

УДК 303.09
303.733.3

ИРИНА ПРИБЫТКОВА,

доктор экономических наук, профессор, ведущий научный сотрудник отдела социальных структур Института социологии НАН Украины

Статистический анализ и измерение в социологических исследованиях

Аннотация

Рассматриваются методы статистического анализа и измерения, используемые в социологических исследованиях для осуществления первичной систематизации и обобщения данных статистического наблюдения и их вторичной переработки в целях более глубокого понимания сущности и закономерностей развития массовых общественных процессов и структур в конкретных условиях места и времени. Обсуждается, в частности, использование марковских моделей, которые дают вероятностное описание механизма перераспределения индивидов по стратификационным группам при моделировании процессов социальной мобильности. Изложены результаты использования шагового регрессионного анализа в целях прогнозирования территориальной мобильности населения. Продемонстрированы возможности использования графических моделей статистических распределений (кривой Лоренца, кумулятивной кривой) при изучении пространственной структуры перемещений сельского населения с трудовыми целями.

Ключевые слова: *количественный анализ как составная часть теории, методы вторичной обработки и анализа статистической информации, моделирование и прогнозирование процессов пространственной социальной мобильности населения, графические модели статистических распределений*

Количественный анализ как составная часть теории

Систематизация накопленных в процессе исследования фактов, их обобщение, анализ и интерпретация, установление связи между ними и уже описанными фактами являются важным моментом разработки теории. Однако не всякий факт можно считать научным. Только будучи включен в

определенную теоретическую систему, он приобретает качество научности. Научно-теоретическое осмысление конкретных фактов, отражающих характер, тенденции и закономерности развития массовых общественных процессов и явлений, позволяет расшифровать информацию об их сущности и взаимосвязях.

На основе количественного анализа эмпирических данных формируются теоретические гипотезы, позволяющие в самом общем виде объяснить социальные, экономические и демографические явления и процессы. Такое объяснение необходимо для того, чтобы направить исследование на поиск систематических связей между фактами. С этой точки зрения количественный анализ можно интерпретировать как составную часть теории. В том же случае, когда рассматриваются конкретные способы обработки и упорядочения эмпирических данных и проверки гипотез статистическими методами, количественный анализ следует отнести к области собственно статистики.

Э.Кейн отмечает, что главную роль в количественном анализе всегда играют формулирование гипотезы (спецификация) и ее проверка. Значение же статистики состоит главным образом в том, что она может по-новому осветить теоретические построения, с помощью которых предполагается описать и предсказать возможные направления, динамику и интенсивность развития разнообразных структур и процессов [Кейн, 1977: с. 9–10].

По мнению Э.Кейна основным предметом профессионального интереса статистиков являются главным образом *гипотезы, которые могут быть опровергнуты*, иными словами такие, ложность которых можно в принципе доказать. Он утверждает, что только тогда, когда гипотеза рискует быть опровергнута, возможен выбор между ней и альтернативным теоретическим предположением. И приходит к заключению, что для опытного аналитика понятия *опровергаемость* и *проверяемость* — это почти синонимы.

Развитие теоретических представлений в любой области научного знания, по Э.Кейну, происходит в такой последовательности:

- разработка теорий, все более поддающихся проверке с помощью эмпирических данных;
- верификация теоретических предположений путем их сопоставления с эмпирическими данными до того момента, пока эти теории не будут опровергнуты;
- замена отвергнутых теорий новыми.

В результате нечеткие теории приобретают более отточенную форму и участвуют в последующих испытаниях и проверках с помощью эмпирических данных. Противоречащие фактам теории пересматриваются. Другие же, успешно преодолевшие очередной тур испытаний, подвергаются новым — более сложным и трудным [Кейн, 1977: с. 10].

Очень важно помнить о недопустимости приспособления (адаптации) своей теории к тем или иным конкретным фактам. Не менее важной задачей является тщательное разграничение проверки уже существующей теории и анализа наблюдений для формулировки новой теории, которая еще ожидает своей проверки. Э.Кейн утверждает, что различные теории, созданные на одном и том же множестве данных, невозможно проверить с помощью тех же самых фактов, поскольку такая проверка не обеспечивает независимого подтверждения теории. Независимым подтверждением могут служить лишь новые факты или сведения, полученные в результате дальнейшего исследования старых фактов [Кейн, 1977: с. 40].

Количественные закономерности массовых общественных процессов проявляются в форме эмпирических статистических закономерностей. Их выявление и исследование лежит в основе обобщенной количественной характеристики этих процессов.

Статистика как наука

Упоминание о статистике как науке появилось впервые во второй половине XVII века в трудах политических арифметиков. Отцом новой науки К.Маркс считал В.Петти, а его метод характеризовал как нетрадиционный. "...Вместо набора целого ряда слов в сравнительной и превосходной степени и спекулятивных аргументов, он (В. Петти) решил говорить посредством ... чисел, весов и мер, пользоваться исключительно аргументами, взятыми из чувственного опыта, и рассматривать только такие причины... которые имеют взаимное основание в природе", — писал классик [Маркс, Энгельс, т. 13: с. 39]. В.Петти показал научную силу и убедительность нового метода и необходимость его применения в противовес доказательствам, основанным лишь на умозрительных аргументах. Дальнейшее развитие статистическая теория получила в работах представителей различных научных школ, а сама статистика и как наука, и как отрасль практической деятельности достигла значительных успехов в технике сбора, обработки, анализа и обобщения статистической информации.

Как наука статистика изучает количественную сторону массовых общественных явлений в неразрывной связи с их качественной стороной, исследует количественное выражение закономерностей общественного развития в конкретных условиях места и времени [Рязов, 1979: с. 8]. Количественные характеристики общественных явлений статистика представляет в цифрах, которые изменяются во времени и пространстве. Близкое к этому определению толкование смысла статистики как отрасли знания дает и распространенный американский справочник по классификации работ в административных учреждениях США: статистика есть *наука о сборе, классификации и количественной оценке фактов в качестве основы для выводов*.

Важное значение в статистике имеет понятие *совокупности*, представляющей собой массу отдельных единиц, объединенных общей качественной основой, однако отличных друг от друга по ряду признаков. Сами же статистические цифры бывают двоякого рода: одни из них составляют обобщающую характеристику объемов совокупности (например, общей численности населения, размеров трудовых ресурсов либо жилищного фонда в стране); другие же дают обобщающие характеристики совокупностей по ряду признаков, которые имеют различное количественное или качественное выражение. Такие признаки называются варьирующими (изменяющимися).

Обобщающие характеристики массовых общественных процессов и явлений статистика использует для выявления закономерностей, проявляющихся в определенном порядке расположения, соотношения или изменения статистических цифр. Это могут быть *закономерности развития (динамики) различных процессов или явлений; закономерности изменения их структуры, распределения единиц внутри совокупности и, наконец, связанного изменения разных варьирующих признаков внутри совокупностей* (например, распределение занятых определенных профессий и квалификации по величине стажа практической работы).

Закономерности общественной жизни всегда отражают действие сложных причин, соотношение которых и совокупный результат изменяются с развитием общества и, следовательно, проявляются *только в условиях конкретной места и времени*. Важнейшая задача статистических исследований заключается в выявлении и измерении закономерностей массовых процессов и явлений с помощью методов, разрабатываемых статистикой как наукой [Рязузов, 1979: с. 14]. Специфические приемы, с помощью которых статистика изучает свой предмет, образуют статистическую методологию. В каждом статистическом исследовании можно выделить три последовательные стадии: *статистическое наблюдение* (сбор первичного статистического материала); *сведение результатов наблюдения в определенные совокупности; анализ полученных сводных материалов*. На всех этих трех стадиях в качестве инструмента статистической методологии используются специальные методы.

Методы статистического анализа дают возможность осуществить первичную систематизацию и обобщение цифровых данных, описывающих исследуемые совокупности. Более глубокое понимание сущности изучаемых процессов или явлений может быть достигнуто лишь в результате *вторичной обработки данных статистического наблюдения*, смысл которой заключается в дальнейшем обобщении уже полученных результатов и их научной интерпретации.

Методы вторичной обработки и анализа данных сводятся, в конечном счете, к математико-статистическому моделированию, которое может быть осуществлено различными средствами. Графики, используемые в математико-статистическом анализе, представляют собой одну из форм математического моделирования, вполне равноправную с аналитической (алгебраической) моделью и имеющую ряд преимуществ перед ней. В частности, в моделях, описывающих статистические совокупности со сложной структурой распределения значений варьирующего признака, графическая форма записи предпочтительнее. Основные функции графических моделей заключаются в отражении объема и структуры статистических совокупностей, их изменения во времени и распределения в пространстве, в выявлении закономерностей, присущих этим совокупностям.

К примеру, анализ структурной диаграммы, известной в практике демографической статистики как *возрастная пирамида*, позволяет составить представление о возрастно-половой структуре населения и выявить отклонения возрастного состава населения от нормального. Кроме того, возрастная пирамида содержит возможность демографического прогноза на ближайшее будущее.

Используя графическое изображение динамических рядов, можно составить наглядное представление о характере развития того или иного процесса, явления либо структуры. Для анализа форм статистического распределения может быть применен ряд моделей: гауссова кривая нормального распределения, модели распределения Пуассона, Пирсона, Фишера, Стьюдента, кривая Лоренца и др.

Задачи графической модели, известной под названием кривой Лоренца, заключаются в определении и измерении различий в степени концентрации больших масс в определенном количестве отдельных крупных единиц, в установлении удельного веса или относительного значения этих крупных единиц во всей генеральной совокупности. В качестве инструмента анализа статистических распределений может быть использована и графическая

модель нарастающих или накопленных частот (кумулятивная кривая), широко применяемая в статистической практике.

Математическое моделирование входит в число общих научных методов, применяемых в экономико-социологических и социально-демографических исследованиях. На определенном этапе построения научной теории возникает необходимость в ее формализации и обеспечении строгой логической основы для выдвинутых гипотез. Математика обеспечивает здесь тот логический фундамент, на котором можно строить теорию [Бунге, 1967: с. 58]. По мнению Д. Харвея, модель можно рассматривать как формализованное изложение теории с помощью средств формальной логики, теории множеств и математической статистики, применение которых позволяет выявить и устранить внутреннюю несогласованность теории [Харвей, 1976: с. 240].

Наиболее распространенными моделями, применяемыми в исследованиях развития пространственных систем и пригодными для анализа миграционных потоков, являются гравитационные, а также модели потенциалов и пространственного взаимодействия. Наряду с гравитационными моделями широко используются предназначенные для практических расчетов регрессионные модели, хотя они нередко дают невысокие значения коэффициентов множественной регрессии. И потому целесообразно дополнять регрессионный анализ качественными исследованиями и тщательной интерпретацией полученных результатов [Прибыткова, 2009: с. 87].

Обнадеживающие результаты приносит комплексное использование статистических методов, в частности сочетание вариационного, дисперсного, корреляционного и регрессионного анализов.

Поскольку регрессионный анализ дает поддающиеся интерпретации результаты лишь применительно к достаточно однородным совокупностям, рекомендуется разбить исходные данные на однородные группы и лишь затем применить регрессионный анализ к каждой из них по отдельности [Айвзян, 1973: с. 968–983]. Однако при выборе одного из многих возможных методов регрессионного анализа необходимо учитывать свойства полученных в результате классификации групп. Пакеты прикладных статистических программ обычно включают ряд методов классификации и регрессии.

Моделирование процессов социальной мобильности

Объектом моделирования процессов социальной мобильности являются различные виды движения множества индивидов и их групп во времени и пространстве: социальные перемещения, различные виды движения трудовых ресурсов (территориального, межотраслевого, межпрофессионального). А первоначальная задача состоит в описании их поведения, которому присущ ряд характерных особенностей.

- Социальное поведение детерминировано множеством личных мотивов, имеющих часто противоречивый характер и определяемых в конечном счете конкретными объективными обстоятельствами.
- Социальное поведение индивидов характеризуется нестабильностью, поскольку определяющие его факторы пребывают в состоянии непрерывного изменения.
- Социальное поведение является четко выраженным вероятностным процессом.

- Социальное поведение отличается существенной неоднородностью совокупности лиц, включенных в процессы социальной мобильности [Движение рабочих кадров, 1974: с. 40].

При моделировании процессов социальной мобильности совокупность включенных в них индивидов разделена на k стратификационных групп, каждая из которых объединяет индивидов, имеющих однородные характеристики. Предполагается, что индивиды могут переходить из одной стратификационной группы в другую и k -мерный вектор $x(t)$ задает численности индивидов в группах в момент времени t . Описание поведения вектора $x(t)$ во времени — задача первостепенной важности, рассмотрение которой при предположении о случайном характере индивидуальных переходов неизбежно приводит к изучению векторного случайного процесса [$x(t), t \geq 0$].

При моделировании социальной мобильности наиболее часто используются модели марковского типа [Stewman, 1976: p. 201–245; Романов, Терехов, 1980: с. 272–291], позволяющие учитывать возможность каждого индивида выбирать с некоторой вероятностью один из нескольких допустимых способов поведения. Это дает возможность получить естественным образом вероятностное описание механизма перераспределения индивидов по стратификационным группам. Наиболее широкое применение находят различные модификации простейшей марковской модели с дискретным временем (например, за год) и постоянной во времени матрицей переходных вероятностей. В их числе следует назвать модификацию модели, учитывающей неоднородность индивидуальной активности, а также модификацию модели, учитывающей изменчивость поведения индивидов во времени [Орлов, 1983: с. 216].

Хороших результатов использования простейшей марковской модели можно ожидать при моделировании процессов, в которых показатели мобильности стационарны, а также при условии, что индивиды из одной стратификационной группы ведут себя одинаково, то есть группа однородна. Такими процессами являются, к примеру, процессы территориальной мобильности (переезды на постоянное место жительства), социальные перемещения индивидов, процессы межотраслевой и межпрофессиональной мобильности.

Для построения интервального прогноза распределения индивидов по стратификационным группам необходимы данные о передвижении индивидов в виде матрицы. Получение такой информации представляет собой самостоятельную проблему и связано, как правило, с проведением трудоемких исследований, как, например, при построении межпрофессионального баланса трудовых ресурсов.

В заключение замечу, что вероятностный подход открывает широкие возможности для объяснения пространственных взаимодействий (связей, потоков) и изучения различных проявлений пространственной деятельности человека.

Использование прикладного регрессионного анализа в прогнозировании процессов территориальной мобильности населения

Развитие процессов территориальной мобильности населения происходит под влиянием множества разнообразных факторов. Некоторые из них являются результатом целенаправленной деятельности человека, другие

же — не зависят от нее. Поэтому при прогнозировании размеров, направлений, интенсивности и структуры потоков территориальной мобильности населения совокупность анализируемых факторов можно разделить на две основные группы: регулируемые и нерегулируемые. Между ними существуют сложные взаимосвязи, они действуют одновременно, и их влияние на размеры и интенсивность межпоселенных и межстрановых потоков индивидов носит комплексный характер. Анализ, оценка и прогнозирование объемов, направлений и структуры этих потоков возможны на основе количественного изучения степени влияния всех факторов — как регулируемых, так и нерегулируемых — на исследуемый показатель.

Многофакторный корреляционно-регрессионный анализ позволяет оценить меру влияния на исследуемый результирующий показатель каждого из включенных в анализ факторов при фиксированном положении (на среднем уровне) всех остальных, а также при любых возможных сочетаниях факторов найти с определенной степенью точности теоретическое значение этого показателя. При этом необходимым условием является отсутствие между ними функциональной связи. На основе предварительного теоретического анализа формулируется гипотеза о причинно-следственных связях.

Лишь после того, как такие предположения выдвинуты, многофакторный корреляционно-регрессионный анализ может либо подтвердить, либо опровергнуть их и дать количественную оценку влияния различных факторов.

В ходе корреляционно-регрессионного анализа осуществляются построение и анализ статистической модели в виде уравнения регрессии, приближенно выражающей зависимость исследуемого показателя от влияющих на него факторов. Математическая задача формулируется следующим образом: из множества функций требуется выбрать аналитическое выражение, которое наилучшим образом будет отражать реально существующие связи между изучаемым показателем и факторами, то есть найти функцию $Y = f(x_1 + x_2, \dots, x_n)$, где в нашем случае Y — территориальная мобильность населения, а x_1, x_2, \dots, x_n — факторы, оказывающие влияние на ее формирование.

Задача состоит в том, чтобы раскрыть характер и степень влияния факторов на показатели территориальной мобильности населения. Если задача успешно решена, мы получаем инструмент для прогнозирования, а в ряде случаев и для управления исследуемым процессом. Выбор типа функции осуществляется на основе либо теоретических знаний об изучаемом процессе, либо предыдущего опыта аналогичных исследований или же определяется эмпирически, путем перебора функций разных типов. Так как в большинстве случаев любая функция многих переменных посредством их замены или логарифмирования может быть сведена к линейному виду, уравнение множественной регрессии строится в линейной форме $Y = f(a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n)$.

Каждый коэффициент этого уравнения показывает степень влияния того или иного фактора на анализируемый показатель при фиксированном положении остальных факторов. С изменением каждого фактора на единицу показатель изменяется на величину соответствующего коэффициента регрессии. Свободный член уравнения не несет никакой социально-экономической нагрузки. Уравнение множественной регрессии учитывает сред-

нее влияние на анализируемый показатель не всех, а лишь наиболее существенных факторов. Поэтому даже при наиболее удачно выбранной функции возникает разброс фактических значений Y вокруг расчетных. При построении уравнения возникает проблема отбора переменных, включенных в регрессионный анализ.

Существует несколько статистических методов отбора переменных: метод всех возможных регрессий, метод исключения, метод включения, шаговый регрессионный анализ, две вариации четырех предыдущих методов, ступенчатый регрессионный анализ. Хорошие результаты приносит шаговый регрессионный анализ, представляющий собой улучшенную модификацию метода включения.

Метод включения позволяет избежать обработки большего числа переменных, чем это необходимо. В анализ включаются переменные по очереди до тех пор, пока регрессионное уравнение не станет удовлетворительным. Порядок их включения определяется с помощью частного коэффициента корреляции, используемого в качестве меры важности еще не включенных в уравнение переменных. Однако этот метод не позволяет исследовать влияние очередной переменной на вклад другой переменной, введенной в уравнение на предшествующей стадии. Этот недостаток преодолен в шаговом регрессионном анализе, который предусматривает дополнительное исследование и оценку на каждой стадии меры важности переменных, включенных в модель на предшествующих стадиях.

Для оценки значимости вклада каждый из переменных при очередном шаге осуществляется дисперсионный анализ, в рамках которого вычисляется частный F -критерий (критерий Фишера). Дисперсионный анализ позволяет математически строго проверить существенность анализируемых связей. Если переменная, вносящая в модель существенный вклад на ранних стадиях ее построения, вводится в регрессионное уравнение после других переменных, она может оказаться излишней из-за значительной корреляции с ними. Таким образом, F -критерий позволяет оценить относительный эффект каждой переменной после других и служит для решения вопроса о добавлении или исключении членов модели.

При отборе переменных для включения в регрессионное уравнение следует установить, является ли та или иная переменная только измеряемой, или же она кроме того еще и управляема. Если модель предназначена для целей управления, в нее должны быть включены управляемые переменные даже в том случае, когда их вклад в уравнение менее значим.

Однако сами коэффициенты регрессии не позволяют сказать, какие из них оказывают наибольшее влияние на размеры потоков территориальной мобильности (поскольку они измерены различными единицами), а также в развитии каких факторов заложены наиболее значительные резервы регулирования этого процесса в заданном направлении. Для этого вычисляются частные коэффициенты эластичности (E_i) и так называемые бета-коэффициенты (β_i).

С помощью частных коэффициентов эластичности устраняются различия в единицах измерения факторов. Они показывают, на сколько процентов в среднем изменяется анализируемый показатель с изменением на 1% каждого фактора при фиксированном положении других факторов. β -коэффициент показывает, на какую часть среднего квадратического отклонения

изменяется исследуемый показатель (зависимая переменная) с изменением соответствующего фактора на величину его среднеквадратического отклонения. Таким образом, на основании частных коэффициентов эластичности и β -коэффициентов можно судить о вкладе, который вносят включенные в уравнение факторы в динамику показателей территориальной мобильности населения.

Многу была предпринята попытка определить наиболее существенные факторы и силу их влияния на формирование потоков территориальной трудовой мобильности сельского населения в Украине. В качестве зависимой переменной Y принята величина контингента лиц, совершающих ежедневные трудовые поездки из сельской местности к месту приложения труда в городских поселениях и обратно. В анализ включены все существенные факторы, которые определяют условия труда, быта и отдыха, а также уровень благосостояния и качества жизни сельского населения страны: уровень трудовых доходов, размеры денежных сбережений и накопления предметов длительного пользования, возможности получения жителями села образования и профессии, доступность и качество медицинской помощи, развитие дорожной сети, обеспеченность учреждениями культурно-бытового обслуживания и торговли и ряд других факторов. Для расчета коэффициентов корреляции и регрессии были использованы динамические ряды включенных в анализ показателей длиной в 15 лет.

Выбор уравнения, которое бы наилучшим образом отражало реально существующие связи между размерами потоков ежедневной трудовой миграции в города из сельской местности и влияющими на этот процесс факторами, был осуществлен с помощью шагового регрессионного анализа. Уравнение приобрело следующий вид:

$$Y = 1779,818 + 1,346X_1 + 0,028X_2 - 255,158X_3 + 9,510X_4 \quad (1),$$

где X_1 — величина розничного товарооборота в расчете на душу населения в сельской местности;

X_2 — строительство жилых домов, тыс. м² общей/полезной площади;

X_3 — число клубных учреждений в сельской местности в расчете на 10000 жителей;

X_4 — густота автомобильных дорог с твердым покрытием, км на 1000 км² территории.

Основная идея шагового регрессионного анализа состоит в нахождении регрессии с несколькими переменными в виде серий линейных регрессионных зависимостей и в перестройке таблицы шаг за шагом до конца. Кроме того, на каждом этапе рассчитывается совокупный коэффициент множественной корреляции R и совокупный коэффициент множественной детерминации R^2 .

Коэффициент множественной корреляции R является основным показателем линейной корреляционной связи. Он измеряет одновременное влияние переменных-факторов на зависимую переменную и является индикатором тесноты связи между ними. Чем меньше наблюдаемые значения исследуемого показателя отклоняются от линии множественной регрессии, тем корреляционная связь является более интенсивной, а следовательно, величина R ближе к единице. Совокупный коэффициент множественной детерминации показывает, какая доля вариации изучаемого показателя объясняется влиянием факторов, включенных в уравнение регрессии. Чем

ближе R^2 к единице, тем в большей мере вариация зависимой переменной характеризуется влиянием отобранных факторов.

В нашем случае R составляет 0,9979, а R^2 равен 0,9959, то есть изменение территориальной трудовой мобильности сельского населения Украины на 99,6% обуславливается 4 факторами, включенными в регрессионное уравнение.

Существенное значение для развития процесса территориальной мобильности сельских жителей имеет их осведомленность об условиях и характере труда в местах их предполагаемого трудоустройства. Перемещению в пространстве как факту всегда предшествует процесс принятия решения об изменении места и сферы приложения труда, характера трудовой деятельности и принадлежности к той или иной социальной группе населения. Включенные в анализ факторы оказывают влияние прежде всего на принятие решения. Само же перемещение индивидов имеет по отношению к решению вторичный характер, подчинено ему и служит целям его реализации.

Полученное уравнение регрессии (1) может быть использовано для целей прогнозирования. Для расчета перспективных значений Y достаточно подставить в него прогнозные значения X_1, X_2, X_3 и X_4 .

Графические модели как инструмент анализа пространственной неравномерности распределения трудовых потоков сельских жителей в города и их интенсивности в Украине

Территориальные перемещения сельского населения в города с трудовыми целями имеют повсеместное распространение в пределах Украины. Однако уровни их развития в отдельных регионах страны отличаются значительной неравномерностью. Для выявления степени равномерности распределения между сельскими административными районами масс индивидов, образующих трудовые потоки, я использовала графическую модель, известную под названием *кривой Лоренца* (рис. 1). График концентрации (неравномерности) распределения по модели Лоренца имеет вид квадрата, по сторонам которого от нулевой точки идут две одинаковые процентные шкалы от 0 до 100%. По горизонтальной шкале даны нарастающие проценты числа сельских административных районов, по вертикальной — нарастающие проценты контингентов живущих в этих районах индивидов, занятых в городских населенных пунктах. Процентные шкалы служат координатами, по которым в поле графика точками отмечены рассчитанные в специальной таблице показатели по одноименным группам (табл. 1). Соединенные между собой точки образуют плавные вогнутые кривые Лоренца.

Проведенная в поле графика из нулевой точки диагональная линия представляет собой теоретический предел равномерности распределения, недостижимый практически. Степень отклонения кривых Лоренца от диагонали отражает концентрацию или степень неравномерности распределения трудовых мигрантов из числа сельских жителей между сельскими районами Украины.

Кривые Лоренца, совмещенные в поле графика, указывают, что в период 1968–1981 годов неравномерность распределения работающих ежедневно в городе сельских жителей между сельскими административными районами Украины уменьшилась. Так, в 1968-м почти половина всех работающих в

городе жителей села (48%) была сосредоточена лишь в 10% общей совокупности этих районов, а на 30% районов приходилось 80% трудовых мигрантов из числа сельских жителей. К 1974 году в 10% районов проживали уже 41%, а в 30% районов — 75% общей совокупности сельских жителей, работающих в городе. К 1981-му в 10% районов концентрация трудовых мигрантов, занятых в городе, уменьшилась до 38%, а в 30% районов — до 72%. И если в 1968 году в половине сельских административных районов проживали всего 8% работающих в городских поселениях трудовых мигрантов, то к 1981-му их удельный вес возрос до 13%. Таким образом, в период 1968–1981 годов в пределах Украины наметилась тенденция к некоторой деконцентрации потоков трудовой маятниковой миграции сельского населения, направленных в города.

Таблица 1

Трудовая маятниковая миграция сельского населения в города Украины на 01.01.1981 года (группировка для составления графика концентрации трудовых мигрантов по модели Лоренца)

Группы по размерам кон- тингентов маятниковых мигрантов, чел.	Число районов в группе	Средняя величина интер- вала, чел.	Общее количество маятни- ковых мигрантов, чел.	Нарастающие итоги			
				Абсолютные данные		Проценты	
				Число районов	Общее количест- во маятниковых мигрантов, чел.	Число районов	Общее количест- во маятниковых мигрантов, чел.
До 100	24	50	1200	24	1200	5,3	0,1
101–200	21	150	3150	45	4350	10,0	0,3
201–400	33	300	9900	78	14250	17,2	0,9
401–600	33	500	16500	111	30750	24,6	2,0
601–800	29	700	20300	140	51050	31,0	3,2
801–1000	16	900	14400	156	65450	34,5	4,2
1001–2000	104	1500	156000	260	221450	57,5	14,1
2001–3000	48	2500	120000	308	341450	68,1	21,7
3001–4000	33	3500	115500	341	456950	75,4	29,0
4001–5000	22	4500	99000	363	555950	80,3	35,3
5001–6000	11	5500	60500	374	616450	82,7	39,2
6001–8000	23	7000	161000	397	777450	87,8	49,4
8001–10000	13	9000	117000	410	894450	90,7	56,8
10001–15000	22	12500	275000	432	1169450	95,6	74,3
15001–20000	14	17500	245000	446	1414450	98,7	89,8
20001–25000	2	22500	45000	448	1452450	99,1	92,7
Свыше 25000	4	28750	115000	452	1574450	100,0	100,0

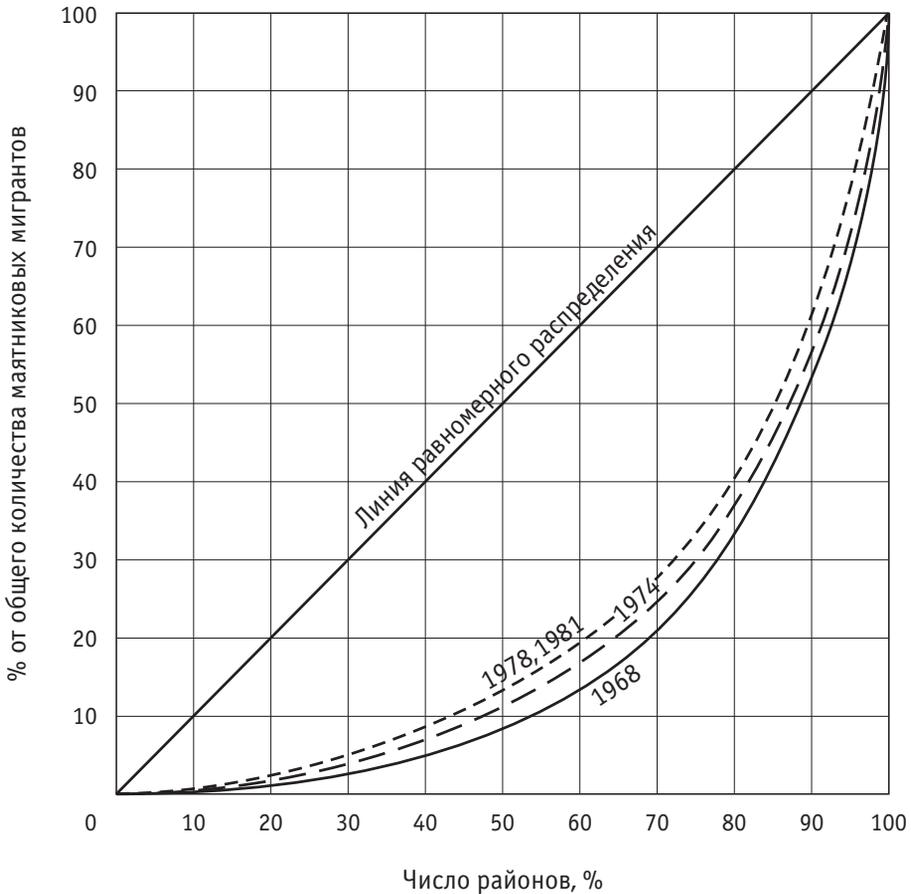


Рис. 1. График концентрации (неравномерности) распределения по модели Лоренца. Распределение общего количества трудовых маятниковых мигрантов между сельскими административными районами Украины в 1968, 1974, 1978 и 1981 годах

Анализ распределения сельских административных районов Украины по уровню интенсивности трудовой маятниковой миграции сельского населения (рис. 2) был осуществлен с помощью графической модели накопленных частот (кумулятивной кривой). Кумулятивный график распределения построен в виде прямоугольника, стороны которого представляют градуированные процентные шкалы: по вертикальной шкале расположены удельные веса трудовых маятниковых мигрантов из числа сельских жителей в общей численности сельского трудоспособного населения административного района; по горизонтальной — нарастающие проценты числа районов. Кумулятивные кривые рассчитаны по принципу «не более чем», их расчетная схема представлена в таблице 2.

Анализ кумулятивных кривых, построенных по данным 1968, 1974, 1978 и 1981 годов, приводит нас к выводу об уменьшении в этот период неравномерности распределения сельских административных районов Украины по уровню интенсивности ежедневных трудовых потоков сельских жителей в города.

В 1968 году подавляющая часть сельских административных районов характеризовались низкой интенсивностью трудовой маятниковой миграции. В 80% районов удельный вес работающих в городах сельских жителей не превышал 10% (в том числе в 20% районов их участие в этом процессе было отмечено показателями менее 1%). В 90% общей совокупности районов интенсивность ежедневных трудовых поездок сельских жителей в города находилась в пределах 20%, и только 10% районов характеризовались показателями выше 20% (от 20% до 69%).

В 1974 году число сельских районов, отмеченных показателями интенсивности ежедневных трудовых потоков в города выше 20%, достигло 16%; сократилось число районов с минимальным уровнем интенсивности ежедневных трудовых поездок сельских жителей в города (менее 1%): их насчитывалось всего 10,3%. Для 80% районов показатели интенсивности этого процесса выросли до 17%, а для 9% районов достигли уровня 27%.

В 1981 году интенсивность трудовых поездок сельских жителей в города в пределах 1% была отмечена только в 7,7% районов. Доля районов с уровнем интенсивности этого процесса выше 20% увеличилась до 22,4%; показатели интенсивности более 30% наблюдались уже в 11,2% районов.

Таким образом, уже к 1981 году произошло сокращение числа районов с низким уровнем интенсивности территориальной мобильности сельского населения в направлении города с трудовыми целями при одновременном увеличении доли районов с наиболее высокими показателями развития этого процесса.

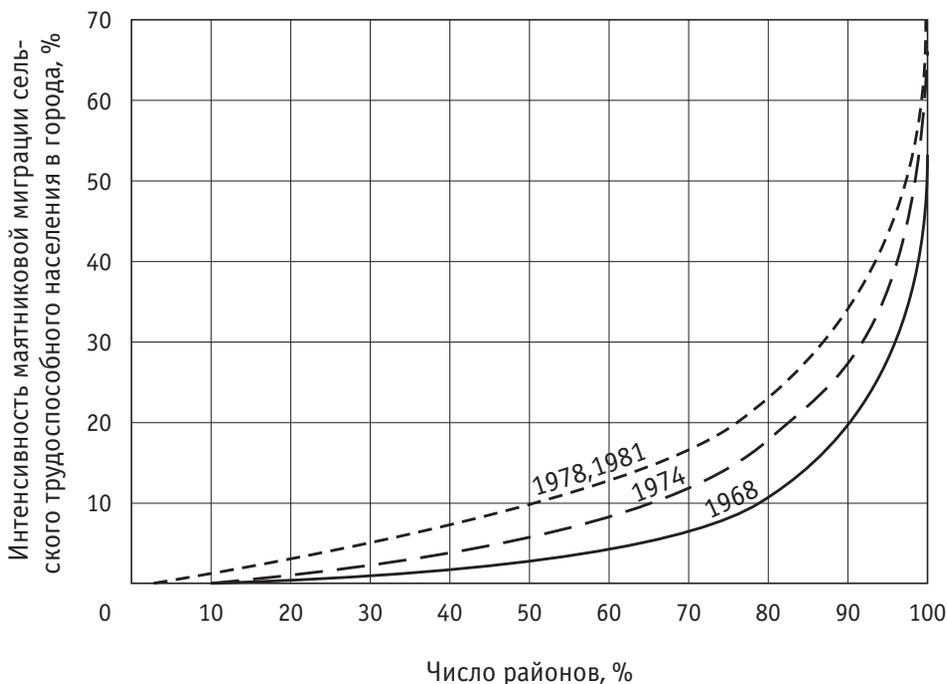


Рис. 2. Кумулятивный график распределения сельских административных районов Украины по уровню интенсивности трудовой маятниковой миграции сельского трудоспособного населения в 1968, 1974, 1978 и 1981 годах

Таблица 2

Распределение сельских административных районов Украины по уровню интенсивности трудовой маятниковой миграции населения на 01.01.1981 года (группировка для составления кумулятивного графика распределения)

Группы по уровню интенсивности трудовой маятниковой миграции сельского населения, в процентах	Число районов в группе	Кумулята по числу районов ^а	
		В абсолютных значениях	В процентах
1,0	35	35	7,7
1,1–2,5	32	67	15,4
2,6–5,0	64	131	28,8
5,1–7,5	60	191	42,0
7,6–10,0	49	240	52,7
10,1–12,5	45	285	62,6
12,6–15,0	29	314	69,0
15,1–17,5	22	336	73,8
17,6–20,0	17	353	77,6
20,1–25,0	22	375	82,4
25,1–30,0	29	404	88,8
30,1–35,0	10	414	91,0
35,1–40,0	9	423	93,0
40,1–45,0	11	434	95,4
45,1–50,0	4	438	96,3
50,1–60,0	15	453	99,6
60,1–70,0	1	454	99,8
70,1–80,0	1	455	100,0

^а Кумулята составлена по принципу “не более чем”. Например, в 35 (или 7,7%) сельских административных районах интенсивность трудовой миграции сельского населения не более чем 1,0%; в 67 (15,4%) — не более 2,5% и т. д.

Заключение

Используя статистические модели, социологи получают инструмент анализа, с помощью которого можно понять многое, но никогда не удастся узнать все. Как правило, это и не нужно. Знать надо только то, что действительно важно. Однако, с другой стороны, никогда нельзя быть уверенным, что в модели учтены все факторы, которые особенно важны при формировании теоретических гипотез или при решении прикладных задач. Статистика как наука пытается получить ответы на вопросы, обладая материалом, который в принципе недостаточен. Поэтому она устанавливает правила, позволяющие судить о степени достаточности или недостаточности того или иного конкретного вывода. Центральной проблемой в освоении статистики

является теория статистического вывода, а понятие вероятности можно назвать подлинной основой статистики.

Источники

Айвазян С.А. Многомерный статистический анализ в социально-экономических исследованиях / Айвазян С.А. — М., 1973.

Бунге В. Теоретическая география / Бунге В. — М., 1967.

Движение рабочих кадров на промышленных предприятиях: теоретические и методические вопросы анализа текучести. — М., 1974.

Кейн Э. Экономическая статистика и эконометрия. Введение в количественный экономический анализ / Кейн Э. — М., 1977.

Маркс Ф. Сочинения / Ф. Маркс, Ф. Энгельс. — 2-е изд. — Т. 13.

Орлов В.И. О прогнозировании процессов социальной мобильности / В.И. Орлов // Прикладная статистика. — М., 1983.

Прибыткова И.М. Пространственная самоорганизация населения: теоретико-методологические предпосылки исследования / И.М. Прибыткова // Социология: теория, методы, маркетинг. — 2009. — № 4.

Романов А.К. Математические модели процессов мобильности : обзор / А.К. Романов, А.И. Терехов. — М., 1980.

Ряузов Н.А. Общая теория статистики / Ряузов Н.А. — М., 1979.

Харвей Д. Социальные процессы и пространственная форма. Анализ концептуальных проблем городского планирования / Д. Харвей // Новые идеи в географии ; вып. 2. — М., 1976.

Stewman S. Markov models of occupational mobility: theoretical development and empirical support / Stewman S. — Pt. I, II. 1976.