомых частот и расчетных вероятностей, так как не ставит своей непосредственной целью достижение их соответствия друг другу.

Если бы такой подход использовался в регрессионном анализе, то для оценки точности полученной модели было бы нужно привлекать систему нормальных уравнений, а не остаточную дисперсию. Так как в числе уже известных методов оценки такого варианта нет, тут лучше провести аналогию с оценкой соответствия эмпирической выборки теоретическому закону распределения. В этом случае, после оценки параметров теоретического закона с помощью какого-то метода (часто метода максимального правдоподобия), рассчитывается мера расхождения эмпирических и теоретических данных. Эта мера никак не связана с методом оценки параметров распределения, но важно, чтобы она, как случайная величина, имела известные параметры распределения и соответствовала целям оценки. После расчета вероятности появления выявленных отклонений можно сделать вывод о степени соответствия эмпирической выборки теоретическому закону распределения. И нужно отметить, что метод максимального правдоподобия разработан для оценки параметров именно в таком случае, и здесь он обеспечивает минимум общеизвестных мер отклонения эмпирических частот от теоретического закона распределения.

При создании моделей дискретного выбора степень соответствия исходных частот и расчетных вероятностей непосредственно не оценивается, поэтому проверка состоятельности результатов оценки параметров методом максимального правдоподобия является очень важной.

Таким образом, можно заключить, что при прочих равных условиях невозможность наблюдать значения результирующего признака в моделях дискретного выбора привела к их кардинальным отличиям от регрессионного анализа на всех этапах моделирования и значительно сократила возможности исследователя по выбору вида связи между изучаемыми векторами. Такие отличия были бы оправданными, если бы обеспечивали наилучшие результаты моделирования, но это требует серьезной проверки.

Для проведения такой оценки необходим критерий, характеризующий меру отклонения наблюдаемых частот и расчетных вероятностей выбора альтернативного варианта пути. Такого общепринятого критерия в математической статистике еще нет, так как задача носит не совсем тривиальный характер. Необходимо оценить отклонения между частотой и вероятностью для большого количества наблюдений (ситуаций выбора), в каждом из которых существует полная группа событий, суммарная вероятность которых всегда равна 1. Больше всего к поставленной задаче подходит критерий Пирсона [6], но его необходимо несколько видоизменить:

$$s_N^2 = \sum_{i=1}^N \left[ \frac{(v_1^{(i)} - P_1^{(i)})^2}{\left(\frac{P_1^{(i)}}{n_i}\right)} + \dots + \frac{(v_{J_i}^{(i)} - P_{J_i}^{(i)})^2}{\left(\frac{P_{J_i}^{(i)}}{n_i}\right)} \right], \quad (7)$$

где  $s_N^2$  – мера отклонения наблюдаемых частот и расчетных вероятностей выбора (критерий точности модели), асимптотически имеющая распределение  $\chi^2$ ;  $N$  – количество респондентов;  $v_j^{(i)}$  – зафиксированная частота выбора альтернативы  $j$  индивидуумом  $i$ ,  $j=1,2,\dots,J$ ;  $P_j^{(i)}$  – расчетная вероятность выбора альтернативы  $j$  индивидуумом  $i$ ;  $n_i$  – количество наблюдений за выбором  $i$ -го индивидуума в одних условиях.

Этот критерий приведен для общего случая, не только дискретного выбора. Многократное наблюдение за одним и тем же выбором позволяет получать оценки вероятности выбора альтернативы, отличные от 0 и 1, и эти оценки будут приближаться к искомой вероятности с увеличением  $n_i$ . В случае моделей дискретного выбора количество наблюдений за каждым выбором всегда равно одному.

Величина  $s_N^2$  является результатом суммирования независимых, центрированных и нормированных ошибок, для которых корректным выглядит допущение о нормальности их распределения. Поэтому корректным будет и допущение о том, что она распределена в соответствии с законом  $\chi^2$  и количеством степеней свободы

$$n_\chi = \sum_{i=1}^N J_i - N - n_p, \quad (8)$$

где  $n_p$  – количество аргументов в функции привлекательности альтернативы.

Эта величина полностью совпадает с прогностической целью построения модели и соответствует всем иным метрикам, которые могут использоваться на этапе выбора критерия для оценки точности построенной модели.

Для оценки результатов моделирования вероятности выбора альтернатив с помощью модели дискретного выбора был выбран простейший пример бинарного выбора из [3]. Для него сначала было рассчитано значение критерия  $s_N^2$ , затем простым подбором был определен параметр, обеспечивающий минимум этого критерия (Табл. 1). Тут же, для сравнения, приведена величина ошибки для равновероятной модели с постоянной вероятностью 0,5. Здесь же содержится значение вероятности того, что случайная величина, распределенная по  $\chi^2$ , превысит значение  $s_N^2$  для  $n_\chi = 2$ .

**Таблица 1.** Оценка точности дискретной модели выбора пути следования

Номер наблюдения	Альтернатива	Параметр поездки	Частота выбора альтернативы	Вероятность выбора для logit-модели		Равновероятный выбор альтернативы
				максимальное правдоподобие	подбор параметра	
1	1	5	1	0,82	0,73	0,5
	2	3	0	0,18	0,27	0,5
2	1	1	1	0,32	0,38	0,5
	2	2	0	0,68	0,62	0,5
3	1	3	0	0,32	0,38	0,5
	2	4	1	0,68	0,62	0,5
Оценка параметра $\beta$ для logit-модели				0,756	0,419	-
Значение критерия $s_N^2$				2,815	2,623	3
Вероятность превышения критерия				0,24	0,27	0,39

Вследствие малого объема наблюдений в примере, все полученные вероятности имеют весьма высокие значения, не дающие оснований отвергать гипотезу о соответствии исходных частот расчетным вероятностям. Этим же объясняется и самое высокое значение вероятности для простейшей модели, т.к. добавление еще одной степени свободы для нее (6 вместо 5, за счет  $n_p = 0$ ) приводит к резкому изменению вероятности по распределению  $\chi^2$ .

Однако по этим результатам можно заметить, что метод максимального правдоподобия обеспечивает для данного случая оценку параметра по критерию  $s_N^2$ , незначительно лучшую, чем равновероятная модель. И конечно, эти результаты свидетельствуют не в его пользу, если только не считать применение метода максимального правдоподобия самоцелью построения модели. Простой подбор параметра модели в большей степени приближает результаты моделирования к исходным частотам и обеспечивает более весомое отличие по значению  $s_N^2$  от равновероятной модели. Это свидетельствует о том, что в данном случае метод максимального правдоподобия не обеспечивает максимальной близости эмпирических и теоретических вероятностей. Зависимость функции максимального правдоподобия от параметра модели сопоставлена с аналогичной зависимостью для  $s_N^2$  на Рис. 1.

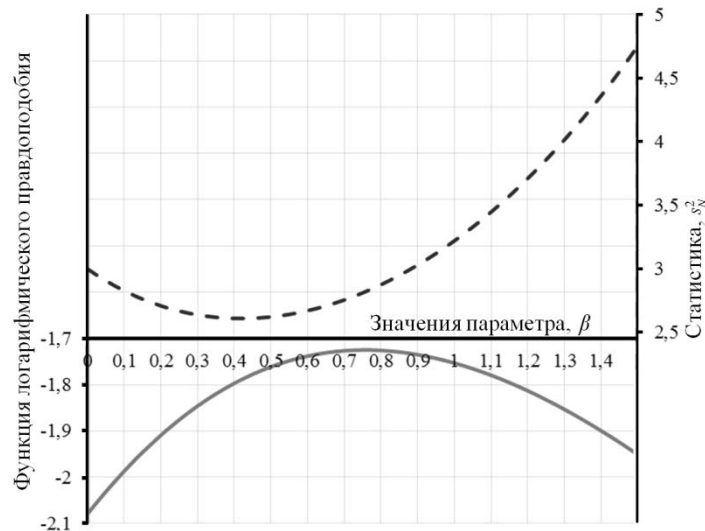


Рис. 1. Зависимость альтернативных оценок от величины параметра функции

Как видно, это две совершенно разные функции, со смещенными друг относительно друга оптимумами. Даже если не считать  $s_N^2$  достаточно убедительным показателем соответствия эмпирических частот и теоретических вероятностей, этот график как минимум порождает сомнения в применимости метода максимального правдоподобия для оценки параметров модели дискретного выбора. Ведь функция максимального правдоподобия непосредственной связи с точностью оценки эмпирической частоты вообще не имеет.

### 3. Постановка задачи

Результаты этой проверки открывают возможности для создания новых моделей оценки вероятностей выбора альтернатив с учетом их зависимостей от тех или иных факторов. Коэффициенты данных моделей должны рассчитываться так, чтобы минимизировать разницу между теоретической вероятностью выбора альтернативы и эмпирической частотой ее выбора. Поэтому дальнейшие исследования были направлены на создание альтернативных методов моделирования поведения субъектов при выборе вариантов передвижения. В связи со спецификой научной деятельности авторов и транспортного рынка Украины, а также вследствие ограниченности объема финансирования исследований единственным объектом внимания в данной работе стало поведение горожан при выборе варианта поездки из дома на работу на маршрутном транспорте во втором по величине городе Украины – Харькове.

Вначале была сделана попытка глубже разобраться в моделях выбора альтернатив, основанных на сознательном принятии пассажирами решений относительно лучшего варианта пути с учетом случайного характера полезности альтернатив.

Уровень информационного обеспечения пассажиров маршрутного транспорта украинских городов и способы организации работы маршрутов в подавляющем большинстве случаев делают время прибытия транспортных средств на остановочный пункт неизвестным пассажиру. В то же время количество транспортных средств на маршрутах достаточно велико для того чтобы считать вероятность отказа пассажирам в посадке незначительной. В этом Харькову сильно помогают метрополитен и недавний переход на обслуживание пассажиров автобусами большой вместимости при сохранении количества транспортных средств на маршрутах. Остальные составляющие затрат на поездку хорошо известны пассажирам, совершающим трудовое передвижение.

Поэтому время ожидания пассажирами транспорта на остановочном пункте вносит основную долю неопределенности в общие затраты на передвижение. Такая ситуация позволила выделить время ожидания в качестве случайной составляющей затрат. При этом в отличие от моделей дискретного выбора здесь считается, что в оставшейся части полезности могут существовать другие случайности. Поэтому аналитическое

решение в общем случае получено быть не может, а для оценки коэффициентов значимости факторов выбора необходимо использовать какой-то из методов выравнивания данных.

Вследствие этого возникла необходимость глубже разобраться в сути уравнений, связывающих вероятность выбора альтернативы со случайной и детерминированной составляющими полезности передвижения.

В уравнении (3) считаются заданными частота использования альтернативы, а также вид и параметры закона распределения случайного элемента. Неизвестной является разность двух детерминированных величин полезности альтернатив

$$v = V_l - V_k. \quad (9)$$

Эта величина имеет размерность случайного элемента полезности и представляет собой такое смещение одного распределения относительно другого по оси  $\varepsilon$ , которое обеспечивает в правой части уравнения (3) значение, равное  $P_k$ .

Можно легко показать, что при одинаковых параметрах распределения  $\varepsilon_l = \varepsilon_k = \varepsilon$  и  $v = 0$ ,  $P_k = f(v, \varepsilon) = 0,5$ . Увеличение величины  $v$  в (3) сдвигает вправо распределение случайной величины  $\varepsilon_l + V_l - V_k$ , уменьшая тем самым значение  $P_k$ . Иллюстрация такого решения для  $\varepsilon_k$  и  $\varepsilon_l$ , распределенных по закону Гумбеля с единичной дисперсией и  $v = 1$ , приведена на Рис. 2.

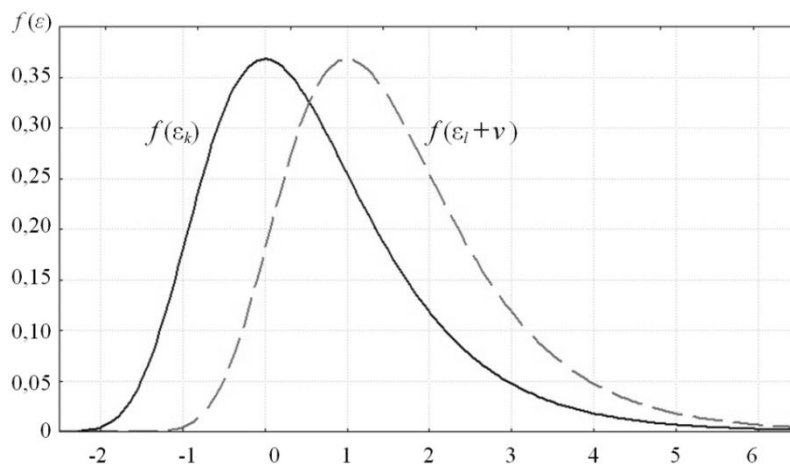


Рис. 2. Пример решения уравнения (3) для двух случайных величин

Такой сдвиг случайной составляющей альтернативы  $l$  приводит вероятность выбора альтернативы  $k$  к значению  $P_k = 0,269$ . Собственно, это легко вычислить из уравнения (4), подставив в него (9).

$$P_k = \frac{1}{1 + e^v}. \quad (10)$$

График соответствующей функции, представляющей собой решение уравнения (3) для двух случайных величин, представлен на Рис. 3.

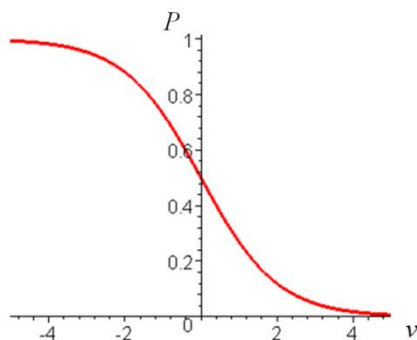


Рис. 3. Решение уравнения (3) для двух случайных величин, распределенных по закону Гумбеля с единичной дисперсией

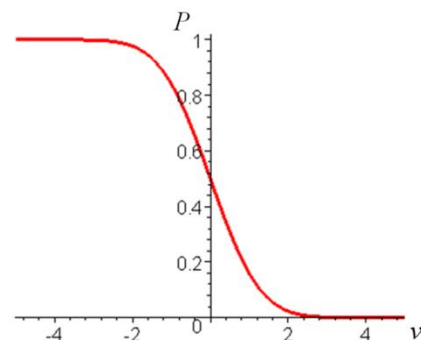


Рис. 4. Решение уравнения (3) для двух нормально распределенных нормированных случайных величин

Не представляет особых трудностей, без использования метода Монте-Карло, получение подобной функции для нормального распределения (Рис. 4).



Порядок альтернатив для решения не имеет значения при любом количестве альтернатив, что будет показано ниже. Для случая с двумя альтернативами это проявляется в относительности сдвига случайных составляющих полезностей. Сдвиг случайной составляющей альтернативы  $l$  в другую сторону, на  $v = -1$ , приводит уже вероятность выбора этой альтернативы к значению  $P_{l \neq k} = 0,269$ . То есть одна пара альтернатив в постановке (2) предоставляет информацию о взаимном расположении плотностей вероятности без привязки к оси  $\varepsilon$ .

Таким образом, абсолютные значения детерминированных полезностей  $V_k$  и  $V_l$  не могут быть определены на основе информации о вероятностях выбора альтернатив из одного набора. Для того чтобы найти одно значение детерминированной полезности, например  $V_k$ , необходимо задать другую полезность  $V_l$ . С практической точки зрения это свойство обусловлено относительным характером полезности альтернатив в рамках одного набора. Добавление в массив данных каждого нового набора альтернатив приводит к увеличению количества свободных переменных на одну. Это свойство также проявляется и в уравнении (5), так как из него следует, что

$$\frac{e^{V_l}}{P_l} = \frac{e^{V_k}}{P_k} \Rightarrow e^{V_k} = e^{V_l} \frac{P_k}{P_l}. \quad (11)$$

Для определения значений этих свободных переменных требуется специальный механизм. Однако в моделях дискретного выбора этой особенности не уделяется особого внимания, она просто обозначается как проблема корреляции альтернатив [5]. А преодолевается эта проблема благодаря вводу дополнительных связей между альтернативами через все ту же функцию максимального правдоподобия.

Также здесь необходимо отметить, что при любом виде распределения случайных величин  $\varepsilon_k$  и  $\varepsilon_l$  дискретные результаты выбора, то есть  $P_k = 1$  и  $P_l = 0$ , создают проблемы в аналитическом решении уравнения (3). При одно- или двухсторонней бесконечности распределений упомянутые вероятности без дополнительных условий могут быть получены только при  $v = \infty$  или  $v = -\infty$ . Такое решение не имеет практического значения, так как не позволяет определить коэффициенты функции полезности. При двухсторонней ограниченности распределений вариантов решения будет бесконечно много. Сочетание односторонне бесконечного и ограниченного распределений будет приводить к одной из описанных выше ситуаций в зависимости от результатов выбора. Это свойство также хорошо видно и в уравнениях (4)-(6), в которых  $P_k = 1$  или  $P_l = 0$  аналитически могут быть получены только при функциях полезности, равных  $+\infty$  или  $-\infty$ , что вызвано бесконечными пределами распределения Гумбеля.

В практическом смысле такая ситуация обусловлена тем, что частота использования альтернатив представляет лишь относительную оценку их привлекательности (11). Чем больше частота использования, тем выше привлекательность альтернативы. Если за период наблюдения из числа альтернатив использовалась только одна, это значит, что она оказалась бесконечно лучше остальных альтернатив, что и отражается в описанном выше свойстве уравнения (3).

Кроме того, дискретные частоты являются далеко не самой лучшей оценкой фактических вероятностей выбора альтернатив. По ним с уверенностью можно сказать только то, что вероятность реализованной альтернативы не равна 0, а нереализованных – не равна 1. В случае фиксации фактического выбора альтернатив дискретная оценка вполне может привести к результатам, противоположным реальным вероятностям выбора.

Проблему несоответствия левой и правой частей уравнений (3)-(6) в моделях дискретного выбора решает все тот же метод максимального правдоподобия. Однако его точность вынуждает к поиску других способов оценки коэффициентов. А неудовлетворительная достоверность оценки вероятностей выбора альтернатив дискретными частотами привела к решению, что строить модель целесообразно по той части наблюдений, где  $0 < P_k < 1$ .

#### 4. Разработка моделей вероятности выбора пути следования при трудовых передвижениях маршрутным транспортом

Вначале была сделана попытка детализации моделей дискретного выбора для случая, когда рассматривается не модальный выбор, а ситуация с выбором маршрута следования без привязки к видам городского пассажирского транспорта. Получение вида функции полезности начинается с того, что для каждого набора альтернатив записывается система уравнений (3)

$$\left\{ \begin{array}{l} P_1 = \int_{-\infty}^{\infty} f(\varepsilon_1) \prod_{l=2}^J F(\varepsilon_l + V_l - V_1) d\varepsilon_l \\ \dots\dots\dots \\ P_k = \int_{-\infty}^{\infty} f(\varepsilon_k) \prod_{l \neq k}^J F(\varepsilon_l + V_l - V_k) d\varepsilon_k \\ \dots\dots\dots \\ P_J = \int_{-\infty}^{\infty} f(\varepsilon_J) \prod_{l \neq k}^{J-1} F(\varepsilon_l + V_l - V_J) d\varepsilon_l \end{array} \right. , \quad (12)$$

которая имеет место и при моделировании дискретного выбора. Однако в случае, когда  $P_k = 1$  или  $P_l = 0$ , ее решение для последующей оценки коэффициентов модели не имеет смысла.

Для получения решения этой системы при  $0 < P_k < 1$  необходимо ввести новую переменную  $v_l$ , которая представляет собой относительную привлекательность  $l$ -й альтернативы относительно первого варианта:

$$v_l = V_l - V_1. \quad (13)$$

Система (12) преобразуется в систему (14) с количеством неизвестных параметров  $v_l$ , равным  $J - 1$ :

$$\begin{cases} P_1 = \int_{-\infty}^{\infty} f(\varepsilon_1) \prod_{l=2}^J F(\varepsilon_l + v_l) d\varepsilon_1 \\ \dots \\ P_k = \int_{-\infty}^{\infty} f(\varepsilon_k) \prod_{l \neq k}^J F(\varepsilon_l + v_l - v_k) d\varepsilon_k \\ \dots \\ P_J = \int_{-\infty}^{\infty} f(\varepsilon_J) \prod_{l \neq J}^{J-1} F(\varepsilon_l + v_l - v_J) d\varepsilon_J \end{cases}. \quad (14)$$

Можно показать, что система (14) имеет единственное решение относительно  $v_l$ . Это свойство было описано выше для двух альтернатив и обусловлено наличием очевидной связи между частотами  $\sum_{l=1}^J P_l = 1$ .

Такое решение весьма просто получить с помощью приближенных методов, для этого только необходимо знать параметры распределения случайной составляющей затрат на передвижение – времени ожидания пассажирами на остановке.

Для выбранного объекта исследования работа транспортных средств на маршрутах в большинстве случаев не регламентируется жестким расписанием движения. Она организуется с помощью их отправления из начальных остановочных пунктов маршрута при достижении некоторой степени заполнения салона пассажирами.

Такой способ организации движения транспортных средств сформирован благодаря тому, что основная масса городских маршрутов является коммерческой, а выручка от перевозки пассажиров поступает исключительно через водителей или кондукторов. Поэтому отправление по заполнению салона является результатом саморегулирования конкурентных взаимоотношений между водителями разных транспортных средств одного маршрута.

Для этого способа организации работы транспортных средств на маршруте возможно показать, что при простейшем пуассоновском потоке пассажиров с некоторым параметром  $\lambda$  интервал отправления транспортных средств имеет плотность гамма-распределения

$$f(\tau) = \frac{\lambda \cdot (\lambda \tau)^{q-1}}{(q-1)!} \cdot e^{-\lambda \tau}, \quad \tau > 0, \quad q \geq 1, \quad (15)$$

где  $\tau$  – промежуток времени между двумя последовательными отправлениями транспортных средств от начальной остановки маршрута, мин.;  $q$  – количество пассажиров в салоне, по достижению которого происходит отправление транспортного средства от начальной остановки маршрута.

Допуская стационарность потока моментов отправления транспортных средств из конечной остановки и зная среднее время между их отправлениями  $\bar{\tau} = \frac{q}{\lambda}$ , можно применить узловую теорему Смита [2] и найти плотность распределения времени ожидания  $t$  пассажиром очередного транспортного средства городского маршрута. Процесс восстановления в данном случае создают моменты отправления транспортных средств от начального пункта маршрута.

$$f(t) = \frac{1 - F(q)}{\bar{\tau}} = \frac{1}{q} \left( \lambda \cdot e^{-\lambda t} + \lambda^2 \cdot t \cdot e^{-\lambda t} + \dots + \frac{\lambda^q \cdot t^{q-1}}{(q-1)!} e^{-\lambda t} \right). \quad (16)$$

Эта плотность является средним арифметическим  $q$  гамма-плотностей, что с практической точки зрения может означать близость фактического времени ожидания пассажирами транспортных средств на маршруте к гамма-распределению. Эта гипотеза была проверена с помощью имитационного эксперимента, в результате которого оказалось, что время ожидания пассажирами транспортного средства одного маршрута хорошо описывается гамма-распределением с параметрами масштаба и формы, которые линейно зависят от среднего интервала движения на маршруте.

При подстановке обоих законов случайной составляющей полезности альтернатив в уравнение (3) аналитическое решение интеграла получено быть не может. Поэтому процедура поиска  $v_l$  была реализована в *Microsoft Excel* с помощью метода численного интегрирования Гаусса по семи точкам. После решения системы (14) неизвестной остается базовая полезность первой альтернативы. Как отмечалось выше, эта ситуация полностью соответствует относительному характеру полезности альтернатив для одного трудового

передвижения, но требует дополнительных усилий по ее определению. В данном случае она решалась в рамках регрессионного анализа путем включения свободных переменных в число независимых факторов, подробное описание приводится в [9].

В конечном итоге, после рассмотрения результатов анкетного обследования передвижений, с помощью метода наименьших квадратов была получена регрессионная модель для полезности трудового передвижения маршрутным транспортом:

$$v_i = 43,25 - 1,029 \cdot t_n - 9,227 \cdot \gamma_n - 10,443 \cdot \Pi_n - 11,430 \cdot \Pi, \quad (17)$$

где  $t_n$  – время передвижения маршрутным транспортом из дома на работу, мин.;  $\gamma_n$  – коэффициент заполнения салона транспортного средства в момент посадки;  $\Pi_n$  – денежная стоимость поездки из дома на работу, грн.;  $\Pi$  – количество пересадок на пути следования на работу, ед.

Модель имеет достаточно высокие статистические характеристики, в том числе множественный коэффициент корреляции, больший 0,9, и весьма значимые коэффициенты регрессии. Однако прогностические способности этой модели оказались неудовлетворительными. Значение критерия  $s_N^2 = 673,3$  для нее оказалось выше аналогичной оценки для равновероятной модели, в случае которой критерий  $s_N^2 = 579,6$ . Вероятность превышения этого значения случайной величиной, распределенной по  $\chi^2$ , в обоих случаях оказалась практически равной 0.

То есть, прямой расчет от частот выбора до регрессионной зависимости полезности от факторов дал приемлемые результаты, а обратный – от параметров пути до вероятности их выбора – неудовлетворительные. Это весьма символичный результат, так как модель (17) была получена после аналитического решения системы (12) для каждого набора альтернатив. Затем полученная аналитическим путем полезность альтернатив была сопоставлена с измеримыми параметрами пути с помощью метода наименьших квадратов, что дало неплохие результаты выравнивания. Очевидно, что при обратном расчете прогнозные значения относительной полезности пути в общем случае не должны совпадать с полезностями, определенными через решение системы (12). Однако полученные расхождения приводят к очень серьезным отклонениям в расчетной вероятности, то есть модель вероятности оказалась очень неустойчивой по отношению к значениям полезности. Аналогичное свойство отмечалось и для моделей дискретного выбора, но там причиной слишком высокой чувствительности модели является экспоненциальный вид функции полезности. В данном случае причина иная, так как никаких ограничений на вид регрессионной зависимости не накладывается, а хорошие статистические характеристики были достигнуты уже для линейной модели.

Для данной методики единственным объяснением плохих прогностических способностей модели является слабая зависимость вероятности выбора пути от случайной составляющей полезности альтернатив. Основная причина этого состоит в том, что выделенная из полезности случайная величина времени ожидания является не единственным случайным фактором выбора альтернативы.

Принятие решения о рациональном варианте текущей поездки, несомненно, зависит не только от интервала движения на маршрутах, но и от целого ряда других факторов, значения которых недоступны транспортным инженерам при проведении расчетов. Это могут быть запас времени на поездку, наличие попутчиков, погодные условия, настроение пассажира и т.д. Считать, что суммарное воздействие этих факторов можно описать одной случайной величиной, измеряемой в каких-то единицах (например, времени), по крайней мере, очень смело. Ведь даже характер взаимодействия этих факторов вряд ли может быть детально описан, а тем более спрогнозирован. В то же время полученные в результате обследования частоты не освобождены от влияния большого количества нетранспортных факторов, а рассчитанные на основе времени ожидания относительные полезности альтернатив никак этого не учитывают. Поэтому даже небольшие отклонения расчетной полезности альтернатив, выраженной в единицах времени ожидания, при подстановке в (14) приводят к существенным изменениям в расчетной вероятности выбора пути.

Вторая причина заключается в неполной адекватности описания времени ожидания гамма-распределением. С практической точки зрения результаты экспериментального исследования времени ожидания и возможность их описания гамма-законом сомнений не вызывают. Сам эксперимент состоял из 48 серий по 50-300 опытов в каждой, общее количество рассчитанных значений времени ожидания составило 8452 значения. Гипотеза о соответствии экспериментальных данных гамма-распределению во всех 48 случаях не была отвергнута ни по критерию Пирсона, ни по критерию Колмогорова-Смирнова при 5%-ном уровне значимости.

Еще одна возможная причина неполной адекватности используемого закона кроется в способе оценки степени соответствия между эмпирическим и теоретическим распределениями, которые допускают определенные расхождения между ними. Это приводит, например, к возможности использования различных законов для одного и того же набора фактических данных. В случае со временем ожидания полным аналогом гамма-закону во всех 48 случаях стало распределение Вейбулла.

Однако с теоретической точки зрения оба этих закона не полностью соответствуют процессу ожидания транспорта на остановке, когда время его прибытия не известно пассажиру. С одной стороны, вероятность нулевого времени ожидания не должна равняться нулю:  $F_i(0) < 0$ , с другой стороны, время ожидания не может быть бесконечным:  $F_i(t_{\max}) = 1$ ,  $t_{\max} = \text{const}$ . Оба из использованных теоретических распределений не соответствуют этим условиям.

Третьей причиной неудовлетворительного уровня прогнозирования вероятностей может считаться не слишком обширный набор данных. Результаты обследования включают в себя 307 участников, каждый из которых по 5 раз выбирал вариант пути, а общее количество реализованных альтернатив составило 681 ед. Способ проведения обследования обеспечивает получение фактических частот использования альтернатив, которые каким-то образом отображают искомые вероятности выбора пути следования. Точность отображения вероятностей в значительной степени зависит как от продолжительности наблюдения за процессом выбора альтернативы каждым участником, так и от количества участников обследования. Поэтому остается открытым вопрос о достаточности такого набора альтернатив для построения адекватной модели.

Таким образом, разработанный в рамках моделей дискретного выбора способ не позволил получить адекватную модель для прогнозирования вероятностей выбора пассажирами пути следования.

К сожалению, сравнить полученные результаты со стандартной моделью дискретного выбора на использованном наборе данных невозможно, так как каждый из респондентов имел не упорядоченный, а свободный набор альтернатив. В этом наборе порядковый номер альтернативы определяется не принадлежностью к способу обслуживания, а его позицией в анкете фиксации фактического выбора пути следования, то есть является случайной величиной. В стандартном подходе такая ситуация не может быть обработанной, хотя именно она необходима для моделирования транспортного процесса.

Однако и для данного набора может быть построена более адекватная модель, если отказаться от попыток аналитического описания причин выбора альтернативы с помощью (1)-(3). Для этого нужно каким-то образом преобразовать исходные частоты выбора в привлекательность альтернативы и установить регрессионную связь между полученным вектором полезности и параметрами пути следования.

Самым подходящим на наш взгляд средством преобразования является простая нормировочная модель:

$$P_k = \frac{V_k(\bar{X})}{\sum_{l=1}^J V_l(\bar{X})}. \quad (18)$$

Тогда, при выборе в качестве базы сравнения 1-й альтернативы

$$V_l = V_1 \frac{P_k}{P_1}. \quad (19)$$

Значения  $P_k$  и  $P_1$  известны по результатам обследования, и встает вопрос об определении значения привлекательности первой альтернативы  $V_1$ . Этот вопрос является естественным вследствие относительного характера альтернатив и необходимости объединения различных корреспонденций в единый массив данных. Именно привлекательность первой альтернативы дает возможность привести затраты на передвижение к единому масштабу. Она может быть определена путем введения ее в массив данных для регрессионного анализа или каким-то другим способом.

Самым простым вариантом  $V_l$  является ее линейная зависимость от параметров пути. Регрессионный анализ результатов обследования для этого варианта функции привлекательности дал следующую модель:

$$V_l = 2,39 - 0,02 \cdot t_n + 0,48 \cdot L_n - 1,46 \cdot \gamma_n - 4,96 \cdot \Pi_n + 1,54 \cdot \Pi, \quad (20)$$

где  $L_n$  – дальность транспортного передвижения, км.

Прогнозная вероятность выбора альтернативы определяется путем подстановки (20) в (18). Разность наблюдаемых частот и расчетных вероятностей составила для этой модели  $s_N^2 = 330,67$  при  $n_\chi = 681 - 307 - 6 = 368$  степенях свободы. Вероятность превышения этого значения случайной величиной, распределенной по  $\chi^2$ , равна 0,919, что свидетельствует об очень высоких прогностических свойствах модели. Однако ряд факторов в ней, таких как время передвижения маршрутным транспортом из дома на работу, коэффициент заполнения салона транспортного средства в момент посадки и количество пересадок на пути следования на работу, оказались по критерию Стьюдента незначимыми. Поэтому модель (20) была упрощена и рассчитаны новые коэффициенты регрессии:

$$V_l = 3,46 + 0,48 \cdot L_n - 4,27 \cdot \Pi_n. \quad (21)$$

Значение критерия для этой модели несколько ухудшилось:  $s_N^2 = 336,82$ , также снизилась и его вероятность 0,898 при  $n_\chi = 681 - 307 - 3 = 371$  степени свободы. Однако считать такие ухудшения значимыми не приходится, поэтому для прогнозирования вероятности выбора пассажиром варианта пути следования можно использовать обе модели – (20) и (21).

Остается открытым вопрос объяснения физического смысла коэффициентов полученных моделей, например положительного коэффициента дальности поездки. Получается, что чем длиннее альтернатива, тем чаще она выбирается. При этом объяснить такой эффект тесной корреляцией между дальностью и стоимостью передвижения не представляется возможным, так как коэффициент парной корреляции между ними равен -0,11.

Но отрицать высокие прогностические способности такого подхода на основе имеющихся данных нельзя. Более того, наверняка существует возможность сократить расхождения между зафиксированными частотами и расчетными вероятностями путем использования нелинейной модели  $V_1$ .

С учетом широких возможностей и простоты подхода построение регрессионных моделей для привлекательности пути следования, которая определяется на основе частоты использования альтернативы, является очень перспективным направлением моделирования поведения пассажиров в транспортной системе. Реализация представленного подхода не требует принятия дополнительных допущений и предположений, позволяет избежать проблемы корреляции альтернатив, что обеспечивает ему значительные преимущества перед моделями дискретного выбора.

### Выводы

Для оценки качества создаваемых моделей выбора альтернатив необходимо использовать характеристику отличия теоретических вероятностей от исходных частот использования альтернатив. Лучше всего для этого подходит аналог критерия Пирсона (7), модифицированный для данного случая. Он обеспечивает получение характеристики отклонений, распределенной по  $\chi^2$ , что позволяет проводить вероятностную оценку расхождений.

Стандартный на сегодня подход к построению моделей выбора альтернатив при реализации потребности в передвижении – моделирование дискретного выбора – не может считаться законченным по целому ряду причин. Он основывается на гипотетических предположениях о характере распределения случайной составляющей полезности передвижения, не может быть использован в ситуациях с разными наборами альтернатив и используется для оценки коэффициентов модели метод максимального правдоподобия без достаточного обоснования. Список проблем этого подхода можно продолжить, но самая главная проблема заключается в том, что он основывается на прогнозе детерминированной части привлекательности с помощью случайной составляющей. Здесь требуются очень точные знания об этой случайной величине, для которой не известен даже полный перечень ее составляющих, что делает задачу получения таких данных практически нерешаемой.

Детализация такого подхода для маршрутного транспорта с использованием нового метода также не привела к получению адекватной модели вероятности выбора пути. При этом разработанный метод получения коэффициентов при факторах в функции привлекательности полностью соответствует относительному характеру альтернатив.

Наиболее эффективным подходом к моделированию выбора пассажирами вариантов пути следования в маршрутной системе города оказался регрессионный анализ привлекательности пути. Последняя может быть получена из зафиксированных частот использования альтернатив на основе принципа «чем чаще используется альтернатива, тем выше привлекательность пути». В этом случае может быть получена модель «черного ящика», имеющая достаточно высокие прогностические способности благодаря гибкости инструмента регрессионного анализа.

### Список литературы

1. Горбачов П. Ф. Концепція формування систем маршрутного пасажирського транспорту в містах: дисс. ... д.т.н.: 05.22.01. Харьков, 2009. 370 с.
2. Bahandur R. On the Asymptotic Efficiency of Tests and Estimates. Sankhya, 1960.
3. Ben-Akiva M. Discrete Choice Analysis: Theory and Application to Travel Demand (Transportation Studies). Massachusetts, 1985. 380 p.
4. Borndörfer R. Models for Fare Planning in Public Transport. Berlin, 2012.
5. Cochran W. G. The  $\chi^2$  Test of Goodness-of-Fit // Annals of Mathematical Statistics. 1952. Vol. 23.
6. Greene W. Discrete Choice Modeling // The Handbook of Econometric. 2008. Vol. 2. № 4.2.
7. Kjær T. A Review of the Discrete Choice Experiment. Southern Denmark, 2005.
8. Ortuzar J. D. Modelling Transport. L., 2006.
9. Sándor Z. Multinomial Discrete Choice Models // Quantile. 2009. Vol. 7.
10. Smith W. L. Asymptotic Renewal Theorems // Proceedings of the Royal Society of Edinburgh. Section A. 1954. Vol. 64.

### METHODS TO FORM MODEL OF CHOOSING VARIANT OF WORKING TRAVEL BY FIXED ROUTE TRANSPORT

Gorbachev Petr Fedorovich, Doctor in Technical Sciences, Professor  
 Makarichev Aleksandr Vladimirovich, Ph. D. in Physical-Mathematical Sciences, Associate Professor  
 Svichinskaja Ol'ga Vladimirovna  
 Kharkiv National Automobile and Highway University, Ukraine  
 gorbachev\_pf@mail.ru; amakarichev@mail.ru; svichinskaya@ukr.net

The questions of using the models of discrete choice for getting the function of travel line attractiveness in urban passenger transport route system when implementing the interval method of passengers' service are considered. Two variants of the model for processing the results of studying passengers' actual travel line choice are suggested. The availability of using regressive analysis for evaluating the coefficients of the route attractiveness model, which was created after the transformation of alternatives use frequency, is shown.

*Key words and phrases:* passenger; models of travel line discrete choice; alternative variant of route; travel attractiveness; parameters of travel line; method of maximum likelihood; regressive analysis.