

ОЦЕНКА ПРИТЯЖЕНИЯ ГРУПП ПО МЕЖГРУППОВЫМ ПОТОКАМ*

© 1995 г. Староверов О.В.

(Москва)

Исследуется движение населения между группами, которые выбирают-ся по воле и желанию людей. Предложен метод объединения макро- и микроподходов для получения индивидуальных и групповых коэффициентов предпочтения. Приведен статистический метод оценки параметров подвижности населения.

ВВЕДЕНИЕ

Наука о народонаселении выделяет обычно три вида движения: 1) естественное, когда группы населения образованы по полу и возрасту; 2) миграцию внутри страны; 3) подвижность между социальными группами.

Движение первого вида здесь рассматриваться не будет, так как переходы из одной возрастной группы в другую происходят предопределенно биологически.

Второй вид отражает пожелания людей и их возможности, поэтому изучение миграции уже требует учета конкретных условий. Очевидно, что лучшие условия определяют перемещение людей, хотя сами возможности лимитированы стоимостью переезда и рядом других обстоятельств, например, ограниченностью подходящего жилья, несмотря на то, что миграция представляется более "человеческим" по сравнению с естественным движением. Именно эти осложнения не дают полностью "очеловечиться" миграции.

Третий вид нуждается в уточнении. Это может быть движение между разными социальными группами (управленцами, инженерами, квалифицированными рабочими и т.д.) или из одних секторов экономики в другие, между группами людей разных профессий. Причем движение между ними возможно и без ограничений (как, например, уже упоминалось – жилищными условиями), т.е. оно совсем "человеческое", происходящее лишь по желанию. Что же касается групп руководителей и профессионалов очень высокого ранга, то движение между ними может и вовсе не иметь ограничений.

Для количественного изучения выбранных видов движения необходимы модели, из которых наиболее известными являются марковские. К ним относятся многостатусные [1], которые описывают переходы не только между регионами, но и между группами. Потоки людей между регионами отражают гравитационные модели – разновидность марковских. Но для описания миграции удобнее всего факторные модели, потоки в которых управляются только условиями или свойствами групп выбытия и прибытия, т.е. они потенциально содержат функции предпочтения. Кроме этого при анализе миграции всегда учитывается еще расстояние.

В [2] есть такая функция, связывающая интенсивность перехода отдельного человека (показатель миграции) из группы i в группу j со свойствами c_i и c_j

$$r_{ij} = F(c_i, c_j, \gamma_{ij}), \quad (1)$$

где γ_{ij} – стоимость перехода (переезда) из i в j . Естественно, что $(\partial F / \partial c_i) < 0$, $(\partial F / \partial c_j) >$

* Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект 93-06-10072).

> 0 и $(\partial F / \partial \gamma_{ij}) < 0$, если улучшение свойств группы ведет к большей привлекательности. Поэтому негативные факторы должны быть взяты со знаком "минус".

Однако попытки использовать макроподход для описания микропроцессов (1) не дали существенных результатов. Это объясняется тем, что формальная замена факторов в (1) на средние свойства группы (региона) не отражает мнения отдельных людей, а поэтому и их группового поведения. Удачнее получилось, когда вместо свойств групп начали применять функции полезности (ФП) этих свойств [3], но подход в [3] был принципиально иным по следующим причинам.

Во-первых, в [3] используется макроподход, а во-вторых, авторы этой работы не выбирали факторы мобильности и не собирали соответствующую статистику. В [3] для некоторых стран на основе потоков v_{ij} людей за единицу времени между регионами i и j рассмотрены оценки средних полезностей u_i и u_j этих регионов и расстояния между ними d_{ij} . Такой подход основан на макростатистике, поэтому использовались средние полезности u_i и u_j групп i и j . Модель выглядела так

$$r_{ij} = A e^{u_j - u_i} / d_{ij} \quad (2)$$

и $Mv_{ij} = n_i r_{ij} h$, где h – единица времени (обычно год); n_i – объем группы в начале периода.

Соотношение макромоделей (2) является частным случаем микромоделей (1), потому что в (2) выбран конкретный вид функции F из (1). Действительно, в (2) полезности

$u_i = \int_{-\infty}^{\infty} u(c) f_i(c) dc$, где $f_i(c)$ – плотность распределения свойств c среди членов группы i .

Поскольку $f_i(z)$ и $u(z)$ – произвольны, то это уравнение имеет достаточно общий вид относительно свойств c .

Остаются открытыми вопросы об обратной пропорциональности зависимости правой части (2) от расстояния; независимости постоянной величины A от исходной группы; зависимости интенсивностей переходов от разности полезностей. Последнее особенно важно пояснить, так как возможность вычитания из u_j полезности u_i позволяет их сравнивать (вычитать), т.е. в (2) предполагается, что функции $u_i(c)$ одинаковы для всех групп i .

Обычно ФП задаются с точностью до монотонного преобразования, поэтому полезности разных индивидуумов или групп могут быть приведены к одной функции, возможно, разными монотонными преобразованиями. Но это, во-первых, может вызвать неодинаковую реакцию людей на одни и те же условия в различных группах, во-вторых, должно по-прежнему отражать реакцию каждого индивидуума определенной группы, т.е. всей группы, на уже приведенную к одинаковым значениям полезность. Поэтому коэффициент A в (2) может зависеть от группы, что и было отмечено.

Соотношения (1) и (2) отражают разные типы как микро- и макроподходов к описанию переходов, так и виды параметризации их интенсивностей r_{ij} в марковских моделях. Поэтому возникают проблемы: 1) параметризации моделей и 2) объединения микро- и макроподходов, особенно актуальные при марковском моделировании, потому что именно эти модели достаточно хорошо описывают реальные процессы движения населения [4].

Цель данной статьи – предложить новую параметризацию вероятностей переходов. На основе этой параметризации можно получить алгоритм оценки параметров без знания статистики о факторах движения и проанализировать адекватность отражения ими реального движения. Подход к такой параметризации был известен раньше в совершенно другой области.

Теория поиска работы (см., например [5]) исследует движение отдельного индивидуума и базируется на изменении заработков. В ней считается, что переход может быть обусловлен предложением большей заработной платы, которая характеризуется

функцией распределения заработков в разных фирмах, отраслях экономики, регионах и т.п.

Однако межрегиональное движение реагирует не только на оплату труда, но и на жилищные условия. Поэтому при исследовании движения людей необходимо рассматривать функции распределения и заработков, и жилья. В общем виде можно изучать функции распределения многих свойств разных групп. При микроподходе обычно множественность факторов приводит к ФП, которая позволяет сравнивать векторы.

Но тут возникает другая задача. Для параметризации интенсивностей (вероятностей в единицу времени) перехода необходимо так соединить микро- и макроподходы, чтобы можно было проводить расчет, используя лишь статистику о потоках переходов без выбора факторов движения и знания их значений. Оценивать предлагаемые параметры на основе потоков переходов нужно так же, как параметры в (2) с помощью хорошо подобранного метода.

В [3] параметры оцениваются методом наименьших квадратов. Этот наиболее распространенный метод следует использовать тогда, когда дисперсии всех потоков одинаковы. В противном случае целесообразнее применять метод наименьших квадратов с весами, но лучше – метод максимального правдоподобия (см. [6]).

1. ПАРАМЕТРЫ МОДЕЛИ МОБИЛЬНОСТИ

Остановимся на содержательном представлении параметров r_{ij} модели движения населения.

Известно, что люди меняют свое место жительства, когда находят лучшие условия жизни и работы. Свойств места (или статуса человека), которые определяют эти условия, может быть несколько. Как уже отмечалось, межрайонная миграция обусловлена не только заработной платой занятых на разных территориях, но и жилищными условиями. Однако межпрофессиональное движение может происходить без смены жилищных условий, но с учетом других факторов, например перспектив роста.

Пусть $m > 1$ – число условий (свойств), x_l – уровень свойства l . Тогда сравнивать места можно с помощью ФП $u(x_1, \dots, x_m)$. Если все население разбито на k групп и имеется распределение $F_i(x_1, \dots, x_m)$ уровней x_l по индивидуумам группы i (например, района, отрасли, предприятия), то существует распределение полезности u уровней $x_i = (x_{i1}, \dots, x_{im})$ группы $i = 1, \dots, k$

$$G_i(y) = P_i(u(x_i) < y) = \int_{u(x_i) < y} \dots \int dF_i(x_i).$$

Человек из группы i с уровнями x_i или ФП u_i будет стремиться попасть в группу j с ФП не меньшей, чем u_j , т.е. область притяжения группы j состоит из всех таких x_j , что $u_i \geq u(x_j) = u_j$. Таким образом, тяга человека группы i , имеющего условия x_i (т.е. ФП u_i), в группу j тем больше, чем больше

$$q_j(y_i) = \int_{y_i}^{\infty} dG_j(z).$$

Поскольку наблюдаются не индивидуальные переходы с каждого уровня u_i , а переходы всех людей, то поток из i в j будет возрастать с ростом

$$q_{ij} = \int_0^{\infty} q_j(y) dG_i(y), \quad i, j = 1, \dots, k.$$

Свойство 1. $q_{ij} \geq 0$ и $q_{ij} + q_{ji} = 1$.

Свойство 2. Величины q_{ij} не зависят от монотонно возрастающих преобразований ни самих $x_1, \dots, x_m$, ни ФП и.

Наблюдаемые потоки v_{ij} между k -группами i и j дают матрицу $N = \|v_{ij}\|$.

Гипотеза. При каждом i параметры потоков v_{ij} для любого $j \neq i$ пропорциональны

q_{ij} .

Если эта гипотеза выполнена, то q_{ij} обусловят внедиагональные элементы матрицы $N = \|v_{ij}\|$. Однако при этом к параметрам q_{ij} добавляются коэффициенты пропорциональности $c_i, i = 1, \dots, k$.

Теперь потоки из i в $j, i \neq j$, будут определяться k параметрами c_i и $(k^2 - k)/2$ -параметрами q_{ij} (число последних не равно $k^2 - k$ из-за справедливости свойства 1). Часто c_i называют склонностью к переходам у людей группы i и она может быть для разных групп различной.

Из свойства 1 следует, что $q_{ii} = 0,5$ для любого i , но вследствие традиций, языка, знания условий, наличия навыков и т.д., внутригрупповые потоки могут не быть одинаковыми не только из-за различий в n_i , поэтому в параметр $n_i r_{ii}$ потока v_{ii} добавим пока чисто формально еще одно слагаемое $-b_i$ и получим $r_{ii} = b_i + c_i q_{ii} = b_i + c_i/2$. Таким образом, для r_{ij} выполняется соотношение

$$r_{ij} = b_i \delta_{ij} + c_i q_{ij}, \quad i, j = 1, \dots, k, \quad (3)$$

а все потоки v_{ij} имеют среднее значение, равное $n_i(b_i \delta_{ij} + c_i q_{ij})$; здесь и далее δ_{ij} — символ Кронекера.

Замечание. Статистика очень часто объединяет людей, которые в обследуемый период времени никуда не переходили, с теми, кто переходил внутри группы. Такая потеря информации заставляет рассматривать лишь межгрупповые потоки. Кроме того, иногда бывают не известны начальные численности n_i групп, от которых сильно зависят все оценки параметров.

Из (3) следует, что при любых c_i и q_{ij} величина b_i определяется однозначно, если задан r_{ij} , поэтому важно найти хорошие оценки c_i и q_{ij} , когда известны потоки v_{ij} . Таким образом, на первых порах можно исключить диагональные элементы матрицы N и рассчитывать оценки c_i и q_{ij} по внедиагональным элементам $v_{ij}, i \neq j$. Для исключения n_i независимо от того, известна она или нет, переобозначим параметры $v_{ij}, i \neq j$, следующим образом: $a_i = n_i c_i$, т.е. $n_i r_{ij} = a_i q_{ij} \forall i \neq j$ при

$$q_{ij} + q_{ji} = 1, \quad q_{ij} \geq 0. \quad (4)$$

После описанной перепараметризации a_i представляет собой коэффициент пропорциональности r_{ij} потока из i в другие группы, имеющие предпочтения q_{ij} : $n_i r_{ij} = a_i q_{ij}$. При известных n_i каждой группы коэффициенты $c_i = a_i/n_i$ связывают q_{ij} при $i \neq j$ с интенсивностями перемещения населения $r_{ij} = c_i q_{ij}$.

2. ОЦЕНКА ПАРАМЕТРОВ ДВИЖЕНИЯ НАСЕЛЕНИЯ

Рассмотрим $k^2 - k$ межгрупповых потоков v_{ij} .

Известно несколько методов оценки параметров, среди них метод максимального правдоподобия позволяет получать оценки, обладающие свойством несмещенности и минимальности дисперсии при большом числе наблюдений.

На основе следующих гипотез ([7]): 1) независимость движения каждого человека между группами от движения других людей; 2) принятие решения о моменте перехода каждым индивидуумом самостоятельно (ординарность); 3) зависимость направления перехода лишь от группы, где человек находится, а не от того, как он в эту группу попал (отсутствие последействия) и 4) вероятность перехода каждого человека в малом интервале времени пропорциональна этому интервалу — можно утверждать, что потоки между группами i и j за фиксированный промежуток времени независимы между собой и распределены по закону Пуассона.

Известно, что параметром закона Пуассона будет среднее значение v_{ij} — численности перешедших (за время наблюдения) из группы i в j (всего групп k). Далее нигде не исключаются переходы внутри группы, которые тоже входят в движение населения, так как не нужно исключать, например, мигрантов внутри отдельного региона.

В таких обозначениях $v_{ij} \sim Pu(n_i r_{ij})$ и независимы для любых i, j , где $Pu(n_i r_{ij}) = P(v_{ij} = m_{ij}) = \frac{(n_i r_{ij})^{m_{ij}}}{m_{ij}!} e^{-n_i r_{ij}}$, $m_{ij} = 0, 1, 2, \dots$. Значения r_{ij} и n_i — прежние.

Методом максимального правдоподобия находятся неизвестные параметры распределения, которые максимизируют вероятность наблюдаемых случайных величин — функцию правдоподобия. В силу независимости всех v_{ij} она равна

$$\prod_{i=1}^k \prod_{j \neq i}^k \frac{a_i q_{ij}^{v_{ij}}}{v_{ij}!} e^{-a_i q_{ij}}. \quad (5)$$

В дальнейшем будет рассмотрена логарифмическая функция правдоподобия

$$\Phi(a, q) = \sum_{i=1}^k \sum_{j \neq i}^k (v_{ij} \ln a_i + v_{ij} \ln q_{ij} - a_i q_{ij} + \ln v_{ij}!). \quad (6)$$

Необходимые условия максимума (6) дают уравнение правдоподобия. Первые из них

$$\frac{\partial \Phi(a, q)}{\partial a_i} = \sum_{j \neq i}^k \left(\frac{v_{ij}}{a_i} - q_{ij} \right) = 0 \quad (7)$$

для произвольных по q_{ij} . Заметим, что с учетом (4) от q_{ij} , $i < j$, зависят лишь два слагаемых (6), в которые входят только следующие величины, зависящие непосредственно от q_{ij} : $(v_{ij} \ln q_{ij} - a_i q_{ij}) + (v_{ji} \ln(1 - q_{ij}) - \hat{a}_i(1 - q_{ij}))$; поэтому вторые уравнения

$$\frac{\partial \Phi(a, q)}{\partial q_{ij}} = \left(\frac{v_{ij}}{q_{ij}} - a_i \right) - \left(\frac{v_{ji}}{1 - q_{ij} - a_j} \right) = 0. \quad (8)$$

Таким образом, необходимо найти решение систем (7) и (8) относительно $k(k-1)/2 + k$ неизвестных q_{ij} , $i < j$ и a_i .

Для оценки параметров b_i , отражающих влияние традиций, языка и т.д. во внутригрупповых переходах в (5), нужно добавить множители для $i = j$

$$\prod_{i=1}^k \frac{[n_i(b_i + c_i q_{ii})]^{v_{ii}}}{v_{ii}!} e^{-n_i(b_i + c_i q_{ii})}.$$

Из (4) следует, что $q_{ii} = 1/2$, поэтому сразу же получаем наилучшие оценки $\hat{b}_i = v_{ii} / n_i - \hat{c}_i / 2$ для $i = 1, \dots, k$.

Если n_i не задано, то легко найти оценку $n_i b_i$ — численности внутригрупповых переходов, определяемых лишь присущим каждой группе бинарным фактором (язык, навыки и т.д.). Для этого достаточно рассмотреть вместо b_i параметр $g_i = n_i b_i$, оценка которого — $\hat{g}_i = v_{ii} - \hat{a}_i / 2$, где $\hat{a}_i = \hat{c}_i n_i$.

3. ВЫБОР КРИТЕРИЯ СРАВНЕНИЯ

Начальные численности перемещающегося населения обычно не приводятся в тех случаях, когда публикуют доли перешедших среди всех выбывших из одних групп во все остальные. Внутригрупповые переходы часто вообще опускаются, так как для подсчета численностей групп в последующий период времени они не нужны. Но они дают очень полезную информацию о переходах в среде, где распределение благ между всеми членами одно и то же.

Часто потоки между группами округляются, т.е. эти показатели приводятся с различным числом значащих цифр. Следовательно, для различных систем групп (стран, моментов обследования и т.д.) данные неодинаковы, т.е. они бывают приведены с разными масштабными множителями. Отсюда – *первое требование* к критерию сравнения качества полученных оценок параметров модели движения населения – он должен быть инвариантным к масштабу статистических данных. Выполнение этого свойства позволит сравнивать движение людей в различных системах групп (регионах стран, межотраслевое, профессиональное, социальное движение и т.д.) как между собой, так и в разные периоды времени.

Обычно исследователи стремятся свести число неизвестных параметров к минимуму. Причина очень проста – большое число параметров дает возможность сколь угодно хорошо описать любые статистические данные. Поэтому критерий не должен реагировать на изменение количества оцениваемых параметров, с помощью которых осуществляется приближение наблюдений. Так получено *второе требование* – инвариантность к числу параметров, которое в основном зависит от количества групп. Это требование необходимо особенно для исследования движения населения, попытки объяснить которое делались с использованием разного набора параметров, среди которых были коэффициенты притяжения и отталкивания групп [8].

В статистических исследованиях встречается критерий, удовлетворяющий первому сформулированному требованию, – коэффициент множественной корреляции. Оценка его такова

$$\hat{R}^2 = 1 - s_1^2 / s_0^2,$$

где $s_0^2 = \frac{1}{t-1} \sum_{i=1}^k \sum_{j \neq i}^k (v_{ij} - \bar{v})^2$; $s_1^2 = \frac{1}{t-l} \sum_{i=1}^k \sum_{j \neq i}^k (v_{ij} - n_i \hat{c}_i \hat{q}_{ij})^2$, $\bar{v} = \frac{1}{t} \sum_{i=1}^k \sum_{j \neq i}^k v_{ij}$; $t = k(k-1)$ – общее количество наблюдений в случае, когда внутригрупповых переходов нет; $l = k(k+1)/2$ – количество неизвестных параметров; $\hat{R}^2$ удовлетворяет первому требованию, а статистика

$$z = \frac{\hat{R}^2}{1 - \hat{R}^2} \frac{t-l-1}{l} \quad (9)$$

– обоем. Здесь z при больших n_i будет иметь распределение, близкое к F -распределению (Фишера) с l - и $t-l-1$ -степенями свободы, так как все v_{ij} близки к нормальным независимым случайным величинам.

Легко видеть, что в критерии (9) при оценках $\hat{g}_i = v_{ii} - \hat{a}_i / 2$ все слагаемые $(v_{ii} - \hat{g}_i - \hat{a}_i / 2)^2$ равны нулю. Поэтому при добавлении еще k параметров g_i величина z не изменится, но не останется постоянным число степеней свободы числителя l , которое теперь равно $k(k+3)/2$. Параметр g_i можно оценивать, если и количество наблюдений станет на k больше – добавятся v_{ij} . Тогда на k увеличится число t и критерий (9) будет менее жестким, чем критерий без добавочных параметров g_i .

Таблица 1

Оценки коэффициентов предпочтения q (на диагонали b), %
(межреспубликанское движение в СССР, 1967 г.)

Из группы	В группу														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-4,6	1,4	0,3	0,0	0,7	0,1	0,0	0,1	0,1	0,1	0,1	0,1	0,0	0,0	0,1
2	98,6	+1,7	7,9	4,2	15,4	2,0	1,7	0,9	4,7	1,4	0,7	0,8	0,7	0,8	0,9
3	99,7	92,1	+1,7	90,5	91,5	3,0	6,6	10,2	3,3	12,8	2,5	2,3	1,6	2,1	3,3
4	100,0	95,8	9,5	+5,1	78,2	1,5	3,8	0,6	2,1	1,6	20,3	16,4	1,1	9,8	0,6
5	99,3	84,6	8,5	21,8	+1,6	1,0	2,1	1,6	3,2	1,2	11,7	2,7	0,4	1,6	0,4
6	99,9	98,0	97,0	98,5	99,0	+0,5	54,2	4,9	80,0	10,1	98,0	97,4	59,0	91,4	8,0
7	99,9	98,3	93,4	96,2	97,9	45,8	+0,6	2,7	12,1	6,3	94,7	91,6	66,1	38,5	2,1
8	100,0	99,1	89,8	99,4	98,4	95,1	97,3	+1,4	95,6	52,0	99,5	98,5	2,4	97,7	56,2
9	99,9	95,3	96,7	97,9	96,8	20,0	87,9	4,4	+0,7	5,9	96,3	92,1	2,3	86,1	3,3
10	99,9	98,6	87,2	98,4	98,8	89,9	93,7	48,0	94,1	+1,7	97,2	95,4	7,9	96,4	50,6
11	99,9	99,3	97,5	79,7	83,3	2,0	5,3	0,5	3,7	2,8	+1,2	26,9	0,3	8,3	0,7
12	99,9	99,2	97,7	83,6	97,3	2,6	8,4	1,5	7,9	4,6	73,1	+0,7	1,8	18,3	5,7
13	100,0	99,3	98,4	98,9	99,6	41,0	33,9	97,6	97,7	92,1	99,7	98,2	+0,4	91,5	95,2
14	100,0	99,2	97,9	90,2	98,4	8,6	61,5	2,3	13,9	3,6	91,7	81,7	8,5	+1,1	1,2
15	99,9	99,1	96,7	99,4	99,6	92,0	97,9	43,8	96,7	49,4	99,3	94,3	4,8	98,8	+2,0

Примечание. Здесь и в табл. 2, 3: 1 – Россия, 2 – Украина, 3 – Белоруссия, 4 – Узбекистан, 5 – Казахстан, 6 – Грузия, 7 – Азербайджан, 8 – Литва, 9 – Молдавия, 10 – Латвия, 11 – Киргизия, 12 – Таджикистан, 13 – Армения, 14 – Туркмени, 15 – Эстония.

Оценки коэффициентов предпочтения q (на диагонали b), %
(межреспубликанское движение в СССР, 1989 г.)

Из группы	В группу														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-4,7	1,7	0,3	0,2	0,6	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,0	0,1	0,0
2	98,3	+1,5	5,8	2,7	5,1	1,2	1,9	0,7	5,0	0,9	0,6	0,7	0,6	1,0	0,6
3	99,7	94,2	+1,9	85,8	94,6	5,2	91,9	14,5	8,4	11,0	3,5	3,1	93,1	5,9	4,4
4	99,8	97,3	14,2	+0,7	81,4	1,0	79,2	1,0	1,6	1,3	19,4	19,2	13,1	14,5	0,5
5	99,4	94,8	5,4	18,6	+1,7	0,6	5,9	0,7	1,5	0,7	10,3	2,0	0,7	1,6	0,5
6	99,9	98,8	94,8	99,0	99,4	+0,5	96,1	2,0	92,8	2,9	3,8	2,2	93,6	3,8	1,9
7	99,9	98,1	8,1	20,8	94,1	3,9	+0,7	0,7	1,1	0,7	3,5	1,0	25,8	6,9	0,5
8	99,9	99,3	85,5	99,9	99,3	98,0	99,3	+0,2	98,8	48,5	97,7	96,7	99,6	96,5	29,2
9	99,9	95,0	91,6	98,4	98,5	7,2	98,9	1,2	+1,3	3,1	2,5	1,7	98,1	3,6	1,5
10	99,9	99,1	89,0	98,7	99,3	97,1	99,3	51,5	96,9	+1,2	95,6	94,4	99,4	89,9	36,3
11	99,9	99,4	96,5	80,6	89,7	96,2	99,5	2,3	97,5	4,4	+1,2	35,3	99,5	11,2	1,2
12	99,9	99,3	96,9	80,8	98,0	97,8	99,0	3,3	98,3	5,6	64,7	+0,7	99,8	41,6	1,3
13	100,0	99,4	6,9	86,9	99,3	6,4	74,2	0,4	1,9	0,6	0,5	0,2	+0,2	0,9	0,3
14	99,9	99,0	94,1	85,5	98,4	96,2	93,1	3,5	96,4	10,1	88,8	58,4	99,1	+1,0	2,0
15	100,0	99,4	95,6	99,5	99,5	98,1	99,5	70,8	98,4	63,7	98,8	98,7	99,7	98,0	+1,2

Склонности к миграции с в СССР по годам, %

Год	Из группы						
	1	2	3	4	5	6	7
1967	14,41	0,41	0,14	2,13	0,68	0,04	0,07
1973	14,06	0,55	0,15	0,16	0,63	0,03	0,06
1989	9,74	0,40	0,09	0,10	0,34	0,08	0,36

Год	Из группы							
	8	9	10	11	12	13	14	15
1967	0,04	0,07	0,06	0,25	0,19	0,03	0,19	0,10
1973	0,03	0,12	0,07	0,25	0,11	0,03	0,12	0,09
1989	0,00	0,12	0,05	0,07	0,04	0,51	0,06	0,04

Таблица 4

Межотраслевое движение в Эстонии ($\hat{q}_{ij}$ и $\hat{b}_i$, %)

Из группы	В группу								
	1	2	3	4	5	6	7	8	9
1	+0,7	76,1	87,1	92,9	89,1	75,0	97,0	99,1	2,9
2	23,9	+5,5	62,4	93,7	96,6	48,2	99,1	89,1	1,2
3	12,9	37,6	+0,6	87,7	92,9	88,3	97,1	94,7	10,7
4	7,1	6,3	12,3	+1,7	33,6	12,1	78,2	51,7	0,2
5	10,9	3,4	7,1	66,4	+1,2	16,9	85,8	58,3	3,3
6	25,0	51,8	11,7	87,9	83,1	+1,8	93,1	84,9	17,0
7	3,0	0,9	2,9	21,8	14,2	6,9	+1,3	16,8	0,6
8	0,9	10,9	5,3	48,3	41,7	15,1	83,2	+2,1	4,0
9	97,1	98,8	89,3	99,8	96,7	83,0	99,4	96,0	+1,9

4. ЧИСЛОВЫЕ ПРИМЕРЫ

Метод максимального правдоподобия позволяют получать $\hat{q}_{ij}$ и в зависимости от того, известны начальные численности групп или нет, $\hat{c}_i$ или $\hat{a}_i$ ($\hat{a}_i$ – оценка $n_i c_i$) и $\hat{g}_i$ ($\hat{g}_i$ – оценка $n_i b_i$), а при известных внутригрупповых потоках v_{ii} еще и $\hat{b}_i$.

Параметр b_i был введен для того, чтобы объяснять большую внутригрупповую мобильность людей, т.е. эффект языка, традиций, навыков группы, где находится человек. Положительность b_i означает, что незнание языка служит ограничением при переезде в другие республики. Однако язык или навыки могут и способствовать перемещению и отрицательная оценка $\hat{b}_i$ либо $\hat{g}_i$ как раз отражает эту помощь или даже выталкивание людей из группы (например, определенное образование).

Очевидно, что люди, имеющие семью, состояние, жилье, передвигаются иначе, чем те, которые такими благами не обладают. Поэтому нужно интерпретировать и эти параметры модели. Так, при заданном q_{ij} – привлекательности группы j для i ; чем больше c_i , тем больше в среднем переходов из группы i , значит c_i можно назвать склонностью к переходам (миграции) людей из i .

Межотраслевое движение в Латвии ($\hat{q}_{ij}$ и $\hat{b}_i$, $\times 1000$)

Из группы	В группу					
	1	2	3	4	5	6
1	-414	32	20	4	21	7
2	968	+194	350	35	378	99
3	980	650	+172	130	708	401
4	996	965	870	+133	907	865
5	979	622	292	93	+310	202
6	993	901	599	135	798	+360
7	996	910	842	192	891	795
8	999	983	932	200	949	879
9	999	982	992	1000	993	864
10	993	837	867	523	942	826
11	1000	982	991	1000	992	985
12	997	963	932	98	943	855

Из группы	В группу					
	7	8	9	10	11	12
1	4	1	1	7	0	3
2	90	17	18	163	18	37
3	158	68	8	133	9	68
4	808	800	0	477	0	902
5	109	51	7	58	8	57
6	205	121	136	174	15	145
7	+73	141	0	0	0	0
8	859	+840	115	533	118	887
9	1000	885	+203	958	515	912
10	1000	467	42	+286	42	889
11	1000	882	485	958	+546	316
12	1000	113	88	111	684	+36

Расчеты были проведены для межреспубликанской миграции СССР в 1967 [9], 1973 [10] и 1989 гг. [11] по межотраслевому движению в ряде регионов бывшего СССР в разные периоды.

В табл. 1 и 2 приведены данные об оценках коэффициентов предпочтения $\hat{q}_{ij}$ и $\hat{b}_i$ в 1967 и 1989 гг., оценки $\hat{b}_i$ помещены на диагонали таблиц и помечены знаками "+" или "-". В табл. 3 показаны значения коэффициентов $\hat{c}_i$ – склонностей к миграции в 1967, 1973 и 1989 гг.

Наиболее высокая склонность к миграции все три года была у граждан Российской Федерации; в 1967 и 1973 гг. – жителей Казахстана, Украины и Киргизии. В 1989 г. общая склонность к миграции упала, хотя неодинаково для разных республик: на первом месте, как и ранее – Российская Федерация, далее – Армения, Украина и Азербайджан. Предпочтения людей в 1967 и 1973 гг. практически постоянны, а в 1989 г. резко изменились особенно в Армении и Азербайджане.

Кроме данных о межреспубликанской миграции сделаны расчеты по межотраслевому движению в трех регионах СССР. Результаты приведены в табл. 4–6. Данные об оценке склонностей к переходам приведены – в табл. 7.

Таблица 6

Межотраслевое движение в Краснодарском крае ($\hat{q}_{ij}$ и $\hat{b}_i$, $\times 1000$)

Из группы	В группу					
	1	2	3	4	5	6
1	+7	267	957	200	957	968
2	733	+24	942	507	975	983
3	43	58	+24	96	480	327
4	800	493	904	+4	939	971
5	43	25	520	61	+18	472
6	32	17	673	29	528	+18
7	15	107	646	141	660	581
8	17	19	277	28	238	166
9	63	100	846	82	894	934
10	33	106	853	105	879	947
11	1	1	7	1	4	6
12	16	25	94	09	131	104

Из группы	В группу					
	7	8	9	10	11	12
1	985	983	937	967	999	984
2	893	981	900	894	999	985
3	354	723	154	147	993	906
4	859	972	918	895	999	991
5	340	762	106	121	996	869
6	419	834	66	53	993	896
7	+31	768	314	273	997	933
8	232	+56	66	132	991	742
9	686	934	-04	717	999	990
10	727	868	283	+21	999	972
11	3	9	1	1	-1027	15
12	67	258	10	28	985	+15

Таблица 7

Склонности к переходам $\hat{c}_i$ по регионам, %

Для групп	Группа											
	1	2	3	4	5	6	7	8	9	10	11	12
Эстонии	0,6	0,5	0,9	2,0	2,0	1,2	3,5	2,4	0,3			
Латвии	115	19	12	5,3	28	12	3,5	13	13	10,3	17	24,3
Краснодарский край	1,4	1,1	2,2	1,5	2,8	4,3	1,6	4,1	2,9	1,1	263,2	2,2
Социальные	11	15	21	33	2409	34	27*					

* Для социальных групп приведены $\hat{a}_i$.

Социальное движение ($\hat{q}_{ij}$ и $\hat{g}_i$, %)

Из группы	В группу						
	1	2	3	4	5	6	7
1	+33,1	59,8	86,6	93,5	100,0	100,0	100,0
2	40,2	+19,1	66,0	88,4	99,9	95,8	96,6
3	13,4	34,0	+7,9	70,2	99,7	83,0	77,3
4	6,5	11,6	29,8	+4,4	99,5	57,7	28,9
5	0,0	0,1	0,3	0,5	-1157	0,7	0,5
6	0,0	4,2	17,0	42,3	99,3	+14,0	36,7
7	0,0	3,4	22,7	71,1	99,5	63,3	+13,6

Таблица 9

Критерии качества

Системы групп	R^2	z	l	$t - l - 1$	1% квантильный
Социальные	0,910	4,7	28	13	3,53
Эстония	0,975	23,1	45	26	2,39
Латвия	0,760	2,2	78	53	1,81
Краснодарский край (12)	0,963	17,7	78	53	1,81
СССР 1967	0,951	14,3	120	89	1,63
СССР 1973	0,934	10,5	120	89	1,63
СССР 1989	0,911	7,6	120	89	1,63
Краснодарский край (20)	0,796	3,1	210	169	1,41
США	0,918	1,1	10	1	6055,8
Венгрия	0,861	2,4	21	8	5,36

В Эстонии были выделены следующие группы – отрасли промышленности (табл. 4): 1) электроэнергетика, 2) топливная, 3) химическая, 4) машиностроение, 5) лесная и бумажная, 6) стройматериалов, 7) легкая, 8) пищевая, 9) прочие. В Латвии (табл. 5): 1) промышленность, 2) строительство, 3) транспорт, 4) связь, 5) жилищное и коммунальное хозяйство, 6) здравоохранение, 7) просвещение, 8) культура, 9) искусство, 10) наука, 11) кредитование, 12) управление. В Краснодарском крае (табл. 6): 1) нефтегазопереработка, 2) энергетика, 3) машиностроение, 4) химическая, 5) лесная, 6) стройматериалов, 7) легкая, 8) пищевая, 9) обработка цветных металлов, 10) бытовое обслуживание, 11) строительство, 12) сельское хозяйство.

Очевидно, что процессы миграции и межотраслевого движения имеют общие черты, так как в обоих случаях (кроме одного исключения) наблюдаются также отрицательные оценки $\hat{b}_i$.

В табл. 8 приводятся результаты расчетов по данным о социальном движении [4] между группами: 1) профессиональные руководители высокого уровня; 2) руководители и административные работники; 3) руководители высокого уровня в сфере умственного труда; 4) руководители в сфере умственного труда; 5) работники простого умственного труда и квалифицированные работники физического труда; 6) полуквалифицированные работники физического труда; 7) неквалифицированные работники физического труда.

Примерно такие же результаты получены по данным о межрегиональном движении в Нидерландах, Франции, Финляндии, США, Италии, Израиле и др. Критерий z позволяет считать, что условия в разных районах некоторых стран одинаковы (например, в США, Канаде, ФРГ, Болгарии и Венгрии). Если это для США ($k = 4$), Венгрии ($k = 6$) и Болгарии ($k = 7$) можно объяснить просто малым числом k выделенных регионов, то

выводы о Канаде и ФРГ нуждаются в уточнении. Значения критериев и некоторых параметров для отдельных систем групп показаны в табл. 9. Данные о миграции в упомянутых странах взяты из публикаций IASA (например, для США из [12]).

ЗАКЛЮЧЕНИЕ

Оценки параметров, полученные методом максимального правдоподобия, обладают очень хорошими свойствами, в частности они асимптотически нормальны, не смещены и эффективны. Из приведенных соотношений легко найти их дисперсии, которые могут пригодиться для проверки, например гипотезы об однородности параметров во времени для межреспубликанского движения в СССР в 1967 и 1973 гг. Более того, дисперсии оценок можно найти при любых ранее упомянутых потерях информации.

Ряд других исследований позволяет надеяться на то, что склонности к переходам слабо зависят от групп. Например, в [3] эта склонность полагалась постоянной для всех групп. Очень часто, когда делаются попытки определить величины потоков между группами на основе знания факторов движения, делается аналогичное предположение, а полученные результаты обычно подтверждают согласие с наблюдениями. Все это говорит в пользу того, что склонности к переходам не зависят от группы, которую люди покидают.

Есть еще одна возможность проверить это предположение. Она состоит в том, что рассчитать оценки склонностей к переходам и коэффициентов $\hat{q}_{ij}$ при гипотезе, что интенсивности переходов между группами $r_{ij} = cq_{ij}$, $i \neq j$, и определить, как было сделано, критерий z . Если z окажется велик, то предположение подтверждается.

Однако этот подход требует по крайней мере информации о численности всех групп, поскольку интенсивности переходов могут быть оценены лишь как отношения численностей людей, перешедших из группы i в j , к численности группы i . Но в настоящей работе этого не потребовалось. Информация о внутригрупповых переходах желательна, поскольку без нее нельзя получить ряда выводов.

Другая задача, подход к которой теперь представляется возможным, состоит в учете емкости групп. Ограниченная емкость не позволяет всем желающим, количество которых пропорционально коэффициентам привлекательности, попасть туда, куда они хотели бы. Наблюдаемое число переходов — это лишь доля всех желающих с коэффициентом пропорциональности, зависящим от группы, в которую человек хочет попасть. Число желающих определяется и склонностью к переходам. Если эта склонность постоянна, то возможен подход, аналогичный предположенному в данной работе, поэтому решение предшествующей задачи очень важно.

Возникает еще одна задача, имеющая фундаментальное значение: что играет большую роль в решении людей поменять группу — перспектива улучшения своего положения, как предполагалось в этой статье, или его среднего улучшения. Ответ можно получить, если иметь данные о распределении каких-либо благ в группах, например данные о распределении доходов или заработков для межотраслевой миграции или о распределении жилья при межрегиональной. Наши выводы свидетельствуют о том, что люди рассматривают главным образом на улучшение своего положения.

Может быть поставлена и решена еще одна проблема — определения распределения полезности в группе по данным о движении населения. Более того, если известно распределение благ в группе, например заработков и жилья, и именно они будут аргументами функции полезности, то по известным q_{ij} , вероятно, можно оценить и параметры самой функции.

ЛИТЕРАТУРА

1. Rogers A. Introduction to Mathematical Demography. New York; Toronto et al.: John Willy, 1975.
2. Filds G. Place-to-Place Migration: Some New Evidences // Rev. Econom. and Statist. 1979. V. 61. № 1.
3. Weidlich W., Haag G. Interregional Migration. Berlin: Springer-Verlag, 1988.

4. Бартоломью Дж. Стохастические модели социальных процессов. М.: Финансы и статистика, 1985.
5. Holmlund B. Labour Mobility. Stokholm: IUI, 1984.
6. Pao C.P. Линейные методы статистики и их приложения. М.: Наука, 1968.
7. Староверов О.В. Модели движения населения. М.: Наука, 1979.
8. Coussinus H. Quelque points de vue sur l'analyse des talleaux d'échanges // Annals de L'ISEE. 1976. № 22-23.
9. Механическое движение населения СССР за 1967 год // Вестн. статистики. 1968. № 10.
10. Миграция населения СССР за 1973 год // Вестн. статистики. 1975. № 2.
11. Демографический ежегодник. М.: Финансы и статистика, 1980.
12. Long L.H., Frey W.H. Migration and Settlement. United States. Laxenburg: IIASA, RR – 82-15, 1982.

Поступила в редакцию
21 IV.1993