

- ◇ участвующие: «псевдоконсультирование» с членами группы для углубления, детализации обсуждения того аспекта предмета дискуссии, который представляет особый интерес для исследователя;
- ◇ ориентированные на достижение цели: умение направлять дискуссию в русле, строго предусмотренном сценарием.

Дискуссия завершается кратким обобщением сказанного. Модератор благодарит участников.

Примерный отчет по итогам фокус-групп приведен в приложении 6.

ВОПРОСЫ ДЛЯ САМОКОНТРОЛЯ

1. Какие задачи призван решать метод фокус-групп?
2. Чем отличается метод фокус-групп от массовых опросов?
3. Перечислите основные требования к составу группы.
4. Какие основные требования предъявляются к модератору?
5. Дайте характеристику технологии дискуссии в фокус-группе.
6. Какие основные методы управления группой использует модератор?

ПРЕДЛАГАЕМ ВЕРНУТЬСЯ В МЕНЮ И РЕШИТЬ ЗАДАЧИ К ЭТОЙ ГЛАВЕ

Глава 14. Выборочный метод в прикладной социологии

Желать, чтобы наука охватывала природу, значило бы заставить целое войти в состав своей части.

А. Пуанкаре

14.1. Логика выборочного метода

Сам по себе объект исследования может быть очень большим: десятки и сотни тысяч людей, составляющих население города, области, региона; тысячи работников промышленного предприятия, научного учреждения или учебного заведения и т.д. Есть ли необходимость в том, чтобы опросить каждого из них? Когда анкетному опросу подлежит группа численностью 20–30 человек, он действительно может быть *сплошным* — охватывать каждого. При помощи анкеты можно

весьма оперативно опросить и большее число людей, если они находятся вместе, скажем, в кино- или концертном зале. Но опросить методом интервью хотя бы 300 человек уже затруднительно. Причина в том, что достоверность данных, как правило, обеспечивается, если интервьюер в течение дня опрашивает не более 5 человек, и опрос 300 человек в течение одного дня потребует одновременного привлечения к работе 60 интервьюеров. Конечно, можно «растянуть» интервью на несколько дней, но при этом будет потеряна оперативность исследования.

Еще сложнее проводить сплошной опрос, если предметом исследования выступает изучение эффективности избирательной кампании в городе, районе, области, когда речь идет о многих тысячах человек. При этом соответственно возрастает число работников, занятых сбором информации, проверкой годности анкет, кодировкой открытых вопросов и т.д. Другими словами, чем больше анкет, тем, естественно, больше времени требуется на их обработку, больше тратится средств и энергии на исследование. Поэтому в тех случаях, когда объект исследования насчитывает 500 человек и более, единственно правильным следует признать применение *выборочного метода*, теория которого заимствована из математической статистики.

Этот метод довольно широко практиковался в России еще во времена Петра Великого для оценки ожидаемой урожайности злаковых; использовался он и земскими статистиками, проводившими перепись населения.

Процесс выборки основан, в о - п е р в ы х, на взаимосвязи и взаимообусловленности качественных характеристик и признаков социальных объектов, в о - в т о р ы х, на правомерности выводов о целом на основании изучения его части при условии, что по своей структуре эта часть представляет собой микромодель целого.

Выборка – это по сути микромодель, которая является одним из наиболее экономных средств для проверки предположений или гипотез о свойствах предметов, явлений.

Н а п р и м е р, чтобы сделать заключение о вкусе яблок, снятых с дерева, достаточно попробовать одно или два яблока. Эта «достаточность» предопределена нашей уверенностью в том, что качественные характеристики яблок, собранных с одного дерева, совпадают. Но как быть, когда надо сделать заключе-

ние о вкусе яблок, собранных, скажем, со 100 деревьев? Обязательно ли для этого пробовать яблоки с каждого дерева? На помощь здесь приходит знание связи внешних признаков (формы, цвета) яблок с их вкусовыми качествами, позволяющее сделать заключение, что все яблоки одинаковой формы и цвета обладают одновременно и идентичным вкусом.

Предположим далее, что упомянутые признаки (форма и цвет) имеют свои варианты. В таком случае выборочная совокупность должна включать большее число элементов, т.е. «пробе» подлежит больше яблок. Представим, что перед нами корзина с яблоками трех форм (мелкие, крупные, грушевидные). Яблоки каждой формы в свою очередь могут быть трех цветов (красные, желтые, зеленые). Значит, корзина содержит $3 \cdot 3 = 9$ сортов яблок и, чтобы сделать верный вывод о каждой разновидности, нужно попробовать не менее 9 яблок.

Из сказанного следует, что достоверные выводы о качестве объекта должны быть основаны на тесно связанных с этим качеством признаках. Поэтому отбор элементов в выборочную совокупность может быть осуществлен не только по изучаемой качественной характеристике объекта (в приведенном примере — вкусу), но и по тесно связанным с ней признакам (цвету и форме). Чем больше таких контрольных признаков, тем больше потребуется отобрать элементов.

При рассмотрении основных методов выборки, применяемых в исследовании социальных проблем, используются такие ключевые понятия, как объект исследования, генеральная совокупность, выборочная совокупность, единица отбора, единица наблюдения.

Объектом социологического исследования выступает носитель той или иной социальной проблемы.

Генеральной совокупностью называется объект исследования, который локализован территориально, во времени, по демографическим или социальным признакам и на который распространяются выводы исследования.

Локализация объекта территориально чаще всего происходит по административному делению — регион, область, район, поселение. Этот критерий лежит в основе локализации объекта в электоральных исследованиях. Локализация по демографи-

ческим признакам осуществляется в исследованиях отдельных групп, например молодежи, женщин, пенсионеров, этносов; по социальным признакам – в целевых исследованиях профессиональных групп, конфессий, политических движений и др. Локализация объекта во времени осуществляется в длительных социальных экспериментах, контент-анализе средств массовой коммуникации или изучении их аудитории, при проведении повторных исследований.

Выборочную совокупность составляет определенное число элементов генеральной совокупности, отобранных по строго заданному правилу. В момент исследования она должна представлять собой как бы микромодель генеральной совокупности. Другими словами, необходимо, чтобы структура выборочной совокупности максимально совпадала со структурой генеральной совокупности по основным изучаемым качественным характеристикам и контрольным признакам. Чтобы добиться этого, нужно строго соблюдать правила выборки.

Единицы наблюдения есть элементы выборочной совокупности (респонденты), подлежащие изучению (например, опросу). Такими единицами могут выступать отдельные индивиды и целые группы (семья, посетители кинотеатра и т.д.).

Правила формирования выборочной совокупности таковы, что в процессе отбора основными элементами не всегда выступают единицы наблюдения, т.е. непосредственно опрашиваемые. Так, вначале могут быть отобраны те или иные административные регионы (области, края, республики), потом в них – города, в последних – семьи, в которых опрашиваются либо все взрослые члены, либо один член семьи, отобранный по заданному принципу (глава, «распорядитель» бюджета, старший ребенок и т.д.). Элементы (регионы, поселения, семьи, респонденты), отбираемые на каждом этапе выборки по особому плану, называются *единицами отбора*.

14.2. Виды и методы выборки

Различают выборки вероятностные и фокусированные.

Модель *вероятностной выборки* связана с понятием статистической вероятности. В самом общем случае вероятность не-

которого ожидаемого события есть отношение числа ожидаемых событий к числу всех возможных. При этом общее число событий должно быть достаточно большим (статистически значимым; числовой порог статистической значимости предопределен известным из математической статистики законом больших чисел). Так, замечено, что при подбрасывании куба, скажем, 600 раз (общее число событий) «шестерка» (ожидаемое событие) выпадает примерно в 100 случаях. Вероятность того, что куб упадет «шестеркой» вверх, определяется по формуле:

$$P = \frac{100}{600} = \frac{1}{6},$$

где P – вероятность ожидаемого события.

Так же определяется вероятность выпадения любой другой стороны куба. Полная вероятность (того, что куб обязательно выпадет одной из любых сторон) в свою очередь состоит из суммы вероятностей всех событий. В нашем примере полная вероятность равна:

$$P = P_1 + P_2 + P_3 + P_4 + P_5 + P_6 = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{6}{6} = 1.$$

Итак, вероятность того, что любое из событий обязательно произойдет, всегда равна единице, т.е. является *достоверным* утверждением. События, приведенные в нашем примере, именуются *равновероятными* (каждая сторона куба имеет равную вероятность выпадать) и обеспечиваются идентичностью условий: равномерным распределением вещества в кубе, недеформированностью его сторон, отсутствием резких колебаний воздуха при его бросании и т.д.

Создание условий равной вероятности отбора единиц наблюдения¹ играет ключевую роль. Оно должно гарантировать для каждого элемента генеральной совокупности равную вероятность попасть в выборочную. Такая ситуация имеет место лишь тогда, когда элементы генеральной совокупности распределены в ней равномерно.

В прикладной социологии используются следующие основные виды выборки: собственно случайная, механическая,

¹ Иногда эту процедуру называют *рендомизацией*.

серийная, гнездовая — основанные на вероятностных методах; квотная и стихийная — как фокусированные, «невероятностные» выборки; многоступенчатая, чаще всего имеющая вид территориальной выборки.

Методы *собственно случайной выборки* основаны на принципах «урновой модели». Чтобы лучше понять ее, обратимся к примеру. Построим «урновую модель», для чего предположим, что в ящике три вида шаров: 200 красных, 300 черных и 500 белых. Необходимо отобрать 100 шаров так, чтобы присутствовали пропорционально шары всех трех цветов. Их доли в ящике (в генеральной совокупности) составляют соответственно 20, 30 и 50 %. Чтобы сохранить те же пропорции должна быть отобрана выборочная совокупность, состоящая из 20 красных, 30 черных и 50 белых шаров. Это возможно при условии, что каждый шар в ящике будет иметь равную вероятность быть отобранным. Данное условие будет выполнено, если тщательно перемешать шары. После этого можно применить два метода отбора:

- ◇ случайный бесповторный, когда каждый отобранный шар откладывается в сторону, а из ящика берется следующий шар и т.д. до необходимой выборочной совокупности (всего 100 шаров);
- ◇ случайный повторный, когда записывается цвет каждого отобранного шара, а сам он возвращается в ящик, шары вновь перемешиваются и отбор повторяется.

Примерно такая же операция проводится и с элементами генеральной совокупности в социологическом исследовании. Все ее элементы (респонденты) пофамильно или посредством кода (числового номера) заносятся на карточки, после чего последние перемешиваются в ящике, из которого производится отбор одним из двух методов. При этом очень важно, чтобы карточки были перемешаны тщательно. Это повышает равную вероятность отбора респондентов.

Рассмотрим пример, построенный на условных обозначениях. Обозначим рабочих высокой квалификации шарами, средней квалификации — конусами, низкой квалификации — кубами. Равная вероятность любого из них быть отобранными обеспечивается лишь в том случае, если они расположены равномерно (рис. 14.1).

У обоих методов есть общий недостаток: их можно успешно применять только для генеральных совокупностей, насчитываю-

щих не более 800–1000 единиц. При увеличении числа единиц эти методы весьма трудоемки: нужно выписывать все фамилии или коды на карточки, равномерно их перемешивать и т.д.

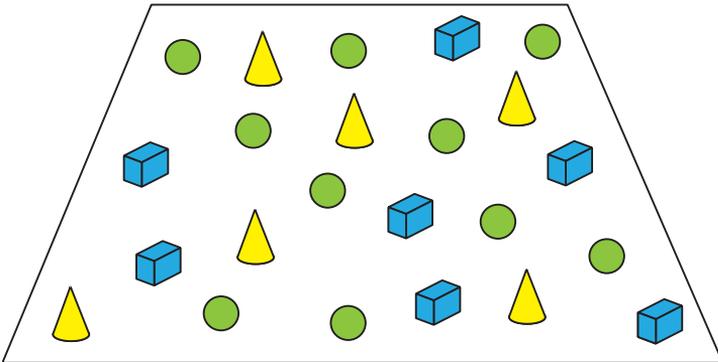


Рис. 14.1. Аналоговая модель равновероятного распределения единиц отбора

Для больших генеральных совокупностей удобнее применять метод *механической выборки*, общий принцип которой заключается в том, что все элементы генеральной совокупности сводятся в единый список и из него через равные интервалы отбирается соответствующее число респондентов. Шаг отбора K рассчитывается по формуле:

$$K = N / n,$$

где N – величина генеральной совокупности; n – величина выборочной совокупности.

Например, пусть $N = 2000$ человек, а $n = 200$ человек. Тогда $K = 2000 / 200 = 10$.

Это означает, что из списка должна быть отобрана каждая десятая единица. Отбор может отсчитываться с начала или с конца алфавитного списка. В качестве таковых для отбора респондентов успешно могут быть использованы списки избирателей или записи в домовых книгах (если опрос проводится в масштабах города, села), бланки учета в отделах кадров предприятий и учреждений, в общественных организациях, перечень статей или номеров газет в случае проведения контент-анализа и т.д.

Метод *стратифицированной* выборки довольно удобный и точный. Его суть в следующем. Если имеется возможность

«разбить» генеральную совокупность на однородные части (страты) по заданному признаку, то отбор респондентов может быть осуществлен из каждой страты отдельно (рис. 14.2). При этом число респондентов, отбираемых из страты, пропорционально общему числу элементов в ней.

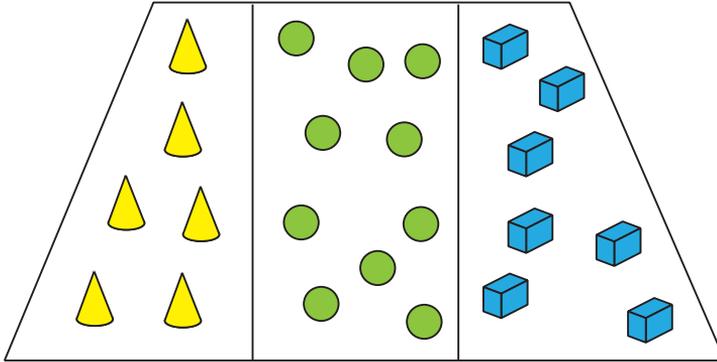


Рис. 14.2. Аналоговая модель стратифицированного распределения единиц отбора

На рис. 14.2 отображены три страты. Из каждой можно осуществить отбор единиц наблюдения при помощи собственно-случайной или механической выборки. Число респондентов n_i , подлежащих отбору из каждой страты в отдельности, определяется из соотношения:

$$n_i = \frac{N_i n}{N},$$

где i – число страт, выделенных в генеральной совокупности; N_i – общее число единиц в i -й страте.

Например, возьмем генеральную совокупность в 2000 человек, которая включает в себя 300 инженеров, 1000 рабочих и 700 служащих. Выборочная совокупность планируется в 200 человек. Следовательно, из каждой серии подлежат отбору и опросу:

$$n_1 = \frac{300 \cdot 200}{2000} = 30 \text{ человек}; \quad n_2 = \frac{1000 \cdot 200}{2000} = 100 \text{ человек};$$

$$n_3 = \frac{700 \cdot 200}{2000} = 70 \text{ человек}.$$

Общая величина выборочной совокупности равна:

$$n = n_1 + n_2 + n_3 = 200 \text{ человек.}$$

Метод *гнездовой выборки* довольно часто приходится применять на практике. Он предполагает, что в качестве единиц наблюдения будут отбираться не отдельные респонденты, а группы, с последующим сплошным опросом в отобранных группах.

Например, из 200 студенческих групп, в каждой из которых по 15 человек, должно быть отобрано на основе случайной выборки 30 групп. Тогда опросу подлежат 450 человек. Гнездовая выборка репрезентативна (представительна, или гарантирует распространимость результатов исследования на генеральную совокупность) в том случае, если составы групп в максимальной степени схожи по основным демографическим признакам респондентов (рис. 14.3).

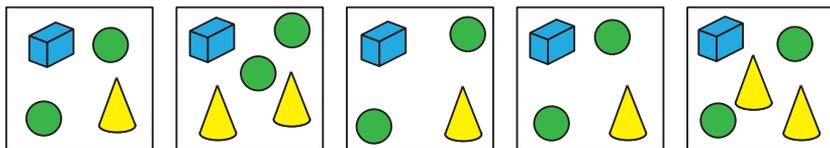


Рис. 14.3. Аналоговая модель гнездового распределения единиц отбора

В методе гнездовой выборки списки или карточки составляются только для групп (каждая из них представляет собой самостоятельную единицу отбора) путем их нумерации.

В качестве отбираемых для обследования могут выступать производственные бригады, группы студентов, школьников и другие общности, представляющие собой целостный объект с точки зрения социологического изучения той или иной проблемы.

Фокусированная выборка используется в социологических исследованиях наряду с вероятностной. К этой выборке неприменимы правила теории вероятности. Ее виды: стихийная выборка и метод квот.

Примером *стихийной выборки* служит почтовый опрос читателей журнала или газеты. В данном случае нельзя заранее определить структуру массива респондентов, возвративших анкеты, что, естественно, затрудняет оценку репрезентативности выборки. Поэтому выводы исследования, как правило,

распространяются лишь на опрошенную совокупность. Например, стихийной является выборка при опросе в магазине в течение дня покупателей какого-то одного вида товара, опрос зрителей спектакля и т.д.

Метод снежного кома — разновидность стихийной выборки. Например, если необходимо провести опрос 300 членов немногочисленной, но территориально чрезмерно рассредоточенной конфессии, а известны адреса только десяти ее членов, то поиск других респондентов можно продолжить по их подсказке. Предположим, каждый из десяти опрошенных сообщил адреса еще двоих своих единоверцев, а те в свою очередь — еще по два адреса и т.д. Тогда запланированный объем выборочной совокупности достигается на пятой ступени отбора. При этом отбор респондентов по этапам будет происходить в следующей пропорции: $10 + 20 + 40 + 80 + 160 = 310$ (опросить можно на 10 человек меньше, т.е. 300).

Метод основного массива применяется в разведывательных исследованиях для проверки какого-нибудь контрольного вопроса, скажем, удобно или неудобно назначенное активистами той или иной партии время массовой демонстрации. В таких случаях опрашивается 30–60 % потенциальных участников демонстрации.

Метод квотной выборки — наиболее точный, широко применяемый, в частности, в опросах общественного мнения. Его используют в тех случаях, когда до начала исследования имеются статистические данные о контрольных признаках элементов генеральной совокупности. Все данные о том или ином контрольном признаке выступают в качестве квоты, а их отдельные числовые значения — в качестве *параметров квот*. При квотной выборке респонденты отбираются интервьюерами целенаправленно, с соблюдением параметров квот. Число признаков, данные о которых выбираются в качестве квот, как правило, не превышает четырех. Дело в том, что при большем числе фиксированных признаков отбор респондентов становится чрезмерно трудоемким.

Квоты могут быть заданы как по *независимым*, так и по *взаимосвязанным* параметрам.

Например, пусть совпадение структуры выборочной и генеральной совокупностей обеспечивается по двум призна-

кам — уровню квалификации и уровню образования работников. Распределение респондентов по этим признакам в генеральной совокупности известно и выражено в процентах. Предположим далее, что опросу подлежат 200 человек. Тогда их состав при опросе формируется так, чтобы в процентном отношении структура распределения элементов выборочной совокупности по отмеченным двум признакам была тождественна структуре распределения элементов генеральной совокупности по этим же признакам. В табл. 14.1 и 14.2 отражено число подлежащих опросу респондентов в соответствии с параметрами квот по признакам «квалификация» и «образование».

Таблица 14.1. Независимые параметры квоты по признаку «квалификация»

Уровень квалификации	Данные о параметрах квоты в генеральной совокупности, %	Число человек, подлежащих отбору по соответствующему признаку	
		Абсолютное число	%
Высокий (5–6-й разряд)	50	100	50
Средний (3–4-й разряд)	35	70	35
Низкий (1–2-й разряд)	15	30	15
Итого	100,0	200	100,0

Таблица 14.2. Независимые параметры квоты по признаку «образование»

Образование	Данные о параметрах квоты в генеральной совокупности, %	Число человек, подлежащих отбору по соответствующему признаку	
		Абсолютное число	%
Начальное	2,5	5	2,5
Неполное среднее	22,5	45	22,5
Среднее	55,0	110	55,0
Незаконченное высшее	5,0	10	5,0
Высшее	15,0	30	15,0
Итого	100,0	200	100,0

Для квоты, заданной по взаимосвязанным параметрам, число подлежащих опросу дано в табл. 14.3.

Таблица 14.3. Взаимосвязанные параметры квоты по признаку «квалификация» и «образование»

Образование	Уровень квалификации							
	Высокий		Средний		Низкий		Итого	
	Состав генеральной совокупности, %	Требуется опросить, человек	Состав генеральной совокупности, %	Требуется опросить, человек	Состав генеральной совокупности, %	Требуется опросить, человек	Состав генеральной совокупности, %	Требуется опросить, человек
Начальное	–	–	–	–	2,5	5	2,5	5
Неполное среднее	2,5	5	12,5	25	7,5	15	22,5	45
Среднее	32,5	65	17,5	35	5,0	10	55,0	110
Незаконченное высшее	2,5	5	2,5	5	–	–	5,0	10
Высшее	12,5	25	2,5	5	–	–	15,0	30
Итого	50,0	100	35,0	70	15,0	30	100,0	200

По форме реализации различают выборки одноступенчатые и многоступенчатые. Методы, изложенные выше, представляют собой пример одноступенчатой выборки.

14.3. Многоступенчатые выборки

Приведенные модели вероятностной и фокусированной выборок эффективны для небольших генеральных совокупностей, включающих не более 1000 респондентов. В действительности объект исследования существенно больше. Так, в опросах общественного мнения речь идет о генеральной совокупности, включающей десятки миллионов человек. Кроме того, такой объект, как население страны, региона, города имеет широкое территориальное рассредоточение. В таких ситуациях невозможно применить единственную простую модель выборки. Приходится поэтапно использовать две, три и даже более одинаковых или различных моделей.

Выборку, основанную на поэтапном многократном использовании одной или нескольких моделей, называют *многоступенчатой*. Она осуществляется в несколько ступеней путем применения на каждой из них единого или разных методов. В последнем случае на первой ступени обычно реализуется гнездовая выборка (отбор поселений или административных районов, областей), а далее проводится случайный отбор респондентов в гнездах.

В прикладной социологии используют три вида такой выборки – территориальную, производственную и комбинированную (рис. 14.4).

Территориальная выборка предполагает поэтапный отбор административных объединений и поселений до отбора собственно респондентов. Эта процедура называется *районированием*. В основе районирования могут лежать географические, урбанистические, этнические, экономические, социальные и другие признаки.

Географические признаки играют важную роль на этапе первоначального районирования, в частности, при исследовании проблем экологии. Критериями выделения территорий здесь могут быть ландшафт (горы, леса, лесостепь, степь), климатические зоны (север, центр, юг) и т.д. В выделенных на ос-

нове перечисленных признаков регионах на втором этапе отбираются поселения, на третьем этапе происходит отбор семей, а в семьях – респондентов. В результате имеем четырех-ступенчатую выборку.

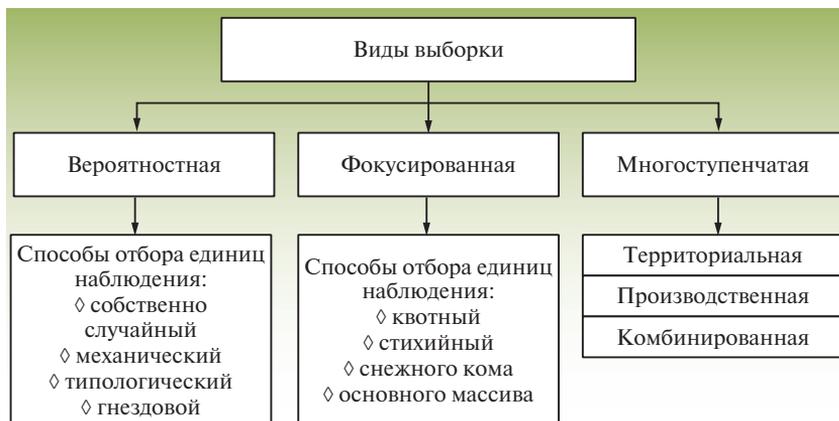


Рис. 14.4. Виды моделей выборки

При использовании урбанистического признака на первой ступени выборки составляются типологические группы городов и иных поселений. В основе группирования могут лежать: типология поселений (мегаполисы, областные, районные центры, рабочие поселки, села и деревни, хутора); численность жителей поселений независимо от их типа (до 1000 человек, от 1001 до 5000 человек, от 5001 до 10000 человек и т.д.); административный статус поселения; расстояние до административного центра и др.

Районирование по этническому признаку предполагает вначале выделение групп территорий и отдельных поселений, где доминирует тот или иной этнос. В Российской Федерации многие этнические общности имеют государственное устройство. Поэтому на первой ступени выборки могут быть сформированы три типологические группы: в первую войдут республики, во вторую – иные этнические территориальные образования (автономные области, округа), в третью – самостоятельные этнические поселения в составе «иноэтнических» регионов. На второй ступени стоит осуществить типологию поселений по

числу жителей, после чего отбор семей и соответственно респондентов.

В качестве экономического критерия районирования могут выступать показатели доминирующих отраслей (аграрная, энергетическая, добывающая, перерабатывающая); уровень дохода, уровень занятости населения; интенсивность миграционных потоков трудовых ресурсов и др.

В основе районирования территорий по социальным критериям могут лежать признаки конфессиональной принадлежности (регионы, населенные преимущественно теми, кто исповедует ислам, буддизм, православие), политических пристрастий (приверженности тем или иным политическим партиям и движениям), зон активных социальных протестов и др.

Пример одной из возможных моделей общероссийской многоступенчатой территориальной выборки приведен далее.

Первая ступень. В качестве массива выступают 11 территориально-экономических районов Российской Федерации, а также Москва и Санкт-Петербург: Северо-Западный (включая Калининградскую обл.), Северный, Волго-Вятский, Центральный, Центрально-Черноземный, Поволжский, Северо-Кавказский, Уральский, Западно-Сибирский, Восточно-Сибирский и Дальневосточный районы. Эти территориально-экономические районы образованы на основе доминирующих в них отраслей экономики. Территориально-административные образования (области, края, республики), которые они включают, по профилю экономики и социально-профессиональному составу населения идентичны.

Вторая ступень. В территориально-экономических районах на основе случайной бесповторной выборки или консультаций с экспертами осуществляется отбор одной административной единицы (области, края). Если в регионе имеются национальные республики, из их состава отбор объекта следует производить отдельно.

Третья ступень. Составляются списки всех поселений области (края, республики), кроме административного центра, которые распределяются в **страты** в соответствии с численностью жителей. В стратах поселения располагаются в алфавитном порядке, по первой букве названия. Из каждой страты при помощи механической выборки отбираются 1–3 поселения в

зависимости от приходящегося на область (край, республику) объема выборочной совокупности. Целесообразно, чтобы на одно отобранное поселение приходилось не менее 10 респондентов.

Четвертая ступень. Осуществляется отбор семей либо путем механического отбора из списка избирателей, либо с помощью трехступенчатого механического отбора. В последнем случае вначале из списка случайным образом отбираются улицы, потом из другого списка — дома на отобранных улицах и, наконец, респонденты из квартир в отобранных домах.

Пятая ступень заключается в отборе респондентов в семьях. В зависимости от цели исследования это могут быть: «распорядитель» семейного бюджета при изучении инвестиционных установок; член семьи в возрасте 18 лет и старше при изучении электоральных установок; мужчины при изучении спроса, скажем, на бритвы, и т.д.

Если в семье несколько членов с идентичными характеристиками, из них можно выбрать респондента при помощи таблицы случайных чисел или случайной бесповторной выборки. При условии, что таковых не более шести, целесообразно применить прием бросания кубика с шестью гранями.

Исследователь, имеющий статистические данные о ряде демографических и социальных признаков генеральной совокупности, на последней ступени территориальной выборки может применить квотный отбор респондентов. Для общероссийской выборки, используемой в опросах общественного мнения, это могут быть квоты по возрастному, половому и социально-профессиональному признакам в целом по населению территориально-экономических районов и мегаполисов.

Если определен объем выборочной совокупности¹, то для всех ступеней он рассчитывается пропорционально населению (или объекту исследования) соответственно в регионах и поселениях.

Производственная выборка используется в исследованиях проблем, связанных с трудовыми ресурсами или иными инсти-

¹ В практике современных всероссийских опросов общественного мнения используются выборочные совокупности объемом от 1600 до 3000 человек.

туционально объединенными группами населения. Например, темпы и причины сокращения численности рабочих, состояние их социальной защищенности на предприятиях различных отраслей экономики; условия **рекреационной** деятельности в учреждениях отдыха; информационная и техническая обеспеченность научных исследований в вузах страны и т.д.

В исследовании проблемы, связанной с деятельностью научных подразделений, на первой ступени научные учреждения объединяются в страты по научному профилю независимо от территориальной расположенности. Из этих страт механическим или собственно-случайным методом отбираются те (пропорционально их численности в каждой страте), в которых будет проводиться исследование.

На второй ступени в отобранных научных учреждениях объединяются в страты научные подразделения (на уровне отделов, секторов, лабораторий) с идентичным научным профилем. Внутри страт научные подразделения объединяются в 2–4 группы в соответствии с численностью сотрудников в них. После этого осуществляется пропорциональный отбор единиц в каждой группе, для чего вначале следует определить общую численность сотрудников в идентичных по профилю научных подразделениях отобранных научных учреждений, а потом – в группах с разной численностью сотрудников. Число отбираемых единиц в каждой группе определяется из процентного соотношения общей численности сотрудников в стратах.

Например, пусть в научных подразделениях гуманитарного профиля заняты 18,6 % сотрудников, технического – 61,9 % и естественного – 19,5 %. В такой же пропорции распределится выборочная совокупность респондентов. По численности сотрудников в подразделениях гуманитарного профиля имеем: 20 групп по 21–30 человек и 30 групп по 11–20 человек; технического профиля: 40 групп по 31–40 человек, 25 групп по 21–30 человек и 60 групп по 11–20 человек; естественного профиля: 25 групп по 21–30 человек и 10 групп по 11–20 человек.

Всего генеральная совокупность включает 100 групп численностью 11–20 человек в каждой, что в сумме составляет 1550 человек, или 32,6 %; 70 групп численностью 21–30 человек в каждой (в сумме 1785 человек) – 37,5 %, 40 групп численностью 31–40 человек в каждой (в сумме 1420 человек) – 29,9 %

генеральной совокупности. Если объем выборочной совокупности равен 1000 человек, то в 100 малочисленных группах отбору подлежат 326 человек, в средних – 375 и в крупных – 295. Из соотношения соответствующих величин в выборочной и генеральной совокупностях нетрудно рассчитать, что при сплошном опросе респондентов отбору во всех трех типах подлежит каждая пятая группа (табл. 14.4). Если исследователем будет принято решение опрашивать в каждой группе не более половины сотрудников, то шаг отбора будет равен двум-трем, чередуясь через каждую отобранную группу. В этом случае предстоит еще одна ступень отбора респондентов в группах с применением одной из вероятностных моделей выборки.

Таблица 14.4. Параметры генеральной совокупности

Профиль научных подразделений	Численность сотрудников в группах					
	11–20		21–30		31–40	
	Число групп	Число сотрудников (всего) в группах	Число групп	Число сотрудников (всего) в группах	Число групп	Число сотрудников (всего) в группах
Гуманитарный	30	465	20	510	–	–
Технический	60	930	25	638	40	1420
Естественный	10	155	25	638	–	–
Итого	100	1550	70	1786	40	1420

Комбинированная многоступенчатая выборка предполагает сочетание территориального и производственного принципов районирования.

14.4. Определение объема выборочной совокупности и обеспечение ее репрезентативности

Данные вопросы являются самыми важными во всей теории выборки. С одной стороны, объем выборочной совокупности должен быть статистически значим, т.е. достаточно большим для того, чтобы получить достоверную инфор-

мацию, а с другой — экономным, в некотором смысле оптимальным. Критерием оптимальности, как доказано математиками, являются числовые усредненные параметры контрольных признаков элементов генеральной совокупности, точнее, их *дисперсия* (разброс). Чем больше дисперсия, тем большим должен быть объем выборочной совокупности.

Например, осуществляется отбор из генеральной совокупности 2000 человек; при этом контролируется состав выборочной совокупности по признаку «пол»: 70 % мужчин и 30 % женщин. Согласно теории вероятности, можно предположить, что среди каждых 10 отбираемых респондентов встретятся три женщины. Если надо опросить по крайней мере 90 женщин, то, исходя из упомянутого выше соотношения, необходимо отобрать не менее 300 человек. Теперь предположим, что в генеральной совокупности 90 % мужчин и 10 % женщин. Чтобы в этом случае в выборочную совокупность попало 90 женщин, необходимо отобрать уже не менее 900 человек. Из примера видно, насколько велико влияние дисперсии (разброса признака) на объем выборочной совокупности. Поэтому его следует вычислять по признаку, дисперсия значений которого наибольшая.

Когда информация о признаках элементов генеральной совокупности отсутствует, исключается возможность определения объема выборочной совокупности при помощи математических формул. В таких случаях можно опираться на многолетний опыт социологов-практиков, свидетельствующий о том, что для пробных опросов достаточна выборочная совокупность объемом 100–250 человек. При массовых опросах, когда объем генеральной совокупности менее 5000 человек, достаточный объем выборочной совокупности — не менее 500 человек; если объем генеральной совокупности превышает 5000 человек, то объем выборочной совокупности 10 % ее состава, но не более 2000–2500 человек. Это гарантирует весьма достоверные результаты исследования. Объемы выборочной совокупности, превышающие 3000 респондентов, используются лишь в комплексных исследованиях, а также при сложной структуре генеральной совокупности.

Определение статистически значимого объема выборочной совокупности — важная, но не достаточная предпосылка пра-

вомерности распространения выводов исследования на всю генеральную совокупность. Дело в том, что из одной и той же генеральной совокупности можно отобрать большое число выборочных совокупностей. Например, даже если мы контролируем совпадение структур выборочной и генеральной совокупностей лишь по полу (включает всего два значения), то из совокупности в 1000 человек, из которых 40 % женщин и 60 % мужчин, и желательном объеме выборочной совокупности, равном 100 человек, можно сформировать не менее 10 таких выборочных совокупностей, где не встретится дважды ни один респондент, а соотношение доли женщин и мужчин в них будет соответственно 40 и 60 %. Можно составить сотни выборочных совокупностей, в которых будет соблюдаться указанное соотношение по признаку «пол» с определенной степенью приближенности.

Из всего многообразия возможных выборочных совокупностей необходимо отобрать одну, наиболее *точную*, т.е. максимально совпадающую по структуре и параметрам контролируемого признака с генеральной совокупностью. Для нашего примера — такую, в которой соотношение представителей двух полов было бы максимально близким их соотношению в генеральной совокупности (соответственно 40 и 60 %). Выборочная совокупность считается *репрезентативной*, если отклонение не превышает в среднем 5 % (в выборочной совокупности соотношение названных величин может составлять, например, 37 и 63 % или 42 и 58 %). Понятие «репрезентативный» заимствовано из математической статистики и символизирует степень совпадения, определенную по итогам измерения, среднего значения признака в выборочной совокупности с его средним значением в генеральной совокупности. Отклонение выборочного значения признака от его истинного значения в генеральной совокупности называют погрешностью или *ошибкой выборки*.

Различают два вида ошибок выборки — случайную и систематическую. Обе эти погрешности связаны с репрезентативностью выборки. Погрешность выборки носит *случайный* характер, если соблюдены все правила отбора единиц наблюдения. Она вычислима только для вероятностных моделей выборки.

Систематические ошибки нарушают точность выборочной совокупности. Дело в том, что репрезентативная выборочная

совокупность (т.е. такая, в которой среднее числовое значение признака отличается менее чем на 5 % от его среднего числового значения в генеральной совокупности) может оказаться неточной. Выявить эту неточность путем вычисления степени репрезентативности невозможно, так как случайная ошибка может иметь одну и ту же величину при разных числовых значениях признака (когда в выборочной и генеральной совокупностях состав респондентов с признаками А и Б соответственно 40 и 60 % и когда в выборочной совокупности по причинам ошибок смещения 60 % респондентов с признаком А и 40 % с признаком Б).

Как правило, систематические погрешности носят субъективный характер и могут присутствовать при реализации и вероятностных, и фокусированных моделей выборки.

Систематические ошибки бывают следствием:

- ◇ неверных исходных статистических данных о параметрах контрольных признаков генеральной совокупности (если используется квотная выборка);
- ◇ слишком малого (статистически незначимого) объема выборочной совокупности;
- ◇ неверного применения способа отбора единиц наблюдения (например, отбор из неверно составленного списка, неудачный выбор места и времени проведения опроса).

Н а п р и м е р, при опросах электората о готовности участвовать или не участвовать в выборах образуется группа «еще не принявших решение». Среди них могут быть и такие, кто просто скрывает свою позицию; так, в 1995 г. накануне выборов в Государственную Думу многие пенсионеры скрывали, что будут голосовать за коммунистов. В итоге большое число таких «колеблющихся» внесет серьезную погрешность в оценку электорального поведения.

При вероятностном отборе семей по механической выборке может получиться так (выбран такой шаг отбора), что каждая отобранная семья проживает в квартире на первом или последнем этажах. В этом случае в выборке будут систематически накапливаться ошибки по социальному статусу респондентов, так как известно, что квартиры на этих этажах многоэтажных домов обычно «достаются» населению с низким социальным статусом.

Так как определить величину систематических ошибок при помощи математических формул практически невозможно, они автоматически переходят на результаты и выводы исследования. Это еще раз свидетельствует об исключительной важности неукоснительного соблюдения правил отбора единиц анализа. Избежать систематических ошибок помогают также точные предварительные сведения о структуре генеральной совокупности. В качестве источников информации могут выступать: документы ведомственного учета, данные опроса экспертов, сведения органов государственной статистики, результаты разведывательного исследования, публикации об итогах проводившихся ранее исследований того же объекта.

14.5. Выборка в повторных исследованиях

Чтобы обеспечить получение репрезентативной информации об объекте исследования на различных этапах его развития в течение определенного, порой достаточно длительного отрезка времени, выборка в повторном исследовании должна быть универсальной, достоверно отражающей состояние изучаемого объекта в каждой его временной точке.

Характер выборки зависит от задач повторного исследования, поэтому не обязательно, чтобы в первоначальном и повторных опросах она проводилась по единой схеме. Главное, чтобы применяемый способ обеспечивал репрезентативность и сравнимость данных. Желательно строить выборку так, чтобы она была репрезентативна для объекта в момент опроса. Как правило, в повторном исследовании необходимо обеспечить сопоставимость данных. Принцип сопоставимости данных предполагает сохранение идентичности выборочной совокупности по основным параметрам (последнее важно для построения прогностических моделей). Поэтому независимо от того, какой метод выборки применялся первоначально, при повторном опросе целесообразно осуществить квотную выборку, взяв в качестве параметров квот числовые значения контролируемых признаков выборочной совокупности первоначального опроса.

В панельных исследованиях совокупность респондентов на каждом этапе сохраняется, однако по разным причинам она

может со временем уменьшиться. Например, в изучении жизненных планов выпускников средних школ первоначальная группа, опрошенная накануне выпускных экзаменов, по причине осложнений в поиске респондентов может не совпасть с группой, подлежащей опросу через год. В этом случае для обеспечения сопоставимости данных исследования, полученных на двух массивах, потребуются коррекция (как правило, уменьшение) первоначальной выборочной совокупности — в ней сохранятся только те респонденты, кого удалось найти при повторном исследовании.

Имеются и более сложные методы выборки, при разработке модели которых прибегают к помощи специального математического аппарата и компьютера. Поэтому при конструировании модели выборки целесообразно консультироваться у специалиста по математической статистике.

Основной математический аппарат, применяемый для вычисления выборочной совокупности и средней случайной ошибки выборки, приведен в приложении 7.

ВОПРОСЫ ДЛЯ САМОКОНТРОЛЯ

1. Чем вызвана необходимость применения выборочного метода в прикладной социологии?
2. Чем различаются вероятностные и невероятностные модели выборки? Перечислите их разновидности.
3. По каким правилам формируется квотная выборка?
4. По каким правилам формируется территориальная выборка?
5. Как определяется оптимальный объем выборочной совокупности?
6. Что представляет собой репрезентативность выборочных данных?
7. Какие виды погрешностей возможны при формировании выборки и чем они вызваны?

ПРЕДЛАГАЕМ ВЕРНУТЬСЯ В МЕНЮ И РЕШИТЬ ЗАДАЧИ К ЭТОЙ ГЛАВЕ