УДК 519.24, 316.4

СТОХАСТИЧЕСКАЯ МОДЕЛЬ СТРУКТУРЫ СОЦИАЛЬНОЙ ГРУППЫ[1]

Савельев Л.Я., Гончарова Г.С.
Новосибирский государственный университет, Новосибирск, Тувинский государственный университет, НИЛ «Теории вероятностей и ее приложений», Кызыл,
Институт философии и права СО РАН, Новосибирск

THE STOCHASTIC MODELS OF SOCIAL GROUPS STRUCTURE

Savelyev L.J., Goncharova G.S.
Novosibirsk State University, Novosibirsk, Tuvan State University, Laboratory "Theory of Probability and its Applications", Kyzyl. Institute of Philosophy and Law, Russian Academy of Sciences, Novosibirsk

В статье рассматриваются простая марковская стохастическая модель социальной группы и примеры ее применения к исследованию динамики структуры конкретных групп и ее прогноза. Описываются группы учащейся молодежи некоторых регионов Сибири. Исследуются условия достижения выбранной стационарной структуры рассматриваемой социальной группы.

Ключевые слова: вероятность, марковское свойство, социокультурные типы, коллективисты, индивидуалисты.

The article discusses a simple stochastic Markov model of a social group, and examples of its application to the study of the dynamics of the structure of particular groups and its prognosis. Describes a group of students in some regions of Siberia. We study the conditions for achieving the selected stationary structure considered a social group.

Key words: probability, Markov property, socio-cultural type, collectivists, individualists.

Введение. Рассматриваются формализованные классификации индивидов данного социума по выбранным личным, общественным и культурным признакам. В простейшей двоичной классификации по каждому выбранному признаку индивид относится к одному из двух уровней: высокому или низкому. В более подробной троичной классификации по каждому выбранному признаку индивид относится к одному из трех уровней: высокому, среднему или низкому. Содержание выбранных признаков и их уровней определяется рассматриваемым социумом и поставленной задачей. Предлагается стохастическая модель, позволяющая описывать состояние социума в данный момент и его динамику в рассматриваемом индивидуально-социально-культурном пространстве, а также делать краткосрочный, среднесрочный и долгосрочный прогнозы состояния

[1] Материалы статьи подготовлены по гранту РГНФ № 12-03-00546 а. «Социокультурные типы молодежи современной России: этническое и региональное измерение»

социума. Термин классификация употребляется в математическом смысле: установление отношения эквивалентности или разбиение множества на части.

Среди стохастических моделей, описывающих различные естественные и социальные процессы, выделяются марковские. Они отличаются сравнительной простотой и часто с успехом применяются. Для описания социальных явлений эффективно используются марковские процессы с дискретным временем и конечным множеством состояний, называемые конечными марковскими цепями [1; 3]. Определения основных понятий элементарной теории вероятностей есть в книге [4]. Важность и эффективность применения статистических методов в социологии подробно описаны в книге [5].

Использование предлагаемых простых математических моделей предполагает существенное упрощение выбранных характеристик сложных представлений о личности, обществе и культуре. Но простые модели позволяют достаточно адекватно описывать некоторые фундаментальные закономерности структуры рассматриваемых общественных групп, связанные с личнымикачествами составляющих их индивидов, социальными и культурными особенностями группы. Деление на два уровня дает возможность по имеющимся неполным данным достаточно уверенно относить индивида в одну из двух групп. Если двух уровней мало для решения поставленной задачи и нужно более подробная характеристика составляющих группу индивидов, их личных, социальных и культурных качеств, то можно использовать три уровня. Деление на три уровня часто тоже дает возможность по имеющимся неполным данным достаточно уверенно относить индивида в одну из трех групп.

В теории вероятностей и математической статистике разработаны модели с любым числом признаков и любым числом уровней. Если есть надежный статистический материал, то можно использовать более общие модели, чем предложенная двоичная и троичная. Использование модели с большим числом признаков и большим числом уровней связано с большим объемом вычислений. Применение современных компьютерных программ позволяет с ним справиться. Главными недостатками применения моделей с большим числом признаков и большим числом уровней является трудность учета взаимной зависимости выбранных признаков и размытость границ выбранных уровней. Кроме того, адекватность модели существенно снижается использованием неточных описательных определений рассматриваемых признаков и уровней в имеющихся материалах. Эти недостатки существенно снижают адекватность моделей с большим числом признаков и большим числом уровней. Наконец, большое число признаков и уровней затрудняет содержательную интерпретацию полученных результатов. Часто для надежной интерпретации приходится группировать признаки и уровни, получая те же результаты, что и при изначальном использовании модели с небольшим числом признаков и небольшим числом уровней.

Особенно хорошо разработаны марковские стохастические модели. Их применение связано с теоретическим или статистическим определением вероятностей перехода рассматриваемой системы из одного выделенного состояния в другое. Определение таких вероятностей обычно требует большого объема статистических данных. Часто нужных данных нет, а их получение связано со значительными трудностями. Для моделей с большим числом признаков и большим числом уровней эти трудности нередко бывают непреодолимы.

Оценить такие трудности можно на примере классификации по трем признакам (личные, общественные и культурные качества) и пяти уровням (очень высокий, высокий, средний, низкий, очень низкий). Точное описание составленной из получающихся 125 групп структуры и ее динамики представляется сложной проблемой. Два уровня (высокий, низкий) дают 8 групп, а три уровня (высокий, средний, низкий) – 27 групп. В этих случаях можно рассчитывать на успех.

Серьезным ограничением для применения марковских стохастических моделей является необходимость выполнения марковского свойства: распределение состояний системы в следующий момент зависит только от ее состояния в данный момент и не зависит от состояний в предыдущие моменты. Марковское свойство часто поясняется фразой: при известном настоящем прошлое и будущее стохастическинезависимы. Термины состояние и момент имеют очень общий смысл, и поэтому модель может описывать широкий круг процессов любой природы. Этим объясняется широкое применение марковских стохастических моделей в самых разных областях. Исключительно важным свойством этих моделей является возможность прогнозировать поведение рассматриваемой системы. Марковские прогнозы часто хорошо согласуются с реальностью.

1. Марковские случайные последовательности. Содержание понятий событие, случайность, вероятность хорошо знакомо каждому. Опишем кратко простейшую математическую модель, используемую в статье.

Основным понятием простейшей стохастической модели является *вероятностное пространство* $\{U,p\}$, составленное из конечного множества $U=\{u\}$ *исходов* u и семейства $p=\{p[u]\}$ *элементарных вероятностей* $p[u]\geq 0,\quad \sum p[u]=1$. Исходами могут быть элементы любой природы. Часто их удобно нумеровать, отождествлять исходы с их номерами и считать $U=\{1,2,\dots,m\}$, где $m=n[U]$ число исходов. Части $A\subseteq U$ множества U называются *событиями*. Сумма $P[A]$ элементарных вероятностей $p[u]$, составляющих событие A исходов u, называется *вероятностью* этого события: $P[A]=\sum_{u\in A}p[u]$. Ясно, что $0\leq P[A]\leq 1$. В частности, вероятность пустого (невозможного) события $\emptyset$ равна нулю, а вероятность полного (*достоверного*) события U равна единице: $P[\emptyset]=0,\quad P[U]=1$. Вместо P иногда пишут $\Pr\square$. Числа $p[u]$ и $P[A]$ служат *мерами реализуемости* исхода u и события A.

Для вероятности, P верно, *правило сложения*: $P[A+B]=P[A]+P[B]$ для *непересекающихся* (не имеющих общих исходов) событий A, B. Равенство $P[A|B]=\dfrac{P[AB]}{P[B]}$ выражает условную вероятность события A при условии, что произошло событие B. Верно *правило умножения*: $P[AB]=P[A|B]P[B]$. Здесь сумма А + В обозначает *объединение* событий A, B, а произведение AB - их *пересечение* (общую часть). Предполагается, что $P[B]>0$ (событие B *возможно*). Если $P[AB]=p[A]P[B]$, то события A и B называются *стохастически независимыми* или просто *независимыми*. Независимость событий A и B выражается также равенством $P[A|B]=P[A]$. Понятие независимости является фундаментальным для теории вероятностей.

Любая функция на множестве исходов называется *случайной переменной*. Она может принимать абстрактные значения. Выделяются случайные переменные со *шкалированными* значениями, в частности *числовыми*. Числовые случайные переменные называются случайными величинами. Они являются главным объектом теории вероятностей. Активно развивается теория *возможностей* для шкалированных случайных переменных [6]. Применения теории шкал в социологии подробно описаны в книге [7].

Случайная величина f на вероятностном пространстве $\{U,p\}$ порождает вероятностное пространство $\{X,q\}$, составленное из множества $X=\{x\}$ ее значений $x=f[u]$ и семейства $q=\{[q]\}$ элементарных вероятностей $q[x]=\sum_{f[u]=x}p[u]$ (складываются элементарные вероятности $p[u]$ всех исходов u, для которых $f[u]=x$). Ясно, что $q[x]\geq 0$, $\sum q[x]=1$. Числоравно в пространстве $\{U,p\}$ вероятности $P[u{:}\,f[u]=x]$ того, что случайная величина f примет значение x. Семейство $q=\{q[x]\}$ называется *распределением* случайной величины f.

Важнейшей характеристикой случайной величины f является ее *среднеезначение* $Ef=\sum_{u\in U}f[u]p[u]$. По связанной с азартными играми традиции число Ef называют также *математическим ожиданием*. Верно удобное для вычислений равенство $Ef=\sum_{x\in X}xq[x]$. Для оценки отклонений значений случайной величины f от среднего значения $a=Ef$ часто применяется ее *дисперсия*, равная среднему квадратичному отклонению: $Df=E[(f-a)^2]$. Верно удобное для вычислений равенство $Df=\sum_{x\in X}(x-a)^2q[x]$. Используется также *стандартное отклонение* $\sigma=\sqrt{Df}$. С ним связано знаменитое *правило трех сигм*: значения

$x = f[u]$, для которых $|x - a| > 3\sigma$, считаются невозможными. Это правило часто применяют необоснованно.

Для лучшего восприятия вероятностной модели полезна ее механическая интерпретация. Семейство $p = \{p[u]\}$ описывает распределение единицы массы на множестве U: в точке u сосредоточена масса $p[u]$, а на множестве A сосредоточена масса $P[A]$. Случайная величина f переносит массу на множество X вещественной прямой: в точке x сосредотачивается масса $q[x]$. Среднее значение Ef случайной величины f является *центром масс*, а дисперсия Df – *моментом инерции* такой механической системы. Применение вероятностных моделей в статистике обосновывает *закон больших чисел*, утверждающий, что при определенных условиях вероятность события в определенном смысле близка частоте его реализаций.

Семейства случайных переменных исследует теория случайных процессов. Элементы этой теории вместе с основами теории вероятностей излагаются в книге [8]. Опишем кратко используемое в статье достаточно простое семейство случайных переменных, не уточняя вероятностное пространство, на котором они определены. Рассмотрим последовательность ξ случайных переменных $\xi(t), t = 0, 1, 2, \dots$ с конечным множеством значений. Эти значения могут быть любыми, но удобно их перенумеровать и заменить номерами. Множеством значений тогда становится конечное множество номеров $C = \{1, 2, \dots, \}$. Последовательность ξ называется *марковской цепью*, если она обладает *марковским свойством*: при известном настоящем $\xi(t)$ прошлое $\{\xi(s), s < t\}$ и будущее $\{\xi(u), \quad u > t\}$ стохастически независимы.

В рассматриваемом конечном случае марковское свойство выражается в том, что совместные распределения конечных семейств $\{\xi(t), \quad 0 \le t \le n\}$ задаются произведениями матриц, составленных из начальных и переходных вероятностей. Строка $A = \{a_i\}$ $(i = 1, \dots, m)$ начальных вероятностей состоит из вероятностей $a_i = Pr[\xi(\mathbf{0}) = i]$ того, что в начальный момент $t = \mathbf{0}$ значением случайной величины $\xi(0)$ будет номер i. Матрица $Q = \{q_{ij}\}$ переходных вероятностей состоит из условных вероятностей $q_{ij} = \Pr[\xi(t+1) = j | \xi(t) = i]$ того, что в следующий момент $t + \mathbf{1}$ значением случайной величины $\xi(t+1)$ будет номер j, если в данный момент t значением случайной величины $\xi(t)$ является номер i. Подчеркнем, что здесь термин *момент* служит синонимом термина *номер* или термина *индекс* и может не иметь никакого отношения ко времени.

Распределение случайной величины $\xi(t)$ выражает строка $p^t = \{p_i^t\}$, составленная из вероятностей $p_i^t = \Pr[\xi(t) = i]$ того, что в момент t значением случайной величины $\xi(t)$ является номер i. В частности, распределение случайной величины $\xi(0)$ выражается равенствами $P^{\mathbf{0}} = A, \quad a_i = p_i^o$ $(i = 1, \dots, m)$. Замена состояний их номерами упрощает обозначения. Состояния могут описываться целыми фразами или сложными графическими изображениями. Матрицы умножаются по правилу *строка на столбец:*

$$p^{t+1} = p^t Q \ , \qquad p_j^{t+1} = \sum_{i=1}^{m} p_i^t q_{ij} \ \ (j = 1, \dots, m)$$

Переход через k моментов, от t к $t+k$ описывается матричным равенством $p^{t+k} = p^t Q^k$. В частности, распределение случайной величины $\xi(u)$ выражается равенством $p^u = AQ^u$. При $m = 2, 3$ эти распределения легко вычисляются в общем виде [4]. Получены также общие формулы для их важных характеристик, связанных со структурой серий в реализациях таких марковских последовательностях.

2. Двоичная и троичная формальные классификации. Опишем подробнее предлагаемые классификации индивидов данного социума по личным (P), социальным (S) и культурным (K) признакам.

Двоичная классификация. По каждому признаку P, S, K социум делится на типы A и B. В результате получаются группы PA, PB, SA, SB, KA, KB. Тип A описывает высокий уровень индивида (личный, социальный, культурный), тип B - низкий. Конкретное содержание этих типов определяется в зависимости от рассматриваемого социума и поставленной задачи.

Из полученных групп образуются 8 классов $AAA, AAB, ABA, BAA, ABB, BAB, BBA, BBB$. В класс AAA относятся индивиды с высоким уровнем по всем трем признакам. В класс AAB относятся индивиды с высоким уровнем по первым двум признакам и низким по третьему. Остальные классы тоже легко интерпретируются. В последний класс BBB относятся индивиды с низким уровнем по всем трем признакам.

Троичная классификация. По каждому признаку P, S, K социум делится на типы A, C, B. В результате получаются группы $PA, PC, PB, SA, SC, SB, KA, KC, KB$. Тип A описывает высокий уровень индивида (личный, социальный, культурный), тип C – средний, тип B – низкий. Конкретное содержание этих типов определяется в зависимости от рассматриваемого социума и поставленной задачи. Уровни качеств можно определять, не уточняя общие определения самих качеств. Это существенно облегчает задачу.

Из полученных групп образуются 27 классов: $AAA, AAC, AAB, ACA, ACC, ACB, ABA, ABC, ABB, CAA, CAC, CAB, CCA, CCC, CCB, CBA, CBC, CBB, BAA, BAC, BAB, BCA, BCC, BCB, BBA, BBC, BBB$. В класс относятся индивиды с высоким уровнем по всем трем признакам. В класс AAC относятся индивиды с высоким уровнем по первым двум признакам и средним по третьему. В класс AAB относятся индивиды с высоким уровнем по первым двум признакам и низким по третьему. Остальные классы тоже легко интерпретируются. В последний класс относятся индивиды с низким уровнем по всем трем признакам. Отступление от алфавитного порядка объясняется соответствием обозначений предложенных классификаций: всюду A – высокий уровень, а B – низкий.

От троичной классификации можно перейти к двоичной, если разделить множество индивидов уровня C на два множества λC и $(1-\lambda)C$, $0 \le \lambda \le 1$, присоединив λ – долю множества к множеству A, а оставшуюся $(1-\lambda)$ – долю – к множеству В. Объединенные группы $\overline{A} = A + \lambda C$ и $\overline{B} = B + (1-\lambda)C$ определят новые высокий и низкий уровни в двоичной классификации.

Выбранный порядок классов в определенной степени отражает их значимость: от высоких уровней к более низким. Это позволяет при анализе результатов переходить от абсолютной шкалы к порядковой. При одинаковом конечном числе градаций эти шкалы асимптотически эквивалентны [9]. Применение марковской модели дает возможность исследовать динамику психологической, социальной и культурной структур выделенного общественного слоя. Можно выделить периоды стабильности и делать краткосрочный, среднесрочный и долгосрочный прогнозы.

Общая классификация. В общем случае выбранное множество делится по определенным признакам на n попарно непересекающихся (не имеющих общих элементов) частей, которым присваиваются номера 1, 2, ..., n. Назовем эти части *большими*. Каждая большая часть в свою очередь по определенным правилам делится на m попарно непересекающихся *мелких* частей, которым присваиваются внутренние номера 1, 2, ..., m. Кроме того, мелкой части с внутренним номером i из большой части с номером j присваивается *двойной номер* ij. В итоге множество разбивается на m^n частей. В частности, при $n = 2, m = 2$ получается частей; при $n = 3, m = 2$ получается частей; при $n = 3$, $m = 3$ получается частей. Содержание такой модели может быть самым разным. Деление больших частей на различные числа маленьких технически усложняет модель.

В последнее время стали иногда вместо множеств рассматривать функции со значениями в отрезке [0,1], описывающие *степень принадлежности* элемента множеству [6]. Такие функции называются *нечеткими множествами*. Нередко они более адекватно описывают содержание задачи, но их корректное применение связано с математическими трудностями.

3. Марковская модель с 2 состояниями. Рассмотрим двоичную марковскую последовательность ξ случайных переменных $\xi(t)$, $t \ge 0$ с множеством значений $C = \{1,2\}$ (номера состояний), начальной и переходной матрицами

$$A = \{a_1, a_2\} = \{a, 1-a\}, \quad Q = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}$$

Здесь $a_1 = Pr\{\xi[0] = i\}$ вероятность появления в начальный момент номера i, а $q_{ij} = Pr\{\xi[t+1] = j | \xi[t] = i\}$ вероятность появления в момент $t+1$ номера j при условии, что в момент t появился номер i $(1 \le i, j \le 2)$. Будем использовать более простые обозначения a, p, q и $d = p + q - 1$, $b = \dfrac{1-q}{1-d}$.

Нетрудно доказать, что верно следующее выражение для t - ой степени Q^t матрицы Q :

$$Q^t = \begin{pmatrix} b + (1-b)d^t & 1 - b - (1-b)d^t \\ b - bd^t & 1 - b + bd^t \end{pmatrix}$$

Из матричного равенства $p^u = AQ^u$ следуют формулы

$$p_1^t = \Pr[\xi(t) = 1] = b + (a-b)d^t, \qquad p_2^t = \Pr[\xi(t) = 2] = 1 - b + (a-b)d^t$$

для вероятностей появления в момент t номеров $j = 1, 2$. Если $|d| < 1$, то при достаточно больших значениях t верны приближенные равенства $p_1^t \approx b$, $p_2^t \approx 1 - b$.

Постепенно последовательность $\xi(t)$ стабилизируется. Распределение случайной величины $\xi(t)$ все ближе к $\{b, 1-b\}$ и все меньше зависит от начального распределения $\{a, 1-a\}$.

Распределение $\{b, 1-b\}$*стационарно*: $\{b, 1-b\}.Q = \{b, 1-b\}$. Если в некоторый моменты распределение случайной величины $\xi(s)$ оказывается равным $\{b, 1-b\}$, то это распределение имеют и все следующие случайные величины $\xi(t)$, $t > s$. Процесс переходит в стационарный режим. Чем меньше $|d| < 1$, тем быстрее убывает $|d|^t$ и быстрее достигается стационарный режим. Выделяется случай $p + q = 1$. В этом случае $d = 0$ и $p_1^t = b = p$ при $t > 0$ и любом начальном распределении, $1 - a$ }. Марковская последовательность превращается в последовательность независимых случайных переменных. Стационарный режим достигается уже в первый момент.

В статье [10, с. 82] описываются 6 социокультурных типов молодежи Республики Саха (Якутия) в целом, а также в этнических группах русских, саха (якутов) и группы КМНС (коренных малочисленных народностей Севера). В разработанную авторами специальную анкету включены вопросы об отношении к обществу, о важности осознания своей национальной принадлежности, знания языка и культуры своего народа. По ответам на эти вопросы были выделены следующие типы:

(1) коллективист с развитым этническим самосознанием, (2) коллективист с менее развитым этническим самосознанием, (3) коллективист с неразвитым этническим самосознанием, (4) индивидуалист с развитым этническим

самосознанием, (5) индивидуалист с менее развитым этническим самосознанием, (6) индивидуалист с неразвитым этническим самосознанием.

Если интересоваться отношением индивида к обществу и отвлечься от этнического самосознания, то естественно объединить типы 1 – 3 в общий тип (*коллективист*), а типы 4 – 6 объединить в общий тип (*индивидуалист*). Для этнических групп русских, саха (якутов) и коренных малочисленных народов Севера (КМНС) доли коллективистов к моменту окончания анкетирования (2010 год) были соответственно равны 0.71, 0.75, 0.87. Статистических данных для определения переходных вероятностей нет. Поэтому рассмотрим различные гипотетические варианты течения процесса.

Предположим, что за интервал времени между моментами t и $t+1$ коллективисты и индивидуалисты с одинаковой вероятностью продолжают придерживаться своих взглядов: $p = q \ (0 < p, q < 1)$. Тогда $b = 1 - b = \frac{1}{2}$ и доли коллективистов и индивидуалистов постепенно становятся равными. Обозначим r отношение вероятностей для коллективистов и для индивидуалистов за интервал времени между моментами t и $t+1$ изменить свои взгляды на противоположные: $r = \frac{1-p}{1-q}$. В этом случае доля коллективистов постепенно становятся равной $b = \frac{1}{1+r}$, а доля индивидуалистов постепенно становятся равной $1 - b = \frac{r}{1+r}$. Пусть, например, $p = 0.8$, $q = 0.4$. Тогда $d = 0.2$, $r = \frac{1}{3}$ и $b = \frac{3}{4} = 0.75$. Будем считать такую долю коллективистов нормальной. Так как $(d^t) = (0.2)^t$ быстро убывает, то признанная нормальной доля коллективистов в этнической группе русских при $a = 0.71$ быстро достигается:

$$p_1^1 = 0.742, \quad p_1^{12} = 0.7484, \quad p_1^3 = 0.7497, \quad p_1^4 = 0.7499, \quad p_1^5 = 0.75.$$

В этнической группе саха (якутов) при $a = 0.75$ и таких переходных вероятностях доля коллективистов стационарна: p_1^t=0.75 для всех $t \geq 0$. В этнической групп КМНС при $a = 0.87$ и таких переходных вероятностях доля коллективистов постепенно размываетсядо признанной нормальной:

$$p_1^1 = 0.774, \ p_1^2 = 0.7548, \ p_1^3 = 0.751, \ p_1^4 = 0.7502, \quad p_1^5 = 0.75.$$

Формально интервалы времени между моментами t и $t+1$ могут быть любыми. Выбор этих интервалов определяется только адекватностью описания матрицей Q переходов от состояния в момент t к состоянию в момент $t+1$.

4. Марковская модель с 4 состояниями. Рассмотрим марковскую последовательность ξ случайных переменных $\xi(t)$, $t \geq 0$ с множеством значений $C = \{1, 2, 3, 4\}$ (номера состояний), начальными вероятностями

$a_i = Pr\{\xi[0] = i\}$ и переходными вероятностями $q_{ij} = \Pr\{\xi[t+1] = j|\xi[t] = 1\}$, составляющими 1х4-матрицу (строку) $A = \{a_i\}$ и 4х4 - матрицу $Q = \{q_{ij}\}$. Здесь, как и прежде, a_i – вероятность появления в начальный момент состояния с номером i, а q_{ij} - вероятность появления в момент $t+1$ состояния с номером j при условии, что в момент t появилось состояние с номером $i(1 \leq i, \quad j \leq 4)$. Можно выписать общие формулы для Q^t и $P^t = AQ^t$, но они громоздкие. Проще рассматривать нужные частные случаи.

Если исследовать отношение индивида к обществу и его этническое самосознание, не выделяя отношение к языку и культуре, то естественно объединить типы 2 - 3 и точно так же объединить типы 5 - 6 . В итоге получаются 4 типа: - коллективист с развитым этническим самосознанием, - коллективист с недостаточно развитым этническим самосознанием, - индивидуалист с развитым этническим самосознанием. - индивидуалист с недостаточно развитым этническим самосознанием. Занумеруем эти классы в алфавитном порядке: $AA - 1$, $AB - 2$, $BA - 3$, $BB - 4$. Будем считать коллективизм и индивидуализм *признаками* (или на первом месте), а этническое самосознание измерять *уровнями* (или на втором месте).

Для этнических групп русских, саха (якутов) и КМНС доли рассматриваемых классов к моменту окончания анкетирования (2010 год) представлены в таблице 1.

Таблица 1

Распределение социокультурных типов среди учащейся молодежи
в этнических группах Республики Саха (Якутия)

Национальность	Социокультурные типы			
	- коллективист с развитым этническим самосознанием	- коллективист с недостаточно развитым этническим самосознанием	- индивидуалист с развитым этническим самосознанием	- индиви-дуалист с недостаточно развитым этническим самосознанием
Русские	0.42	0.29	0.16	0.13
Саха (якуты)	0.63	0.12	0.19	0.06
КМНС	0.67	0.2	0.08	0.05

Статистических данных для определения переходных вероятностей нет. Поэтому рассмотрим различные гипотетические варианты течения процесса.

Сделаем упрощающие предположения. Будем считать, что вероятность перехода от признака или уровня к признаку или уровню равна p, а вероятность обратного перехода равна q . Одновременную смену признака и уровня будем считать невозможной. При этих условиях матрица переходных вероятностей имеет вид$(r = p + q)$:

$$Q = \begin{pmatrix} 1-2p & p & p & 0 \\ q & 1-r & 0 & p \\ q & 0 & 1-r & p \\ 0 & q & q & 1-2q \end{pmatrix}$$

Рассмотрим произвольную строку начальных вероятностей с безиндексными обозначениями $A = \{\propto, \quad \beta, \quad \gamma, \quad 1-\alpha-\beta-\gamma\}$. Все ненулевые элементы матрицы Q и строки A больше 0 и меньше 1. Будем предполагать, что $0 < p, \quad q < \frac{1}{2}$. Естественно выделяется частный случай $q = p$, когда переходы отк и обратно одинаково вероятны. В этом случае матрица Q симметрична и вычисления проще.

Стационарным является распределение

$$B = \{({}^{q}/_{r})^2, \quad ({}^{p}/_{r})({}^{q}/_{r}), \quad ({}^{p}/_{r})({}^{q}/_{r}), \quad ({}^{p}/_{r})^2\}$$

Равенство $BQ = B$ легко проверяется. По стандартным правилам вычисляется определитель матрицы Q : $d = (1-r)^2\,(1-2r)$. Степени множителей $1-r, 1-2r$ входят в выражения для разности распределения случайной величины $\xi[t]$и стационарного. Если эти множители достаточно малы, то стационарное распределение быстро достигается. В случае $q = p$, когда переходы от к и обратно одинаково вероятны, верны равенства $\frac{p}{r} = \frac{q}{r} = \frac{1}{2}$ и стационарное распределение равномерно: $B = \left\{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right\}$. Доли всех рассматриваемых классов становятся одинаковыми.

Рассмотрим сначала распределение для этнической группы русских. Пусть, например, $p = 0.1, \quad q = 0.4$. Тогда $1-r = 0.5, \quad 1-2r = 0$ и $b = \{0.64, \quad 0.16, \quad 0.16, \quad 0.04\}$. Будем считать такие доли рассматриваемых классов нормальными. Так как множители $1-r, \quad 1-2r$ малы, то признанные нормальными доли классов достигаются быстро и при сохранении переходных вероятностей в дальнейшем не меняются. Они выписаны в таблице 2.

При тех же переходных вероятностях динамику распределений в этнических группах саха (якутов) и КМНС описывают таблицы 3 и 4.

Таблица 2

Гипотетическая динамика изменения социокультурных типов у русских до признанного нормальным при заданных переходных вероятностях $p = 0.1, \quad q = 0.4$

Время t	Социокультурные типы			
	- коллективист с развитым этническим самосознанием	- коллективист с недостаточно развитым этническим самосознанием	- индивидуалист с развитым этническим самосознанием	- индивидуалист с недостаточно развитым этническим самосознанием
$t+1$	0.52	0.24	0.17	0.07
$t+2$	0.58	0.2	0.17	0.06
$t+3$	0.61	0.18	0.16	0.05
$t+4$	0.62	0.17	0.16	0.04
$t+5$	0.63	0.16	0.16	0.04
$t+6$	0.64	0.16	0.16	0.04

При тех же переходных вероятностях динамику распределений в этнических группах саха (якутов) и КМНС описывают таблицы 3 и 4.

Таблица 3

Гипотетическая динамика изменения социокультурных типов у саха до признанного нормальным при заданных переходных вероятностях $p = 0.1, \quad q = 0.4$

Время t	Социокультурные типы			
	- коллективист с развитым этническим самосознанием	- коллективист с недостаточно развитым этническим самосознанием	- индивидуалист с развитым этническим самосознанием	- индивидуалист с недостаточно развитым этническим самосознанием
$t+1$	0.63	0.15	0.18	0.04
$t+2$	0.63	0.15	0.17	0.04
$t+3$	0.64	0.16	0.17	0.04
$t+4$	0.64	0.16	0.16	0.04

Таблица 4

Гипотетическая динамика изменения социокультурных типов у коренных малочисленных народов Севера до признанного нормальным при заданных переходных вероятностях

$p = 0.1, \quad q = 0.4$

Время t	Социокультурные типы			
	- коллективист с развитым этническим самосознанием	- коллективист с недостаточно развитым этническим самосознанием	- индивидуалист с развитым этническим самосознанием	- индивидуалист с недостаточно развитым этническим самосознанием
$t+1$	0.65	0.19	0.13	0.04
$t+2$	0.64	0.17	0.14	0.04
$t+3$	0.64	0.17	0.15	0.04
$t+4$	0.64	0.16	0.16	0.04

Вычисления всюду проводились с округлением до сотых.

Замечание. Из таблиц 2-4 видно, что у саха и КМНС при заданных одинаковых переходных вероятностях время достижения стационарного распределения социокультурных типов, условно признанных нормальными, совпадает, а у русских – больше на два периода. Это косвенно может говорить о близости личных, социальных и культурных характеристик у саха и КМНС.

5. Марковская модель с 8 состояниями. Рассмотрим марковскую последовательность ξ случайных переменных $\xi(t)$, $t \geq \mathbf{0}$ с множеством значений $C = \{1,2,3,4,5,6,7,8\}$ (номера состояний), начальными вероятностями $a_i = \mathbf{Pr\{\xi[0]\}} = i$ и переходными вероятностями $q_{ij} = \Pr\{\xi[t+1] = j|\xi[t] = 1\}$, составляющими 1x8-матрицу (строку) $A = \{a_i\}$ и 8x8-матрицу $Q = \{q_{ij}\}$. Здесь, как и прежде, a_i - вероятность появления в начальный момент состояния с номером i , а q_{ij} - вероятность появления в момент $t + \mathbf{1}$ состояния с номером j при условии, что в момент t появилось состояние с номером i $(1 \leq i,j \leq 8)$. Будем рассматривать подходящий для выбранного примера частный случай такой модели с достаточно простыми выкладками.

В статье [10, с. 96] приведено распределение группы молодежи Якутии по личным, социальным и культурным признакам, в каждом из которых выделены два уровня. По этой классификации группа была разделена на 8 классов $AAA, AAB, ABA, BAA, ABB, BAB, BBA, BBB$. Классы занумерованы в алфавитном порядке:

$AAA - 1, AAB - 2, ABA - 3, BAA - 4, ABB - 5, BAB - 6, BBA - 7, BBB - \mathbf{8}$.

Отступая от принятой в [9, с. 82] терминологии, можно так описать эти классы. По личным качествам выделялись индивиды с высоким уровнем *активности* (буква на первом месте в обозначении класса). По социальным качествам выделяются индивиды (коллективисты) с высоким уровнем социального взаимодействия (буква на втором месте в обозначении класса). По культурным качествам выделяются индивиды с высоким уровнем этнического самосознания (важно осознавать свою национальную принадлежность, знать язык и культуру своего народа) (буква на третьем месте в обозначении класса). Альтернативы отмечаются буквой на соответствующем месте.

В качестве начального выберем распределение для молодежи 20 – 24 лет: $A = \{0.14, \quad 0.05, \quad 0.05, \quad 0.03, \quad 0.42, \quad 0.11, \quad 0.14, \quad 0.06\}$ Достаточных статистических данных для определения переходных вероятностей нет. Поэтому рассмотрим различные гипотетические варианты течения процесса.

Сделаем упрощающие предположения. Будем считать, что вероятность перехода от признака или уровня к признаку или уровню равна p , а вероятность обратного перехода равна q . Одновременную смену признака и уровня будем считать невозможной. При этих условиях матрица переходных вероятностей имеет вид $(2p + q = r, p + 2q = s)$:

$$Q = \begin{pmatrix} 1-3p & p & p & p & 0 & 0 & 0 & 0 \\ q & 1-r & 0 & 0 & p & p & 0 & 0 \\ q & 0 & 1-r & 0 & p & 0 & p & 0 \\ q & 0 & 0 & 1-r & 0 & p & p & 0 \\ 0 & q & q & 0 & 1-s & 0 & 0 & p \\ 0 & q & 0 & q & 0 & 1-s & 0 & p \\ 0 & 0 & q & q & 0 & 0 & 1-s & p \\ 0 & 0 & 0 & 0 & q & q & q & 1-3q \end{pmatrix}$$

Все ненулевые элементы матрицы Q больше 0 и меньше 1. Будем предполагать, что $0 < p, \quad q < \frac{1}{3}$. Естественно выделяется частный случай $q = p$, когда переходы от к и обратно одинаково вероятны. В этом случае матрица Q симметрична и вычисления проще.

Стационарным является распределение

$$B = \{c^3, \quad bc^2, \quad bc^2, \quad bc^2, b^2c, \quad b^2c, \quad b^2c, \quad b^3\}$$

где $b = \frac{p}{p+q}, \quad c = \frac{q}{p+q}$. Равенство $BQ = B$ легко проверяется. По стандартным правилам вычисляется определитель матрицы Q: $d = (1-u)^3(1-2u)^3(1-3u)$, где $u = p + q$. Степени множителей $1-u, \quad 1-2u, \quad 1-3u$ входят в выражения для разности распределения случайной величины $\xi[t]$ и стационарного. Если эти множители достаточно малы, то стационарное распределение быстро достигается. В случае $q = p$, когда переходы от к и обратно одинаково вероятны, верны равенства $b = c = \frac{1}{2}$ и стационарное распределение равномерно: $B = \left\{\frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}\right\}$. Доли всех рассматриваемых классов становятся одинаковыми.

Пусть $p = q = 1/5$.

Тогда $u = {}^2/_5, \quad 1 - u = 3/5, \quad 1 - 2u = 1/5, \quad 1 - 3u = \frac{2}{5}$ и стационарное распределение B равномерно. Умеренно малый размер множителей $1-u, 1-2u, 1-3u$ определяет умеренную скорость приближении распределения к равномерному. Эти распределения в моменты $t = 3, \quad 6, \quad 12, \quad 15$ с округлением результатов до тысячных выписаны в таблице 5.

Таблица 5

Гипотетическая динамика изменения социокультурных типов у молодежи возрастной группы 20-24 лет до приближения распределения к равномерному при заданных переходных вероятностях

$$p = q = \frac{1}{5}$$

Время t	Социокультурные типы							
	AAA	AAB	ABA	BAA	ABB	BAB	BBA	BBB
3	0.116	0.132	0.136	0.101	0.15	0.114	0.117	0.134
6	0.123	0.127	0.127	0.12	0.13	0.123	0.123	0.127
12	0.125	0.125	0.125	0.124	0.126	0.125	0.125	0.125
15	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125

Приведем также пример неравномерного стационарного распределения. Пусть $p = \frac{1}{12}$, $q = \frac{1}{6}$.

Тогда $u = \frac{1}{4}, 1 - u = \frac{3}{4}, 1 - 2u = \frac{1}{2}, 1 - 3u = \frac{1}{4}$ и $B = \{0.30, 0.15, 0.15, 0.15, 0.07, 0.07, 0.07, 0.04\}$ при округлении до сотых. Будем считать такие доли рассматриваемых классов нормальными. Умеренно малый размер множителей $1-u, 1-2u, 1-3u$ определяет умеренную скорость стабилизации распределения. Эти распределения в моменты $t = 5, 10, 15, 20$ с округлением результатов до сотых выписаны в таблице 6.

Таблица 6

Гипотетическая динамика изменения социокультурных типов у молодежи возрастной группы 20 – 24 лет до стацонарного распределения при заданных переходных вероятностях $p = \frac{1}{12}$, $q = \frac{1}{6}$

Время t	Социокультурные типы							
	AAA	AAB	ABA	BAA	ABB	BAB	BBA	BBB
5	0.23	0.14	0.16	0.12	0.11	0.08	0.08	0.05
10	0.28	0.15	0.15	0.14	0.08	0.08	0.08	0.04
15	0.29	0.15	0.15	0.15	0.08	0.07	0.07	0.04
20	0.3	0.15	0.15	0.15	0/07	0.07	0.07	0.04

Можно провести аналогичные вычисления и в общей марковской модели с 8 состояниями, не выводя общих формул и не делая упрощающих предположений о переходных вероятностях. Существующие математические программы позволяют производить достаточно точные вычисления и с матрицами больших порядков.

Заключение. Если имеющихся статистических данных достаточно для определения переходных вероятностей, то можно выбрать стохастическую марковскую модель, с нужной адекватностью описывающую динамику исследуемой общественной группы по индивидуальным, социальным и культурным признакам. Марковская модель дает возможность прогнозировать эту динамику на близкую, среднюю и далекую перспективу. Использование аппарата неоднородных марковских цепей и последовательное изменение переходных вероятностей позволяет учитывать изменение условий. Прогнозы по удачно выбранным марковским моделям хорошо согласуются с реальностью.

Приведенные в статье примеры хотя и используют реальные статистические данные, но имеют иллюстративный характер. Отсутствие достаточных статистических данных для определения переходных вероятностей позволяет рассматривать только их различные гипотетические варианты. Если определить желательную по социальным соображениям стабильную структуру рассматриваемой группы населения, то можно выбрать специальные переходные вероятности так, чтобы они обеспечивали постепенный переход к стационарному распределению по выделенным классам за приемлемое время. Переходные вероятности могут давать представление о направлениях усилий общества для стабилизации, рассматриваемой структуры. В частности, речь может идти о воспитании личности, этнического сознания, общей и национальной культуры.

Библиографический список

1. Кемени Дж., Снелл Дж. Конечные цепи Маркова. Москва, Наука, 1970. С. 271.
2. Пытьев Ю.П. Возможность. Элементы теории и применения. Москва, УРСС, 2000. С. 190.
3. Жданок А.И. Основы теории конечных цепей Маркова. Теория и практика. Кызыл, ГУП «Тываполиграф», 2008. С. 168.
4. Савельев Л., Балакин С. Конечные марковские цепи и серии (теория и приложения). Saarbrucken, LAMBERT Academic Publishing, 2012. С. 150.
5. Савельев Л.Я. Элементарная теория вероятностей, 1-2. Новосибирск, НГУ, 2005. С. 158, 190.
6. Толстова Ю.Н. Математико-статистические модели в социологии. Москва, ГУ ВШЭ, 2008. С. 243.
7. Толстова Ю.Н. Измерение в социологии. Москва, 2007. С. 287.
8. Феллер В. Введение в теорию вероятностей и ее приложения, 1-2, Москва, Мир, 1984. С. 528, 752.
9. Загоруйко Н.Г., Савельев Л.Я. Относительная мощность измерительных шкал. Вычислительные системы, вып. 101. Новосибирск, 1984. С. 114-129.
10. Абрамова М.А., Костюк В.Г., Гончарова Г.С. Жизненные планы и адаптация различных социокультурных типов молодежи. Журнал социологии и социальной антропологии. Т. 16, № 1, 2013. С. 78-99.

Bibliograficheskij spisok

1. Kemeni Dzh., Snell Dzh. Konechnye tsepi Markova. Moskva, Nauka, 1970. C. 271.
2. Pytev Yu.P. Vozmozhnost. Elementy teorii i primeneniya. Moskva, URSS, 2000. S. 190.
3. Zhdanok A.I. Osnovy teorii konechnykh tsepej Markova. Teoriya i praktika. Kyzyl, GUP "Tyvapoligraf", 2008. S. 168.
4. Savelev L., Balakin S. Konechnye markovskie tsepi i serii (teoriya i prilozheniya). Saarbrucken, LAMBERT academic publishing, 2012. S. 150.
5. Savelev L.Ya. Elementarnaya teoriya veroyatnostej, 1-2. Novosibirsk, NGU, 2005. S. 158, 190.
6. Tolstova Yu.N. Matematiko-statisticheskie modeli v sotsiologii. Moskva, GU VSHE, 2008. S. 243.
7. Tolstova Yu.N. Izmerenie v sotsiologii. Moskva, 2007. S. 287.
8. feller v. vvedenie v teoriyu veroyatnostej i ee prilozheniya, 1-2, moskva, mir, 1984. s. 528, 752.
9. Zagorujko N.G., Savelev L.Ya. Otnositelnaya moschnost izmeritelnykh shkal. Vychislitelnye sistemy, vyp. 101. Novosibirsk, 1984. S. 114-129.

10. Abramova M.A., Kostyuk V.G., Goncharova G.S. Zhiznennye plany i adaptatsiya razlichnykh sotsiokulturnykh tipov molodezhi. Zhurnal sotsiologii i sotsialnoj antropologii. T. 16, № 1, 2013. S. 78-99.

Савельев Лев Яковлевич – кандидат физико-математических наук, профессор Новосибирского государственного университета, профессор Тувинского государственного университета, НИЛ «Теории вероятностей и ее приложений», E-mail: savelev@math.nsc.ru

Гончарова Галина Савитовна – научный сотрудник Института философии и права СО РАН, E-mail: socis@philosophy.nsc.ru

Savelyev Lev – PhD Candidate of Physics and Mathimatics, professor of Novosibirsk State University, Professor of Tuvan State University, Laboratory "Theory of Probability and its Applications", E-mail: savelev@math.nsc.ru

Goncharova Galina – Researcher, Institute of Philosophy and Law, Russian Academy of Sciences, E-mail: socis@philosophy.nsc.ru.

УДК 519.217.2 +314.3

МАРКОВСКАЯ МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДИНАМИКИ КОЛИЧЕСТВА ДЕТЕЙ В СЕМЬЯХ ТУВЫ

Хурума А.К.
Тувинский государственный университет,
НИЛ «Теории вероятностей и ее приложений», Кызыл

THE MARKOV MATHEMATICAL MODEL THE DYNAMICS OF THE NUMBER OF CHILDREN IN FAMILIES OF TUVA

Khuruma A.K.
Tuvan State University, Laboratory "Theory of Probability and its Applications", Kyzyl

В работе строится и исследуется Марковская математическая модель динамики количества детей в семьях Республики Тыва. Дается полное количественное описание параметров и характеристик построенной цепи Маркова, основанное на данных Тывастата РТ. Доказывается, что построенная цепь Маркова является неразложимой и эргодической, найдено ее стационарное распределение и дана оценка скорости сходимости к нему ежегодных абсолютных распределений вероятностей. Произведено прогнозирование динамики изменения количества детей в семьях на ряд лет и сделано сравнение с уже имеющимися статистическими данными.

Ключевые слова: Конечные цепи Маркова, Марковская модель, эргодичность, стационарное распределение, структура семьи, прогнозирование

This article presents the constructed mathematical model of Markov dynamics of the number of children in Tuvan families. Completed quantitative description of the parameters and characteristics of the constructed Markov chain was given, which based on databases of Tuvan Statistics Center of Republic of Tuva. It is proved that the constructed Markov chain is indecomposable and ergodic. The stationary distribution was found and evaluated the rate of convergence to it annual absolute probability distributions. The forecasting dynamics of