

УДК 311.213

А. Е. Жминько

ФГБОУ ВО «Кубанский государственный аграрный университет имени И.Т. Трубилина»,
Краснодар, e-mail: albina_evgenevna@mail.ru

Ю. Н. Шоль

ФГБОУ ВО «Кубанский государственный аграрный университет имени И.Т. Трубилина»,
Краснодар, e-mail: un_shol@mail.ru

И. Г. Сидорчукова

ФГБОУ ВО «Кубанский государственный аграрный университет имени И.Т. Трубилина»,
Краснодар, e-mail: sev_2111@mail.ru

ОСОБЕННОСТИ СБОРА И АНАЛИЗА ЭМПИРИЧЕСКОЙ ИНФОРМАЦИИ

Ключевые слова: анализ, методы, наблюдение, группировки, респонденты.

При организации наблюдения важным аспектом экономико-статистического исследования имеющейся информации является выбор и обоснование методов и способов анализа. При этом теоретико-концептуальный аппарат исследования форматируется за счет экономических и социологических понятий. Анализ эмпирических данных предполагает в полной мере учитывать специфические особенности имеющихся признаков и показателей, их интерпретацию, взаимосвязи и взаимозависимость. Широко используемый метод группировок, который является основным приемом статистического исследования, позволяет изучать особенности размещения отраслей экономики с отражением влияния факторов применительно к единицам наблюдения и их признакам. Метод группировок базируется на расчленении единиц, зарегистрированных при статистическом наблюдении, на группы по принципу однородности. Комбинационные и территориальные группировки совместно с другими показателями могут служить основанием при районировании отраслей народного хозяйства.

A. E. Zhminko

Kuban State Agrarian University named after I.T. Trubilin, Krasnodar,
e-mail: albina_evgenevna@mail.ru

Yu. N. Shol

Kuban State Agrarian University named after I.T. Trubilin, Krasnodar,
e-mail: un_shol@mail.ru

I. G. Sidorchukova

Kuban State Agrarian University named after I.T. Trubilin, Krasnodar,
e-mail: sev_2111@mail.ru

THE STRUCTURE OF THE EMPIRICAL INFORMATION ANALYSIS PROCESS

Keywords: analysis, methods, observation, groupings, respondents.

When organizing observation, an important aspect of the economic and statistical study of available information is the choice and justification of methods and methods of analysis. At the same time, the theoretical and conceptual apparatus of the study is formatted at the expense of economic and sociological concepts. The analysis of empirical data assumes that the specific features of the available signs and indicators, their interpretation, interrelationships and interdependence are fully taken into account. The widely used grouping method, which is the main technique of statistical research, allows us to study the peculiarities of the placement of economic sectors, reflecting the influence of factors in relation to observation units and their characteristics. The grouping method is based on the division of units registered during statistical observation into groups based on the principle of uniformity. Combinational and territorial groupings together with other indicators can serve as a basis for the zoning of branches of the national economy.

Из методов, применяемых для анализа эмпирической информации, в социально-экономических исследованиях шире всего используется метод группировок. Для организации обработки существенны следующие особенности эмпирической информации: структурированность представления информации, не альтернативная структура значений признаков, наличие функциональных признаков.

Первая особенность состоит в том, что информация представляет собой данные о нескольких типах объектов (единиц наблюдения). Это может быть населенный пункт, семья, член семьи, находящиеся в отношении включенности. Смысл названного отношения состоит в том, что характеристики (значения признаков) одной единицы наблюдения могут быть рассмотрены как характеристики другой единицы наблюдения.

Такие отношения имеют древовидную (иерархическую) структуру. В процессе анализа и обработки информации единицы наблюдения каждого типа выступают в двух ролях: как элементы одноименной статистической совокупности и как источники дополнительных характеристик для единиц наблюдения другого типа, с которыми они связаны отношением включенности.

Целью данной работы является анализ и оценка широко применяемого группировочного метода для анализа эмпирической информации, кроме того, оценка положительных и отрицательных особенностей его использования.

Любое статистическое исследование начинается со сбора сведений или фактов об изучаемых явлениях и процессах. Научно организованное статистическое наблюдение должно быть проведено с соблюдением определенных правил с целью получения сведений о каждой единице наблюдения. Задача второй стадии статистического исследования состоит в том, чтобы упорядочить и обобщить первичный материал, свести его в группы и на этой основе дать обобщенную характеристику совокупности. Статистическая группировка сводится к выделению групп по существенному для единиц совокупности признаку.

Рассмотрим основные процедуры, входящие в состав этого метода.

Во-первых, производится структурирование теоретической модели изучаемого явления с целью выделения относительно независимых блоков. Типичным примером может служить иерархическая структура те-

оретического понятия образа жизни. Корнем дерева служит само понятие образа жизни. Затем оно разбивается на ряд компонент: труд, досуг, быт, участие в управлении. Каждая из компонент, в свою очередь, расчленяется на такие подкомпоненты, как содержание труда, условия труда, его режим и прочее. Затем, каждой подкомпоненте ставится в соответствие один или несколько признаков. Например, подкомпоненте «содержание труда» соответствует один признак, а подкомпоненте «условия» – несколько, такие как характеристика рабочего места, психологический климат в коллективе.

Во-вторых, в рамках выделенных структурных блоков теоретической модели явления проводится эмпирический анализ. В частности, при изучении образа жизни эмпирический анализ состоял в последовательном: построений сначала шкал подкомпонент на основе признаков, затем шкал компонент на основе шкал подкомпонент, наконец, шкалы (типологии) образа жизни на основе шкал компонент.

В-третьих, необходим процесс соединения теоретических и эмпирических построений, который на каждом этапе изучения носит, как правило, итеративный характер. Опыт показывает, что правила агрегирования эмпирических данных, вытекающие из априорных теоретических построений, почти всегда нуждаются в корректировке. Дело в том, что большое число признаков приводит к большому числу логически возможных типов (групп объектов) и тем самым затрудняет построение теоретических критериев группировки объектов. В этой ситуации исследователи руководствуются различными приемами.

При определении системы признаков наряду с теоретическими соображениями учитывается число логически возможных групп объектов, порождаемых данной системой признаков. Вычисление эмпирического распределения по системе признаков дает возможность:

а) значительно сократить число анализируемых типов путем исключения статистически незначимых (малочисленных) групп объектов из числа логически возможных;

б) выделить соответствующие многочисленным группам объектов, значения признаков в качестве типобразующих. На базе таких значений могут быть выделены основания для построения критерия группировки объектов. Возможен и другой

исход анализа распределения, согласно которому система признаков может быть признана неудовлетворительной. В результате потребуется корректировка теоретических построений в формирование новой системы признаков.

Таким образом, процесс анализа объектов конкретных социально-экономических исследований методом группировок является многоэтапным и итеративным, причем связь между итерациями осуществляется на базе содержательных соображений. Структурная схема процесса представлена на рисунке.

Входом процесса является теоретическая концепция изучаемого явления и совокупность данных массовых наблюдений о нем, выходом – статистическое описание типологии изучаемого явления. Переход от входной информации к выходной осуществляется путем последовательного применения к ней пяти чипов преобразований.

Преобразование П1 состоит в формулировании теоретической типологии изучаемого явления на основе его теоретической концепции;

П2 – в постановке задачи обработки данных, т. е. в указании правил группирования единиц наблюдений и построения статистических описаний группировок. Назначение преобразования П3 – представить задачу обработки данных на формальном языке, на котором осуществляется общение человека с ЭВМ. Преобразование П4 сводится к обработке данных, т. е. преобразованию эмпирической информации в соответствии с описанием этого преобразования на формальном языке. Преобразование П5 состоит в анализе результатов обработки данных с целью выяснения степени адекватности полученного статистического описания той теоретической типологии, для которой оно предназначено.

Как видно из рисунка 1, связь между названными типами преобразований информации носит циклический характер. В роли управляющего элемента в цепи обратной связи выступает преобразование П5.



Анализ эмпирической информации методом группировок

Наличие обратной связи обусловлено тем, что в социально-экономических исследованиях статистические данные, как правило, не полностью удовлетворяют требованиям теоретической концепции. В свою очередь, это связано с тем, что операционализация теоретической концепции в эмпирические понятия не заканчивается разработкой статистического инструментария. Она может существенно дополняться и уточняться на стадии анализа данных.

Не альтернативная структура значений признаков имеет место, когда вопросы закрываются с помощью набора не взаимоисключающих подсказок, предлагаемых для выбора. Например, в анкете «Сельский житель» па вопрос «Почему Вы хотите уехать из деревни?» предлагаются следующие варианты ответов: 1) трудно продолжить образование, получить специальность; 2) нет возможности удовлетворять культурные запросы; 3) трудно найти работу, удовлетворяющую склонностям и интересам; 4) не устраивают заработки; 5) плохое снабжение; 6) семейные причины.

Этот набор подсказок не представляет собой континуума какого-либо одного свойства (в данном случае – причин отъезда на деревню), а дает перечисление разных причин. Характеристикой объекта в данном случае является набор подсказок, указанных респондентом. Естественно, эти наборы можно рассматривать в качестве символов, как и для качественных альтернативных признаков. Но тогда теряется возможность группировать объекты как по отдельным подсказкам, так и по числу подсказок в наборе.

Покажем, какие преимущества дает использование не альтернативных схем сбора информации. Эти преимущества становятся очевидными при сравнении с повсеместно распространенными альтернативными схемами сбора и анализа информации. Прежде всего, надо сказать, что не альтернативная информация возникает из двух источников. Во-первых, существуют вопросы, которые нельзя задать в полной форме. Например, ответ на вопрос «Какие вещи длительного пользования имеются в Вашей семье?» может представлять собой только одну форму: перечень вещей, которыми обладает данная семья. Во-вторых, не альтернативная информация появляется при формализации ответов на открытые вопросы анкет. Остановимся на втором источнике более подробно.

Общепризнано, что открытые вопросы целесообразно использовать на начальной стадии изучения проблемы, когда пути «закрывания», вопросов еще не ясны. Вторым этапом работы, как правило, является уяснение характера ответов, получаемых от населения, и формулировка адекватных им подсказок в альтернативной форме. Тогда непосредственно населению начинают адресоваться вопросы, заданные в альтернативной форме. Такая практика работы имеет ряд недостатков. Главный из них состоит в том, что собираемая информация оказывается искаженной. Если не вдаваться в частные случаи и примеры, то в целом искажение возникает как следствие требования выбора одного из нескольких вариантов ответов. Так как возможные ответы чаще всего не являются взаимоисключающими (ибо характеризуют равные стороны того или иного целого), то выбор одного и только одного из них чаще всего оказывается весьма искусственной процедурой, своего рода насильем над психикой респондента и социальной сущностью изучаемого явления.

Второй недостаток использования альтернативных схем сбора информации состоит в том, что с ним связаны потери информации, причем объемы теряемой информации чаще всего остаются неизвестными. Первое, чем вызываются потери – это необходимость ограничения числа подсказок. Само по себе ограничение преследует цель облегчить требуемый выбор одной позиции из всех предложенных: при прочих равных условиях, чем меньше число вариантов, тем легче выбор. Но другая сторона этого уменьшения – потеря информации, ибо чем меньше число вариантов, тем беднее представляется объект изучения.

Наконец, потери информации связаны с тем, что, как правило, большая часть опрашиваемых способна выбрать более чем один вариант из предложенных подсказок. Следовательно, пользуясь альтернативной схемой, мы обедняем описание каждого из респондентов в отдельности, упрощаем занимаемую им позицию.

В соответствии с двумя источниками возникновения не альтернативные схемы выполняют две функции. Они служат средством сбора информации в случаях, когда речь идет о явлениях большой сложности, и средством закрытия открытых вопросов, уменьшающим потерю информации по сравнению с альтернативными схемами.

Перейдем к рассмотрению последней из указанных выше особенностей эмпирической информации – наличия функциональных признаков. Функциональным будем называть признак, значения которого у некоторого объекта есть функция от значений признаков только этого объекта. Простейший пример функционального признака – возраст респондента, представленный в интервалах, которые вычисляются по абсолютным значениям возраста, зафиксированным в массиве данных.

Причины использования в анализе функциональных признаков состоят в следующем.

Во-первых, в терминах первичных признаков не может быть реализован один из важных приемов анализа методом группировок, каким является специализация признаков. Поясним этот прием на следующем примере. Предположим, что ставится задача классификации работников по тяжести труда. Конечной целью решения этой задачи должна быть некая классификация в стандартных терминах (работа тяжелая, средней тяжести, легкая и т. д.). Использование соответствующего вопроса сглаживает и огрубляет реальную картину тяжести труда. Так как содержание труда работников различных групп (например, у инженерно-технических работников и животноводов) чрезвычайно различается, то и критерии тяжести труда в таких случаях часто несопоставимы. В связи с этим для каждой группы работников приходится вводить разные критерии тяжести труда, которые и выступают в виде специальных признаков.

Во-вторых, в процессе анализа эмпирической информации последователь

получает новые знания об изучаемом явлении, которые приводят к корректировке первоначальных теоретических представлений. В результате признаки, составляющие эту информацию, становятся неадекватными решаемым задачам.

В-третьих, в ряде случаев характеристики изучаемого явления собираются в детализированном виде. Детализированная форма не только упрощает процедуру сбора информации, но и, что значительно важнее, обеспечивает возможность более углубленного и разностороннего анализа, сопоставления с другими исследованиями. В этой связи, например, информацию о доходе от личного подсобного хозяйства целесообразно собирать как данные в натуральной, а не в денежной форме.

Заключение

Наряду с названными объективными причинами ограниченных возможностей использования первичных признаков имеется ряд обстоятельств субъективного характера. Это некорректная формулировка вопросов в анкетах, несогласованность действий интервьюеров в процессе обследований. Следствием таких обстоятельств является определенная доля «шума» или выбросов в эмпирической информации. Один из способов его учета состоит в формировании новых функциональных признаков.

Исх проведенного исследования, анализ эмпирической информации требует построения группировок и их статистических описаний в терминах функциональных признаков.

Библиографический список

1. Vasilieva N.K., Reznichenko S.M., Vasiliev V.P., Trubilin A.I., Bershitskiy Yu.I. Economic stability of agricultural organizations in the region: conceptual-theoretic and applied aspects // International Journal of Economic Research. 2016. Т. 13. № 6. С. 2525-2540.
2. Vorobyova O., Morozov I., Shol Yu., Reshetnikova S., Valiev G., Improving customs control upon arrival of goods to the territory customs of the Eurasian Economic Union // Universidad y Sociedad. 2020. № 12.2. С. 447-451.
3. Бондаренко Н.Н., Бузыгина Н.С., Василевская Л.И. и др. Статистика: показатели и методы анализа: справочное пособие. Мн.: Современная школа, 2005. 628 с.
4. Елисеева И.И. Статистика: учебник. М.: Высшее образование, 2009. 566 с.
5. Жминько Н.С. Губиева С.Ю. Прогнозирование в экономике как универсальное средство предвидения развития производства и бизнеса // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2016. № 117. С. 1162-1172.
6. Жминько Н.С., Черненко Т.А. Способы минимизации предпринимательского риска // Формирование экономического потенциала субъектов хозяйственной деятельности: проблемы, перспективы, учетно-аналитическое обеспечение: материалы VI Международной научной конференции (Краснодар, 20-22 апреля 2016 г.) / ФГБОУ ВПО «Кубанский государственный аграрный университет». 2016. С. 345-352.
7. Кацко И.А., Ворокова Н.Х., Жминько А.Е., Сенникова А.Е. Статистические методы анализа данных: учебное пособие. Краснодар, 2017. 203 с.