

НЕКОТОРЫЕ ВОПРОСЫ КОЛИЧЕСТВЕННОГО ИЗУЧЕНИЯ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

Ю. Н. ГАВРИЛЕЦ

(Москва)

Экономико-математические методы, математическая экономика, теория оптимального планирования — эти понятия сейчас не нуждаются в особом разъяснении, они стали довольно привычными, характеризуют целое направление в науке об общественной жизни, а точнее, об «экономической стороне» общественной жизни. Слова «социологические исследования» или «математическая социология» только начинают завоевывать права гражданства в нашей науке, но интерес к этому направлению, а также отдельные успехи дают основания не сомневаться в таких правах.

В последние годы советскими социологами были проведены многочисленные и интересные практические и теоретические исследования жизни нашего общества, а августовское (1967 г.) постановление ЦК КПСС подчеркнуло все возрастающую роль общественных наук, и в частности социологии, в деле строительства коммунизма. Развитие марксистской социологии тесно связано с дальнейшим развитием исторического материализма и научного коммунизма, влияние которых на жизнь общества огромно. Реальные основы социологической науки были созданы Карлом Марксом еще в прошлом веке, «хотя он сам никогда не употреблял слово „социология“» [1]. Анализируя вклад Маркса в социологическую науку, В. И. Ленин использовал и термин «социология», вложив в него новое научное диалектико-материалистическое содержание: «Маркс положил конец воззрению на общество, как на механический агрегат индивидов..., возникающий и изменяющийся случайно, и впервые поставил социологию на научную почву, установив понятие общественно-экономической формации, как совокупности данных производственных отношений, установив, что развитие таких формаций есть естественно-исторический процесс» [2].

В настоящей статье делается попытка обрисовать в общих чертах одну из возможных точек зрения на взаимосвязь экономической науки и конкретных социальных исследований, остановиться на некоторых понятиях и подходах к изучению социальных явлений и, кроме того, рассмотреть примеры «социально-экономических» задач и подходы к их решению.

1. ЭКОНОМИКА И ЕЕ СРЕДА

Теория оптимального планирования или, как ее стали называть, теория оптимального функционирования изучает функционирование экономической системы и способы сознательного воздействия на экономику. Однако экономическая система (т. е. совокупность процессов производства, распределения и потребления материальных благ и некоторых услуг) не изолирована от других процессов в обществе, так что изменение течения этих процессов оказывает существенное влияние на саму экономику. Конечно, границы между экономикой и неэкономикой определяются довольно условно (это зачастую зависит от удобства решения практических и тео-

ретических задач), но если таковые вообще не будут проведены, то появляется возможность неверного объяснения многих явлений, а практические рекомендации могут стать бессмысленными или даже вредными.

Такие социальные явления, как социальная и классовая структура, уровень научно-технических знаний общества, юридические нормы взаимодействия производств между собой и с человеком, культурный уровень трудящихся, система здравоохранения страны, отношение людей к труду, удовлетворенность своим трудом, коллективом, зарплатой, семейное положение, субъективные оценки человеком материальных и духовных благ и т. д. и т. п., строго говоря, не являются экономическими. Они характеризуют состояние «среды экономической системы». Но без этих явлений не бывает реальной экономической системы. Поэтому правильно понять законы протекания экономической жизни можно лишь тогда, когда будет известно, какова среда экономической системы, как среда влияет на экономику, как вообще происходит взаимодействие экономической системы со средой, что единственно и является процессом функционирования экономики.

Общество в целом представляет собой совокупность взаимодействующих с природой и между собой социальных, экономических, политических, культурно-бытовых и т. п. процессов. Тогда общество в целом будет надсистемой экономической системы, состояния которой зависят от одной группы общественных явлений (их можно называть пограничными процессами входа экономической системы) и определяют протекание другой группы общественных явлений (называемых тогда пограничными процессами выхода экономической системы).

К объектам собственно социологических исследований относятся такие «элементы входа» экономической системы (ЭС), как трудовая деятельность человека в сфере экономики и явления, определяющие эту деятельность. Из «элементов выхода» ЭС к области социологических исследований относятся такие явления, как уровень удовлетворения потребностей населения, влияние потребления на внеэкономическое поведение человека, явления, определяющие спрос, и т. д. Ясно также, что характер трудовой деятельности человека, его образ жизни, семья, поведение в обществе непосредственно влияют на трудовую деятельность. Весь этот круг процессов и явлений непосредственно попадает в сферу конкретных социологических исследований, посвященных изучению общественной жизни человека.

Не учитывая в достаточной полноте роль человека как создателя потребительских благ и как главную цель всего функционирования производства, оптимальное планирование не способно решить до конца ни стоящих перед ним теоретических проблем, ни тем более вопросов практического внедрения достижений науки в реальное планирование.

Большинство макроэкономических моделей народного хозяйства рассматривают человека в качестве трудовых ресурсов — в виде некоторого числа ограничений — и как потребителя — в виде групп семей-потребителей, которые характеризуются либо функциями спроса, либо целевыми функциями потребления. Не вдаваясь в подробности, сразу можно указать на принципиальную ограниченность такого общепринятого подхода. Во-первых, необходимые количества труда (определяемые из решения экстремальной задачи) не могут приниматься в расчет, как, например, требуемые количества сырья и оборудования, поскольку «переливание» труда из одних производств в другие связано с желанием человека, с дополнительными затратами на это перераспределение и т. п. Учет же таких факторов сам по себе очень непрост, для этого нужна специальная социолого-экономическая модель. Во-вторых, учет обратного влияния уровня жизни на производительность также нетривиален в силу того, что субъектом потребления является не индивид-производитель, а семья. По этой же причине труду

может соответствовать лишь зарплата, а не потребление индивида. В-третьих, система действительно оптимального планирования должна рассматривать труд не только как необходимость (которую надо минимизировать), а и как благо (социологи только подступают к этой проблеме через коэффициенты привлекательности профессий и т. п.). Это, в частности, означает, что оптимальный план должен предполагать такое распределение трудовых ресурсов, которое с точки зрения максимизации общей производительности может быть и не наилучшим, поскольку труд используется не там, где наибольшая производительность, а там, где людям «приятнее работать». На самом деле, коллизия только кажущаяся, и исчезает, если понимать, что потребности людей в широком смысле слова включают в себя и потребности трудиться там, где этого хочется *. В-четвертых, с проблемой затрат труда тесно связана проблема использования вне рабочего времени, досуга и вообще социального поведения. Человеку не только хочется увеличить время, свободное от обязательного труда, но улучшить структуру затрат этого времени, что связано с проблемой услуг, культуры и духовного роста, с внеэкономической деятельностью. Учет этих обстоятельств в целевой функции или ограничениях основной модели представляет собой отдельную проблему, не решаемую введением в модель дополнительных формул.

Краткое перечисление особенностей проблемы оптимального использования трудовых ресурсов показывает, что задача определения оптимального плана развития народного хозяйства есть задача социально-экономического развития и что системе моделей оптимального экономического планирования должна соответствовать определенная система моделей «пограничных» социальных процессов.

2. МОДЕЛИ ПОВЕДЕНИЯ ЧЕЛОВЕКА

Экономическая система включает в себя два типа поведения человека, занятого в сфере экономики: производственную деятельность (обязательное выполнение определенных общественных функций) и потребительский спрос (в основном свободная деятельность, к которой здоровое общество не принуждает своих членов насильно). Одна из главнейших задач конкретной социологии — изучение законов формирования поведения человека в обществе, позволяющих предсказывать это поведение и оказывать на него допустимое влияние. Предположение причинной обусловленности социального поведения позволяет построить модель этого поведения. Будем предполагать, что всем признакам, которые характеризуют социальные поведение и параметры. При этом переменными будем считать некоторые наборы формальных элементов (числа, векторы, индексы и др. — см. раздел 4), которые могут меняться в течение нашего исследования (они называются еще неизвестными). Параметрами будем считать неизменные «величины», еще характеризующие весь класс изучаемых ситуаций. Непосредственное разбиение на переменные и параметры производится в зависимости от задачи [3, 4].

Пусть переменная (и параметры) y выражает такие признаки самого социального поведения, которые нас интересуют. Ими могут быть: те или иные варианты производительности труда, спрос, использование досуга, общественная деятельность, принятие решений по различным вопросам и т. п. Пусть далее переменная (и параметр) $x = (x_1, x_2)$ выражает условия, в которых находится индивид x_1 , а также тип самого индивида x_2 . К усло-

* Может показаться, что в оптимальном плане люди желают работать и работают именно там, где выгодно с точки зрения общей производительности. Однако при ограниченном количестве трудовых и других ресурсов подобные надежды, мягко говоря, иллюзорны.

виям могут относиться местожительство, доход, образование, семья индивида, общественное мнение социальной группы, в которую входит индивид, количество и качество информации, которой располагает индивид, и т. д. Сам индивид может характеризоваться такими признаками, как пол, возраст, национальность, характер, темперамент и другие био-психологические

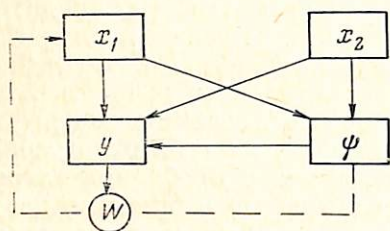


Рис. 1

особенности. Наконец, пусть ψ означает основные характеристики отношения индивида к внешнему миру, к условиям своей и чужой жизни, к возможным актам поведения, субъективные желания и устремления и т. п. — все то, что называют системой ценностной ориентации личности. Теперь можно представить модель социального поведения личности в виде некоторой функциональной схемы (рис. 1), где стрелками обозначено причинное влияние условий

$$x \in X(w), \quad y \in Y(x), \quad \Psi_{x,w}(y) = \max. \quad (1)$$

жизни x_1 и x_2 и отношения ψ на поведение y , а пунктиром — зависимость условий и ценностной ориентации от самого общества (условно обозначаемого буквой w). Если предположить, что множество возможных «поведения» y зависит от условий x , а отношение ψ , также зависящее от условий, упорядочивает это множество по предпочтению некоторой целевой функцией, то можно символически так выразить правила формирования поведения:

В такой записи первая формула показывает, что условия жизни индивида зависят, вообще говоря, от жизни самого общества. Вторая формула означает, что осуществляется лишь одно из многих возможных поведений, определяемых теми же условиями. Третья формула выражает факт выбора конкретного поведения самим индивидом в соответствии со своими целями, желаниями и взглядами, которые, правда, зависят от условий жизни индивида и от самого общества. Отметим еще, что если не предполагается влиять на поведение изменением условий или отношения индивида, то x и w будут параметрами модели, характеризующими среду, в которой находится индивид, а y будет переменной, подлежащей определению.

При конструировании моделей поведения надо иметь в виду еще два обстоятельства. Первое — это то, что моделируют и статические ситуации (т. е. относящиеся к фиксированному отрезку времени) и динамические (когда рассматривается поведение в течение нескольких интервалов времени, причем поведение в будущем зависит от поведения прошлого). Моделирование поведения в фиксированных, стабильных условиях является необходимой предпосылкой для построения динамических моделей и в настоящее время наиболее развито. С формальной стороны динамические модели более сложны. Теория позиционных игр как раз рассматривает некоторые формальные аспекты динамического поведения так называемых игроков в зависимости от ситуаций, в которых они находятся. Второе обстоятельство относится к статистической природе социального поведения человека. При всей «индивидуальной неповторимости и своеобразии» отдельного человека наличие общих законов развития природы и общества приводит к тому, что и в условиях жизни и в самом поведении отдельных личностей проявляется много общих черт. Поэтому, хотя предсказать поведение отдельного человека в рамках социологии и невозможно (в детерминистском понимании этого слова), поведение массового, «статистического» индивида вполне предсказуемо в силу возможности применения законов теории вероятностей и статистики к социальной действительности.

Задачу моделирования поведения человека с целью прогноза и управления можно рассмотреть на примере изучения использования человеком своего свободного времени. Введем следующие обозначения: t_j — количество свободного времени, затрачиваемого на определенный тип j время препровождения в единицу времени, например, неделю, месяц (занятия спортом, чтение, самообразование, посещение кино, театра, хождение в гости и т. п.); T — общий резерв свободного времени; b_j — некоторые естественные ограничения (которых может и не быть) на использование свободного времени j -м способом (например, на селе показывается ограниченное число кинофильмов в неделю, поэтому индивид может посмотреть не больше кинофильмов, чем их показывают). Иногда имеет смысл вводить и нижнее ограничение; a_j означает ту величину затрат свободного времени j -м способом, которая заведомо будет иметь место.

Будем также предполагать, что индивиду не безразлично, как он проводит свое время и что этот факт может быть выражен с помощью функции предпочтения, принимающей тем большее значение, чем предпочтительнее данный способ использования времени. Обозначим эту функцию через $U(t) = U(t_1, t_2, \dots, t_n)$, где n — число возможных затрат времени. В наших обозначениях модель поведения, в принципе, может быть представлена в виде задачи математического программирования

$$\begin{aligned} t_1 + t_2 + \dots + t_n &\leq T, \\ a_j &\leq t_j \leq b_j, \quad j = 1, 2, \dots, n, \\ U(t_1, t_2, \dots, t_n) &= \max. \end{aligned} \quad (2)$$

Смысл этой модели в том, что на основании ее можно определять затраты свободного времени по тем условиям, в которых находится изучаемый индивид. Иными словами, зная общее количество свободного времени и некоторую специфику жизни изучаемых людей, можно указать, сколько раз эти люди будут ходить в кино, театр, в гости, сколько прочтут книг, сколько времени уделят спорту и т. д. Но для этого необходимо знать структуру предпочтений индивидов, даваемую функцией $U(t)$. Эта функция, как легко видеть, имеет природу, аналогичную известной функции потребления в теории потребительского выбора. И сама модель (2) представляет собой аналог модели потребления. Как и в случае потребления, при наличии только одного первого ограничения в (2) оптимальное решение задачи, а значит и конкретное поведение, определяется лишь общим фондом свободного времени T , т. е.

$$x_j^0 = f_j(T), \quad j = 1, 2, \dots, n.$$

На материалах одного конкретного социологического обследования были построены функции $f_j(T)$ * в предположении параболической зависимости. Так, для женщин-рабочих в возрасте 19—28 лет, имеющих детей, затраты времени определяются формулой

$$\begin{aligned} \text{чтение } y &= 1,061 + 0,2227 T - 0,0001 T^2, \\ \text{«зрелища» } y &= 0,387 + 0,1983 T + 0,0034 T^2. \end{aligned}$$

Зависимости оказались почти линейными, что объясняется большим разбросом точек вокруг среднего и вообще малой достоверностью данных. Однако при наличии качественной информации могут быть получены зависимости и от общей величины вне рабочего времени, что позволит де-

* См. кандидатскую диссертацию С. Ф. Фролова «Социальное значение вне рабочего времени в период перехода к коммунизму», и Отчет Лаборатории социологических моделей ЦЭМИ АН СССР за 1968 г.

лать заключения об ожидаемых расходах людьми времени при сокращении рабочего дня и т. п.

Следует отметить, что при анализе поведения людей через затраты ими своего вне рабочего времени особенно необходимо учитывать совмещения (одновременные занятия разными делами), а также четко выделять самого субъекта поведения (семья, индивид, коллектив).

3. ЗАДАЧА АНАЛИЗА СИСТЕМЫ ВЗАИМОСВЯЗАННЫХ ПРИЗНАКОВ

Общественный индивид, которого изучает конкретная социология, представляет собой сложную систему в том смысле, что число переменных, выражающих существенные признаки данного индивида, может быть очень большим, а сами эти признаки (и переменные) находятся в сложной взаимной связи.

Пусть, например, исследуется успеваемость студентов. В этом случае каждому студенту из обследуемой совокупности может быть поставлен в соответствие набор признаков: пол, возраст, место окончания средней школы, успеваемость в школе, образование родителей и т. д. Естественно поставить вопрос: в какой степени влияют все эти признаки на успеваемость? Что произойдет, если изменится один или несколько признаков? И т. п. Очень часто на подобные вопросы отвечают с помощью корреляционного и регрессионного анализа. Однако для удовлетворительного применения этих методов необходимо выполнение ряда предположений относительно переменных: непрерывность, нормальность распределений и т. д. Кроме того, знания коэффициентов корреляции или вида регрессионной зависимости не всегда достаточно, чтобы ответить, например, на такой вопрос: с какой вероятностью данный тип индивида в данных условиях будет вести себя каким-то определенным образом?

Другая трудность статистического изучения «общественного индивида» связана с проблемой преодоления огромной размерности, сложности тех показателей, которые существенны для исследователя. Особенно это дает себя чувствовать при попытках изучения структуры трудовых ресурсов и населения. Разберем эту задачу более подробно на конкретном примере*. Предположим, нас интересует структура труда на некотором производстве, определяемая следующими шестью признаками: 1) возраст, 2) зарплата, 3) профессия, 4) образование, 5) «образованность семьи» (максимальный уровень образования в семье); 6) стаж работы по профессии.

По каждому признаку j ($j = 1, 2, \dots, 6$) будем различать несколько градаций r_j , обозначая k -ю градацию j -го признака через (j, k) . Содержательный смысл всех градаций приводится в табл. 1.

Набор этих признаков с фиксированными градациями определяет конкретную социально-производственную группу. Например, набор

$$h = [(1.1), (2.2), (3.3), (4.3), (5.3), (6.1)] \quad (3)$$

определяет группу столяров (3.3) до 20 лет (1.1), получающих зарплату от 51 до 90 руб. (2.2), имеющих образование 8—10 классов (4.3), не имеющих в семье членов с более высоким образованием (5.3), и со стажем менее одного года (6.1). Очевидно, что общее число всевозможных групп в нашем случае равно

$$N = r_1 \cdot r_2 \cdot r_3 \cdot r_4 \cdot r_5 \cdot r_6 = 4 \cdot 3 \cdot 6 \cdot 4 \cdot 4 \cdot 4 = 3456.$$

* Далее рассматриваются некоторые из результатов обследования Лабораторией социологических моделей ЦЭМИ АН СССР одного из промышленных предприятий.

Таблица 1

j	$(j \cdot k)$	Градации	r_j
(1)	(1.1)	До 20 лет	4
	(1.2)	21—30 лет	
	(1.3)	31—60 лет	
	(1.4)	Свыше 60 лет	
(2)	(2.1)	До 50 руб.	3
	(2.2)	50—90 руб.	
	(2.3)	Свыше 90 руб.	
(3)	(3.1)	Подсобные рабочие	6
	(3.2)	Станочники	
	(3.3)	Столяры	
	(3.4)	Слесари	
	(3.5)	Служащие низкой квалификации	
(4)	(4.1)	Служащие средней и высокой квалификации	4
	(4.2)	До 4 классов	
	(4.3)	5—7 классов	
	(4.4)	8—10 классов	
(5)	(5.1)	Среднетехническое и высшее образование	4
	(5.2)	До 4 классов	
	(5.3)	5—7 классов	
	(5.4)	8—10 классов	
(6)	(6.1)	Среднетехническое и высшее образование	4
	(6.2)	До 1 года	
	(6.3)	1—5 лет	
	(6.4)	6—13 лет	
		Свыше 15 лет	

Другими словами, знать структуру труда данного изучаемого производства — это значит знать, сколько имеется человек (в абсолютных или относительных единицах) в каждой из 3456 групп, т. е. знать функцию распределения индивидов по группам. Если признаков будет больше, то число возможных групп может стать совсем фантастическим. Интересно, что число лиц в исследуемой совокупности может быть в десятки и сотни раз меньше, чем число «потенциально-возможных лиц». В этом как раз проявляется сложность человека как системы, т. е. «разнообразие», заложенное в выделяемые признаки индивида, оказывается необыкновенно большим. При анализе структуры труда встает задача: как выразить (в аналитическом или табличном виде) доли лиц в каждой из потенциально-возможных групп. Даже в нашем случае задача оказывается довольно-таки сложной, поскольку таблица из 3456 клеток — не очень удобный способ хранения информации, а также по причине неудобства перебора всех группировок признаков с целью установления числа членов во всех группах. Тем не менее надо уметь находить любое из значений функции распределения* индивидов по всем возможным группам. Обозначим эту функцию распределения через $F[h] = F[(1), (2), (3), (4), (5), (6)]$, значения которой в «точках» $(j \cdot k)$ определяют долю лиц в данной группе. Так, для группы (3) функция имеет значение

$$F = F[(1.1), (2.2), (3.3), (4.3), (5.3), (6.1)]. \quad (3')$$

Обозначим еще через $F[(j) / (k), (l), (m), \dots, (q)]$ условную функцию распределения j -го признака при фиксированных значениях признаков $k, l, m, \dots, q$.

* Далее под функциями распределения понимаются безусловные и условные плотности распределения вероятностей и не делается различия между такой функцией и обычной гистограммой.

Если бы все признаки (j) были независимыми, то функция $F[h]$ выражалась бы довольно просто

$$F(h) = F_1[(1)] \cdot F_2[(2)] \cdot F_3[(3)] \cdot F_4[(4)] \cdot F_5[(5)] \cdot F_6[(6)]. \quad (4)$$

Однако при наличии взаимной зависимости формула (4) не имеет места. Но если ввести предположение о том, что между признаками существуют непосредственные зависимости и опосредованные, то можно будет вывести формулу, аналогичную (4). Для более строгого утверждения необходимо воспользоваться понятием графа. Будем называть сим-

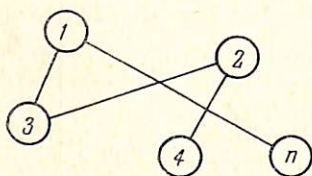


Рис. 2

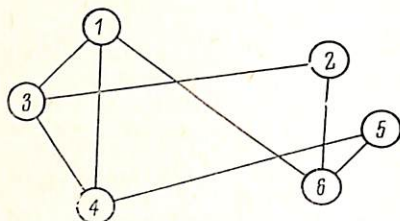


Рис. 3

метричным графом множество точек на плоскости $\{1, 2, \dots, n\}$, часть которых соединена попарно отрезками линий (i, k) . Точки называют вершинами графа, соединяющие их отрезки — ребрами. Путем из вершины i в вершину j назовем последовательность ребер, таких, что конец предыдущего ребра является началом следующего, и начало первого ребра в этой последовательности есть вершина i , а конец последнего — вершина j (при этом каждое ребро встречается не более одного раза). Так, на рис. 2 путем из вершины 3 в вершину n будет последовательность $\{(3,1); (1,n)\}$.

Сопоставим теперь множеству исследуемых признаков некоторый граф*. Будем называть такой граф гипотетическим графом связи всех n признаков, если для функции распределения $F[(1), (2), \dots, (n)]$ имеет место

$$F[(j) / (i), \dots, (l); (m), \dots, (q)] = F[(j) / (i), \dots, (l)] \quad (5)$$

всякий раз, когда не существует путей из вершины $(m), \dots, (q)$ в вершину (j) , которые не проходили бы через вершины $(i), \dots, (l)$. Физический смысл этого определения довольно прост. Он состоит в том, что если признаки $m, \dots, q$ связаны с признаком j только через признаки $i, \dots, l$, то для вероятностных утверждений относительно признака j при известных градациях признаков $i, \dots, l$ нет необходимости знать что-либо о градациях признаков $m, \dots, q$.

При изучении взаимных связей шести упоминавшихся признаков было сделано предположение, что для них гипотетический граф связи имеет вид, показанный на рис. 3.

В частности, согласно этому графу, зарплата индивидов связана с «образованностью семьи» только через стаж и образование самого индивида, а с образованием — через профессию и стаж. Непосредственно связь между образованностью семьи и стажем кажется несколько странной и, по-видимому, объясняется неполнотой выделенных признаков.

Теперь функцию распределения F можно выразить через условные функции распределения, каждая из которых зависит от небольшого числа

* Ниже излагается опыт применения теории случайных величин, задаваемых при помощи некоторого графа. Формальный аспект этой теории будет опубликован в готовящихся к изданию других работах.

аргументов. В самом деле, можно написать

$$F[(1), (2), (3), (4), (5), (6)] = F[(1), (2), (3), (4), (6)] \times \\ \times F[(5) / (1), (2), (3), (4), (6)]. \quad (6)$$

Но согласно формуле (5) и виду графа (рис. 3), признаки (1), (2), (3) связаны с (5) только через (4) и (6). Поэтому формула (6) примет вид

$$F[(1), (2), (3), (4), (5), (6)] = F[(1), (2), (3), (4), (6)] \times \\ \times F[(5) / (4), (6)]. \quad (7)$$

Первый сомножитель последней формулы снова может быть представлен в виде произведения

$$F[(1), (2), (3), (4), (6)] = F[(1), (3), (4), (6)] \times \\ \times F[(2) / (1), (3), (4), (6)]. \quad (8)$$

Так как (1) и (4) признаки связаны с признаком (2) только через (3) и (6), то во втором сомножителе (8) их можно опустить

$$F[(1), (2), (3), (4), (6)] = F[(1), (3), (4), (6)] \cdot F[(2) / (3), (6)]. \quad (9)$$

Упрощенное представление функции $F[(1), (3), (4), (6)]$ нуждается в дополнительном предположении. Как обычно, имеем

$$F[(1), (3), (4), (6)] = F[(1), (4), (6)] \cdot F[(3) / (1), (4), (6)]. \quad (10)$$

Рассмотрим второй сомножитель формулы (10). Признаки (1) и (4) связаны с признаком (3) непосредственно (см. рис. 3), поэтому их отбрасывать нельзя. Но, с другой стороны, признак (6), хотя и не связан с признаком (3) непосредственно, все же связан с (3) не только через (1) и (4), но и через посредство признака (2). Поэтому на основании формулы (5) его также нельзя отбрасывать. Если же предположить дополнительно, что опосредствованное влияние профессии (3) на стаж (6) через зарплату (2) незначительно, т. е.

$$F[(3) / (1), (4), (6)] \approx F[(3) / (1), (4)], \quad (11)$$

то, подставляя (11) в (10), (10) в (9), (9) в (8), а (8) в (7), получим

$$F(h) = F[(1), (2), (3), (4), (5), (6)] = \\ = F[(5) / (4), (6)] \cdot F[(2) / (3), (6)] \cdot F[(3) / (1), (4)] \cdot F[(1), (4), (6)]. \quad (12)$$

Или, вводя простые обозначения для условных и безусловных функций распределения, получим

$$P_{1,2,3,4,5,6} = P_{4,6}^5 \cdot P_{3,6}^2 \cdot P_{1,4}^3 \cdot P_{1,4,6}^1. \quad (13)$$

Таким образом, чтобы найти всю функцию распределения шести признаков, достаточно знать три таблицы условных распределений и одну безусловного. Подсчитаем число клеток в этих таблицах.

Для таблицы	$P_{4,6}^5: r_4 \cdot r_5 \cdot r_6 = 4 \cdot 4 \cdot 4 = 64,$
»	$P_{3,6}^2: r_2 \cdot r_3 \cdot r_6 = 3 \cdot 6 \cdot 4 = 72,$
»	$P_{1,4}^3: r_1 \cdot r_3 \cdot r_4 = 3 \cdot 4 \cdot 6 = 72,$
»	$P_{1,4,6}^1: r_1 \cdot r_4 \cdot r_6 = 3 \cdot 4 \cdot 4 = 48.$

Общее число необходимых значений равно $64 + 72 + 72 + 48 = 256 < N = 3456$.

Итак, знание 256 чисел позволяет нам вычислить любое из 3456 значений искомой функции распределения.

Формула (13) была получена и проверена на материалах одного конкретного социологического обследования, результатом которого явились 275 анкет случайной выборки, представляющей $1/6$ генеральной совокупности. Для этого сначала были составлены 15(6·5/2) табличек парных группировок всех признаков. Так, парные таблицы возраста (1) с «образованием семьи» (5) и стажем (6) имеют вид в процентах (см. табл. 2а и 2б, 3а и 3б):

Таблица 2а						Таблица 2б					
	5.1	5.2	5.3	5.4	Σ		6.1	6.2	6.3	6.4	Σ
1.1	00	11	71	18	100	1.1	70	30	00	00	100
1.2	04	19	60	17	100	1.2	14	51	35	00	100
1.3	08	26	33	33	100	1.3	11	43	41	05	100
1.4	25	25	25	25	100	1.4	25	25	25	25	100

а таблицы парной связи профессии (3) с образованием и стажем (6):

Таблица 3а						Таблица 3б					
	4.1	4.2	4.3	4.4	Σ		6.1	6.2	6.3	6.4	Σ
3.1	50	25	25	0	100	3.1	13	44	25	18	100
3.2	9	41	45	5	100	3.2	14	41	42	3	100
3.3	5	32	59	4	100	3.3	30	44	26	0	100
3.4	3	49	35	10	100	3.4	17	40	40	3	100
3.5	0	6	71	23	100	3.5	12	53	35	0	100
3.6	0	0	26	74	100	3.6	4	57	37	2	100

Анализ этих таблиц показывает, что в одних таблицах колеблемость значений вокруг среднего для столбца велика, в других мала. В качестве меры такой колеблемости можно взять величину

$$\Delta = \frac{1}{m \cdot n} \sum_{i=1}^m \sum_{j=1}^n \frac{|x_{ij} - \bar{x}_j|}{\bar{x}_j}, \quad (14)$$

где x_{ij} — значение элемента i -й строки и j -го столбца, $\bar{x}_j$ — среднее значение для j -го столбца.

Для наших четырех таблиц значения Δ равны

$$\begin{aligned} \Delta(1,5) &= 45\%, & \Delta(3,4) &= 79\%, \\ \Delta(1,6) &= 64\%, & \Delta(3,6) &= 15\%. \end{aligned}$$

Можно предположить, что непосредственные связи между признаками вызывают колеблемость (величина Δ) не меньше 50%. Тогда необходимо заключить, что между признаками (1) и (6), (3) и (4) может существовать непосредственная связь, а между признаками (1) и (5), (3) и (6) она отсутствует. На основании подобных рассуждений, а также некоторых общих соображений и был составлен весь гипотетический граф связи*.

* На самом деле при нахождении графа связи использовался критерий $\Delta = \frac{1}{2}(\Delta^1 + \Delta^2)$, где Δ^1 получалась по формуле (14), а Δ^2 — по аналогичной формуле, но для которой в таблицах парных распределений строки и столбцы менялись местами.

После составления графа связи (рис. 3) и выведения формулы (13), все 275 анкет были подвергнуты четыре раза трехкратным группировкам с целью получения таблиц условных и безусловных частот (вероятностей): $P_{4,6}^5$; $P_{3,6}^2$; $P_{1,4}^3$; $P_{1,4,6}^1$. Приведем первую из этих таблиц в абсолютных (табл. 4а) и в относительных (табл. 4б) единицах.

Таблица 5

Группы	Теоретическая доля (%)	Истинная доля (%)	Расчетные числа анкет	Истинное число анкет
(1.1), (2.2), (3.3), (4.3), (5.1), (6.1)	001	000	0	0
(6.2)	003	007	1	2
(6.3)	020	029	6	8
(6.4)	004	007	1	2
(1.2), (2.3), (3.6), (4.4), (6.1)	002	004	1	1
(6.2)	019	015	5	4
(6.3)	019	025	5	7
(1.3), (2.3), (3.4), (4.1)	002	004	0	0
(4.2)	038	040	10	11
(4.3)	006	007	2	2
(4.4)	004	003	2	1
(1.2), (2.3), (3.1)	007	004	2	1
(3.2)	043	040	11	11
(3.3)	096	094	27	26
(3.4)	024	022	6	6
(3.5)	019	018	5	5
(3.6)	044	047	12	13
(1.2), (2.1)	001	004	0	1
(2.2)	194	193	53	53
(2.3)	229	225	68	62

После этого было выбрано 70 комбинаций признаков («социально-производственных групп»), для них рассчитаны по формуле (13) теоретические доли индивидов от всей совокупности (275) и сравнены с действительными долями, определенными многократными группировками. Приведем несколько результатов такой проверки. Так как относительное количество индивидов (анкет) в фиксированной группе определяется суммированием функции $F(h)$ по всем индексам, не входящим в данную комбинацию, то

$$F[(1), (2), \dots, (6)] = P_{4,5}^5 \cdot P_{3,6}^2 \cdot P_{1,4}^3 \cdot P_{1,4,6}^1,$$

$$F[(1), (2), (3), (4), (6)] = \sum_{(5)} P = P_{3,6}^2 \cdot P_{1,4}^3 \cdot P_{1,4,6}^1,$$

$$F[(1), (2), (3), (4)] = \sum_{(5)(6)} P = P_{1,4}^3 \cdot \sum_{(6)} P_{3,6}^2 \cdot P_{1,4,6}^1,$$

$$F[(1), (2), (3)] = \sum_{(4)} P_{1,4}^3 \cdot \sum_{(6)} P_{3,6}^2 \cdot P_{1,4,6}^1,$$

$$F[(1), (2)] = \sum_{(3)(4)} P_{1,4}^3 \cdot \sum_{(6)} P_{3,6}^2 \cdot P_{1,4,6}^1$$

(суммирование функции по верхнему индексу дает единицу, см., например, табл. 4б). Сравнение теоретических (расчетных) и действительных

величин дается в табл. 5. Сопоставление истинных и теоретических величин показывает довольно удовлетворительное совпадение. При этом, как правило, относительная ошибка меньше для величин, абсолютные значения которых больше. Формулой (13) можно пользоваться не только для определения количества индивидов в той или иной группе. Она еще может показывать «очищенное» влияние одних признаков на другие и главное — дает непосредственные вероятности одних значений признаков при фиксировании любых других (то, о чем говорилось в начале этого раздела).

4. О ПРОБЛЕМЕ ИЗМЕРЕНИЯ В СОЦИОЛОГИИ

До сих пор мы говорили о различных признаках индивидов, о переменных, параметрах, характеризующих их, пользовались условными обозначениями различных градаций этих признаков без особых оговорок. Однако в современной науке существует целое направление исследований, которыми занимаются социологи, математики, психологи, посвященное разработке методов измерения социальных и психологических признаков или методам их количественного представления (см. [5]). Собственно говоря, без формального выражения тех явлений, которые изучаются, не может быть и математической социологии (или экономики). А формальное представление, если ему соответствует строгая и допускающая последующие повторения процедура, позволяет оставлять и исследовать лишь самое существенное из всего массива информации об объектах мира, конструировать и изучать математические модели вместо непосредственного изучения самих объектов, использовать ЭВМ для обработки огромного количества данных, на основании чего можно делать практические выводы и выбирать наилучшую линию поведения.

В чем состоит эта проблема? И что значит измерить некоторый признак? Прежде всего должно быть ясно, что до измерения некоторого признака необходимо иметь способ: 1) установления самого наличия признака у исследуемого объекта и 2) установления какого-то (>1) числа различий (градаций) у данного признака. После выполнения указанных физических операций можно говорить о задаче измерения признака, заключающейся в установлении соответствия между множеством различных градаций признака и некоторым формальным множеством (например, между множеством целых чисел, положительных чисел, каких-то других символов и т. д.). При этом формальное множество, элементы которого теперь будут значениями нашего признака, должно быть не каким угодно, а таковым, что формальные операции над его элементами (например, сложение, умножение и т. п.) будут строго соответствовать некоторым реальным операциям над изучаемыми объектами (перемещение, объединение и т. д.). Известная проблема критерия оптимальности экономической системы как раз является примером измерения такого признака состояния экономики, как общественное благосостояние или общественная полезность, соответствующие данному состоянию. В этом случае каждому значению (вектору состояний экономики) пытаются приписать определенное число, поэтому задача состоит в определении функции общего благосостояния. Но, как часто оказывается, над этой функцией допустимы не все операции, поскольку она может быть определена лишь с точностью до монотонного преобразования (однако мы не будем здесь больше касаться данной темы, которой посвящено и так достаточно большое количество работ).

Обратимся теперь к признакам, с которыми мы уже работали и которые со своими градациями приведены в табл. 1. Как нетрудно видеть,

такие признаки, как возраст (1), зарплата (2), стаж (6), сами по себе допускают естественное выражение через число — в годах или рублях. Образование индивида можно также измерять в годах, но встает дополнительная проблема соизмерения различных видов специального образования. Если же перейти к «образованности семьи» (6), то по сравнению с предыдущим случаем надо образование всех членов семьи сводить каким-то образом к одному показателю (мы это осуществляли сведением к наибольшему образованию членов семьи). Однако если попытаться количественно «измерить», т. е. выразить в числе, профессию индивида (3), то здесь «естественной» меры можно уже и вообще не отыскать. В ряде случаев пытаются характеризовать профессию числом, обозначающим сложность данного вида труда, или что-либо подобное. Действительно, иногда можно располагать профессии в условный ряд «по порядку», но тогда «мера», «число» уже будут не обычным числом, а лишь порядковым номером, и характеризовать не профессию как таковую, а некоторое ее свойство (сложность, важность и т. п.).

Множество чисел, которые могут соответствовать тем или иным признакам и с которыми допустимы те или иные операции, называются шкалами. Существует много различных типов шкал, отметим лишь: 1) обычную числовую шкалу (как в случае возраста или зарплат), 2) порядковую шкалу рангов (как в случае распределения мест на шахматном турнире) и 3) шкалу наименований (когда числа вообще не приписываются градациям, которым могут соответствовать любые различающиеся символы).

Надо ясно понимать, что не существует никаких абсолютных шкал и способов представления признаков в числах. Поэтому результаты исследований, утверждения относительно признаков, связей между ними и видов количественных закономерностей справедливы лишь с точностью до типов шкал, которыми мы пользуемся, до числа градаций и способов выделения этих градаций. При одном способе разбиения на градации признаки будут зависимыми, при другом могут быть независимыми и т. д. Существует много способов измерения качественных признаков, для этого разработаны специальные математические алгоритмы и приемы, но о них имеет смысл рассказать в специальной статье.

ЛИТЕРАТУРА

1. К. Маркс и социология. Информ. бюл. № 3 Научного совета АН СССР по проблемам конкретных социальных исследований. М., 1968 (ротапринт).
2. В. И. Ленин. Полн. собр. соч., т. 1, стр. 139.
3. Социология и идеология. Сборник статей. М., «Наука», 1969.
4. Количественные методы в социальных исследованиях. М., «Наука», 1967.
5. Математические методы в современной буржуазной социологии. М., «Мир», 1967.

Поступила в редакцию
23 XII 1968