

В. Н. КРЫЛОВ

ВЫБОРОЧНЫЙ
МЕТОД
В СТАТИСТИКЕ



ГОССТАТИЗДАТ
МОСКВА ~ 1957

В. Н. КРЫЛОВ

ВЫБОРОЧНЫЙ МЕТОД В СТАТИСТИКЕ

ГОСУДАРСТВЕННОЕ СТАТИСТИЧЕСКОЕ ИЗДАТЕЛЬСТВО

МОСКВА

1957

Выборочный метод, широко применяемый в статистике, сокращает объем статистической работы за счет уменьшения количества обследуемых единиц, дает возможность более детально изучить каждую из них в отдельности и нередко получать выводы более глубокие, чем достигаемые при сплошном наблюдении. Кроме того, выборки обеспечивают получение информации о таких явлениях, сплошное наблюдение которых невозможно или нецелесообразно.

Измерение урожайности и валового урожая сельскохозяйственных культур, изучение структуры розничного товарооборота, спроса покупателей в государственной и кооперативной торговле, привозов, цен и реализации на колхозных рынках, исследование бюджетов семей трудящихся и ряд других вопросов хозяйственного и культурного строительства — во всех этих отраслях выборочный метод используется с большой эффективностью.

Выборочный метод применяется также в промышленном производстве для организации предупредительного контроля качества продукции в ходе самого производства, при приемке готовой продукции, при исследовании отдельных факторов, влияющих на ее качество, и т. д.

Наиболее успешное применение выборочного метода требует знания методов отбора, умения производить расчеты статистических показателей и т. д.

В данной брошюре содержится изложение сущности выборочного метода и рассматриваются отдельные области его применения.

ГЛАВА I

СУЩНОСТЬ ВЫБОРОЧНОГО МЕТОДА

1. Несплошное наблюдение и его виды

Основным этапом, с которого начинается статистическое исследование, является наблюдение, собирание статистических данных для последующей обработки и анализа.

Статистическое наблюдение подразделяется на сплошное массовое наблюдение, т. е. обследование всех без исключения единиц изучаемой совокупности, и несплошное (частичное) наблюдение, при котором обследуется только некоторая часть совокупности.

К несплошному наблюдению относятся: 1) монографический метод, заключающийся в детальном описании единичных объектов, 2) метод основного массива, при котором наблюдению подвергается часть совокупности, имеющая высокий удельный вес во всей совокупности, 3) анкетный метод — рассылка анкет (вопросников) всем лицам, от которых желают получить интересные исследователей сведения, и, наконец, 4) выборочный метод, содержание которого сводится к обследованию некоторой, специально отобранной части изучаемых единиц для получения данных, характеризующих всю совокупность.

Сплошное наблюдение обеспечивает наиболее полную информацию об общем количестве единиц, относительных, средних и других сводных показателях изучаемой совокупности. Однако бывают случаи, когда сплошное наблюдение невозможно или нецелесообразно. Так, например, невозможно организовать сплошное углубленное изучение бюджетов семей трудящихся с ежедневными записями их доходов, расходов, потребления и т. д. Для этого необходимо такое большое число статистиков по бюджетам, какое практически выделить невозможно: один статистик в состоянии вести такое наблюдение не более чем в 20—25 семьях одновременно и для обследования бюджетов всех трудящихся СССР понадобилось бы не менее 3 млн. статистиков. В таких случаях средством получения необходимых данных остается несплошное наблюдение.

По части совокупности не всегда можно получить более или менее точное представление о целом, и отдельные виды несплош-

ного наблюдения различаются главным образом тем, насколько их результаты дают возможность судить обо всей совокупности изучаемых явлений. Рассмотрим с этой точки зрения перечисленные выше методы сплошного наблюдения.

Монографический метод заключается в подробном описании и глубоком изучении отдельного, специально отобранного для этой цели объекта. Отрицать монографический метод, мотивируя это тем, что задачей статистики является характеристика совокупностей, а не отдельных явлений, в корне неправильно. Ф. Энгельс указывал, что единица (Einheit) и множественность (Vielheit) являются неразделимыми понятиями и что множественность так же содержится в единице, как единица во множественности. Статистическая совокупность познается через отдельные индивидуумы так же, как индивидуумы через совокупность. При сплошном наблюдении внимание исследователя сосредоточивается на том, чтобы охватить всю массу изучаемых явлений по сравнительно ограниченному числу наиболее существенных признаков и рассмотреть взаимосвязи этих признаков внутри всей совокупности, выделив то общее, существенное, что присуще всей массе в целом. При монографическом наблюдении, наоборот, основное внимание направляется на детальнейшее изучение всех сторон, всех черт одной наблюдаемой единицы. Монографическое наблюдение является необходимым дополнением к сплошному и в сочетании с последним обеспечивает всесторонний анализ изучаемых явлений.

Монографическое наблюдение производится в целях выявления тенденции развития прогрессивных явлений и распространения передового опыта. Монографическое наблюдение помогает вскрыть неиспользованные резервы. Это достигается монографическим описанием передовых заводов, колхозов, совхозов, МТС и опыта отдельных передовиков производства. Таким образом, монографический метод является весьма ценным дополнением к сплошному наблюдению, так как обеспечивает получение данных, необходимых для корректирования выводов последнего.

Метод основного массива — несовершенный сплошной метод или, как его иногда называют, группировочно-цензовый метод — заключается в том, что наблюдению подвергается часть единиц, сравнительно незначительная по численности, но имеющая преобладающий удельный вес в совокупности. Остальными единицами совокупности, несмотря на их многочисленность, пренебрегают как несущественными. Так, данные годового обследования цензовой промышленности СССР за 1925/26 г. показывают высокую концентрацию последней: 475 предприятий, составлявших только 5% общего их числа, имели 70% основных фондов всей цензовой промышленности. Не менее высокой была и концентрация грузооборота на железнодорожном транспорте: 350 железнодорожных станций, или 10% общего их числа, охватывали 55—65% грузооборота по 30 важнейшим товарам.

Поэтому можно было характеризовать промышленность по небольшому числу предприятий, которые, однако, по количеству рабочих, продукции и основных фондов составляли подавляющую массу и могли определять структуру и динамику всей государственной промышленности.

Для иллюстрации репрезентативности такой выборки приводится таблица, где дается динамика выработки и номинальной заработной платы за 1 человеко-день в хлопчатобумажной промышленности СССР в 1927/28 г.: а) по всей отрасли, б) по нескольким предприятиям (выборка) ¹.

Кварталы	В руб.				В процентах к I кварталу			
	средняя выработка		номинальная зарплата		средняя выработка		номинальная зарплата	
	а	б	а	б	а	б	а	б
I	17,21	17,67	2,273	2,346	100,0	100,0	100,0	100,0
II	17,88	18,26	3,367	2,423	103,0	103,3	104,1	103,3
III	18,07	18,34	2,406	2,463	105,0	103,8	105,0	105,0
IV	18,48	18,96	2,448	2,510	107,4	107,0	107,7	108,8

Метод основного массива применялся в нашей промышленной и транспортной статистике для изучения динамики отдельных показателей, однако большого распространения он не получил. Репрезентативность такой выборки целиком зависит от удельного веса отбираемой части по изучаемому признаку во всей совокупности, но удельный вес может непрерывно меняться.

Разновидностью метода основного массива являются способы построения некоторых индексов. Например, при изучении реализации и цен на колхозном рынке материалы собираются не по всем товарам, а по группе товаров, имеющих наибольший удельный вес в колхозной торговле.

Анкетный метод, или метод добровольной анкеты, иногда называют корреспондентским, так как сведения поступают по почте и связь между исследователями и лицами, дающими сведения, осуществляется в форме корреспонденции: В качестве корреспондентов, от которых запрашиваются сведения, часто избираются не сами обследуемые, а лица, сведущие в изучаемом вопросе. В нашей дореволюционной статистике и в первые годы Советской власти сведения об урожае собирались при помощи так называемых добровольных корреспондентов: учителей сельских школ, агрономов и грамотных крестьян.

¹ См. М. Смит и др., К теории выборочного метода, „Вестник статистики“ № 3—4, 1929, стр. 148.

При анкетном методе результаты обследования могут искажаться, когда программа его затрагивает интересы опрашиваемых. Последние в своих ответах могут давать такую окраску сообщаемым сведениям, которая им наиболее выгодна, или умалчивать о невыгодных им фактах. Кроме того, ответы присылают лишь те, кто заинтересован в собирании запрашиваемых сведений, и не присылают те, кто заинтересован в том, чтобы их скрыть. Поэтому анкета дает нерепрезентативные результаты. Из этого некоторые статистики делают вывод о полной непригодности анкеты в качестве средства статистических обследований.

Наиболее совершенным видом несплошного наблюдения является выборочный метод. Выборочное исследование организуется так, чтобы по отобранной части совокупности можно было судить не только о том, из каких единиц и с какими признаками состоит вся совокупность, но также и о количественных соотношениях этих единиц, о средних величинах и других сводных статистических показателях. Отобранная часть совокупности, выборка, является представителем всей совокупности, представляет ее, причем принимаются специальные меры, чтобы отклонения выборочных характеристик от характеристик изучаемой совокупности, называемые ошибками выборки или ошибками репрезентативности, по возможности были ничтожно малыми, несущественными.

2. Принципы отбора при выборочном методе

В основе выборочного метода лежат, с одной стороны, принципы планомерно организованного отбора — принцип дифференциации изучаемых явлений и принцип пропорционального представительства отдельных типических групп — и, с другой стороны, принципы случайного отбора — принцип независимости способа отбора от изучаемых признаков и принцип равной вероятности для всех единиц совокупности оказаться включенными в состав выборки. На принципах планомерно организованного отбора основана, например, механическая выборка, при которой отдельные единицы совокупности отбираются через равные промежутки, равномерно по всей массе изучаемых явлений. На принципах случайного отбора основана собственно случайная выборка, повторная и бесповторная, заключающаяся в отборе единиц по жребию, наудачу. Промежуточное место занимает так называемая типическая районированная или расслоенная случайная выборка, при которой изучаемая совокупность разбивается на ряд более или менее однородных районов (групп); внутри каждого из этих районов производится случайный отбор, чаще пропорциональный численностям районов, а иногда и не пропорциональный.

Выборочный метод является научно обоснованным приемом статистического исследования, обеспечивающим с достаточной точностью и надежностью изучение общественных явлений на основе обследования не всей их массы, а лишь некоторой части, надлежащим образом отобранной.

Чем подробнее и глубже программа наблюдения вникает во все особенности изучаемых явлений, чем больше признаков она охватывает, чем детальнее исследование каждой отдельной единицы, тем меньше должен быть объем наблюдения. Это требует приемов исследования, обеспечивающих достаточно точное представление обо всей совокупности путем обследования только некоторой ее части.

Основанием для применения выборочного метода служит диалектическое единство отдельного, частного и общего. По отдельному можно судить об общем, по части можно судить о целом, если, конечно, правильно уловлена связь между отдельным и общим, если методика исследования разработана с учетом этой связи. Кроме того, относительная качественная однородность совокупностей дает возможность разбивать их на типические группы и выделять одну или несколько частей, каждая из которых будет представлять совокупность, т. е. обладать характеристикой всей совокупности.

Выборочное исследование, как и всякое частичное наблюдение, естественно, вынуждено отказываться от задачи учета общей численности изучаемой массы явлений, которую может решить только сплошное исследование. Его возможности ограничены получением относительных, средних и других сводных характеристик. Однако в сочетании с данными сплошного наблюдения оно дает возможность исчислить и общий объем совокупности. Так, определив выборочным путем средний урожай в центнерах с гектара и зная на основании сплошного наблюдения общую площадь изучаемой культуры, нетрудно исчислить и общий объем урожая последней.

Из изложенного следует, что выборочным является только такое несплошное наблюдение, при котором наблюдаемая часть фактов становится подлинным представителем всей их совокупности. Выборка репрезентирует совокупность, т. е. отобранная часть совокупности в уменьшенном масштабе воспроизводит всю совокупность. В отобранной части все изучаемые признаки должны находиться в тех же соотношениях, иметь такие же доли, средние значения и т. д., какие имеются во всей совокупности.

В практике выборочного исследования неизбежно возникают погрешности или ошибки, заключающиеся в том, что отобранная часть не вполне точно воспроизводит характеристики целого. Размеры этих ошибок зависят прежде всего от применяемого способа отбора, а также от объема выборки. В связи с этим перед теорией выборочного метода встает вопрос об оценке точ-

ности того или иного способа отбора и конкретных результатов произведенной выборки, откуда вытекают две основные задачи. Первая требует решения таких вопросов: известны данные обо всей совокупности; производится выборка таким-то способом; каковы будут результаты, т. е. какими свойствами будет обладать выборочная совокупность. Будут ли эти свойства и характеристики выборочной совокупности отличаться от известных нам свойств и характеристик всей совокупности? Насколько будут велики отклонения (ошибки выборки)? Какими условиями они определяются? Могут ли они быть уменьшены и насколько? Решение этой задачи должно основываться на всестороннем анализе свойств изучаемой совокупности и учете особенностей применяемого способа отбора. Здесь важно подчеркнуть роль и значение выводов математики, в частности теории вероятностей, при оценке точности выборки. При этом исследователь рассматривает все возможные варианты выборок, сопоставляя результаты своего анализа с характеристиками заданной генеральной совокупности. Таковы так называемые «прямые» теоремы теории вероятностей: теорема Бернулли, теорема Муавра-Лапласа, теорема Чебышева и др.

Вторая задача теории выборок заключается в определении пределов, в которых находятся характеристики всей совокупности, если при применении какого-нибудь способа отбора получены определенные результаты. Эта задача принципиально отличается от предыдущей, хотя и связана с последней. При решении первой задачи, когда генеральная совокупность нам известна, всегда может быть построена схема всех возможных случаев выборки из этой совокупности. При решении второй задачи такую схему построить уже невозможно, и количественные характеристики генеральной совокупности не могут быть использованы для оценки получаемых выборочных характеристик.

Для оценки разности между выборочной и генеральной характеристиками «прямые» теоремы применимы, когда генеральная характеристика величина постоянная, а выборочная — случайная переменная. При решении второй задачи исходят из предпосылки, что неизвестная нам генеральная характеристика может иметь те или иные значения с соответствующими им вероятностями, т. е. она рассматривается в качестве случайной переменной. Поэтому для решения второй задачи теория вероятностей применяет так называемые «обратные» теоремы Бернулли, Лапласа, «второй закон больших чисел» и др. Они доказуемы только для выборок неограниченного объема и не применимы для оценки точности конечных выборок. Для последних большую роль играют так называемые безусловные вероятности того, что разность между выборочной и генеральной характеристиками окажется сколь угодно малой, и основанные на них «доверительные» вероятности, что генеральная характеристика

находится в тех или иных пределах. Количественное отличие безусловной и доверительной вероятностей от возможных значений «обратной» вероятности тем меньше, чем ближе к единице безусловная вероятность или чем больше объем выборки. Поэтому если вообще логическая ошибка, заключающаяся в смешении «прямой» и «обратной» вероятностей, и может привести к грубым количественным ошибкам, то в этих случаях они оказываются несущественными, и использование «обратных» вероятностей практически допустимым.

Когда говорят об ошибках или погрешностях отбора или, как их называют, ошибках репрезентативности, то имеют в виду не те ошибки, которые могут иметь место при всяком статистическом наблюдении и зависят от качества проведения последнего — точности измерений или подсчетов, правильности ответов опрашиваемых, добросовестности и квалификации регистраторов и т. д. Последние называются ошибками точности или ошибками регистрации. Ошибками репрезентативности («представительства») называются расхождения между статистическими характеристиками выборочной и генеральной совокупностей, возникающие от того, что первая неточно воспроизводит распределение признаков второй, недостаточно ее репрезентирует. Эти ошибки возникают только при выборочном обследовании в результате недостаточной точности отбора и не имеют места при сплошном наблюдении.

Переход от сплошного наблюдения к выборочному, сокращая объем наблюдения, уменьшает ошибки точности, но приводит к ошибкам репрезентативности. Понятно, что такой переход выгоден только в том случае, если ошибки отбора будут компенсированы уменьшением ошибок регистрации. На этом основано применение выборок для проверки и корректировки результатов сплошных наблюдений, проведенных методами, чреватými ошибками точности.

Ошибки отбора или погрешности выборочного обследования, как и всякого другого статистического наблюдения, делятся на случайные и систематические. Случайные ошибки относятся к несущественным, так как они базируются на законе ошибок Ляпунова — Лапласа — Гаусса, по которому отклонения (погрешности) в сторону преувеличения и в сторону преуменьшения встречаются одинаково часто и вследствие этого взаимно погашаются, не отражаясь или мало отражаясь на общих результатах обследования. Систематические ошибки, наоборот, заключаются в постоянных отклонениях в одну и ту же сторону и существенно отражаются на результатах исследования. Поэтому одной из основных задач теории выборок является разработка методов отбора, исключающих возможность появления систематических ошибок и ограничивающих случайные ошибки.

В теории статистики к выборочному методу нередко относили только такие виды несплошного наблюдения, которые ос-

нованы на принципах случайного отбора, поскольку этот отбор в силу действия закона больших чисел гарантирует невозможность систематических ошибок, и отбрасывали как «ненаучные» выработанные практикой приемы, основанные на принципах планомерно организованного отбора.

Исторически теория выборочного метода возникла не в порядке обобщения практики выборок в социально-экономических исследованиях, а была привнесена в статистику извне, из обобщения практики многократных измерений одних и тех же величин в астрономии и других науках в форме теории случайных ошибок измерений. Поэтому теория выборочного метода с самого начала приняла форму и развивалась исключительно как теория случайной выборки, основанная на положениях математической теории вероятностей. Характерно, что теорией выборочного метода занимались не экономисты-статистики, а математики, рассматривавшие ее как прикладную теорию вероятностей. Между тем планомерно организованный отбор, возникший из практики социально-экономических исследований и опирающийся на совершенно иную базу, чем случайная выборка, теорией статистики не рассматривался, несмотря на то, что он оказался очень плодотворным и выработал ряд весьма целесообразных и репрезентативных приемов исследования.

Практика применения выборок показывает, что принципы случайной выборки оказываются более применимыми в области изучения естественных явлений и в технике, поскольку там мы имеем дело с однородными массовыми совокупностями, настолько многочисленными, что практически их можно считать неограниченными. Что касается социально-экономических явлений, то вследствие резкой их дифференциации на типы и группы к ним более целесообразно применение планомерно организованного отбора. Как показывает опыт, планомерно организованный отбор может быть плодотворно применен и к изучению явлений природы, а также и в технике.

Как уже указывалось, случайная выборка заключается в отборе по жребию, обеспечивающем соблюдение основных принципов случайного отбора — принципа независимости отбора от изучаемых признаков и принципа равной для всех единиц генеральной совокупности вероятности оказаться включенными в состав выборки. Принцип независимости отбора от изучаемых признаков требует, чтобы последние не могли оказывать влияния на включение или невключение в выборку тех или иных единиц совокупности. Соблюдение этого принципа призвано обеспечить объективный характер отбора и избежать в нем тенденциозности. Принцип равной для всех единиц генеральной совокупности вероятности оказаться включенными в состав выборки также требует, чтобы отбор не давал предпочтения каким-либо из этих единиц. Этот принцип направлен против возможной тенденциозности отбора. Кроме того, соблюдение этого принципа

дает возможность применять теорию вероятностей к оценке результатов выборки.

Случайный отбор обеспечивает условия для действия закона больших чисел, согласно которому в массе случайностей проявляется общая закономерность. Это проявление, в частности, заключается в том, что выборочная средняя имеет тенденцию совпасть с генеральной средней, являющейся в этом случае ее математическим ожиданием.

Теория вероятностей доказывает, что если организуется отбор по принципам случайной выборки и статистическая характеристика является математическим ожиданием выборочной характеристики, последняя при достаточном числе наблюдений как угодно мало отличается от искомой характеристики, более или менее точно отображает ее величину.

Наиболее существенным преимуществом случайной выборки является возможность объективной оценки результатов обследования. Теория вероятностей разработала систему положений, позволяющих судить о возможных размерах ошибки такой выборки с практически достаточной точностью и надежностью. Достаточной точностью потому, что, увеличивая число наблюдений, всегда можно достигнуть такого предела, когда ошибка выборки будет как угодно мала; достаточной надежностью в том смысле, что указанная точность может быть обеспечена с любой, как угодно близкой к единице вероятностью; кроме того, строгое соблюдение принципов случайного отбора гарантирует от систематической ошибки репрезентативности.

Полное и безоговорочное осуществление принципов случайной выборки в чистом виде, или собственно случайной выборки, предполагает повторный отбор по жребию.

Математическая схема случайного повторного отбора наиболее простая, особенно при многократных измерениях какой-либо постоянной величины, с которыми мы встречаемся в естественных науках. Эти измерения должны быть организованы таким образом, чтобы систематические ошибки были невозможны, а размах колебаний размеров случайных ошибок должен ограничиваться какими-то определенными пределами. Если размах вариации ошибок неограничен, то достижение необходимой точности невозможно. Поэтому иногда считают, что схема бесповторного и других видов случайного отбора при увеличении объема выборки в конечном счете имеет своим пределом те же законы распределения возможных значений выборочных средних, что и повторный отбор.

Очевидна нецелесообразность повторного отбора в социально-экономических исследованиях, при котором нужно было бы вторично регистрировать, а затем учитывать с двойным, тройным и т. д. весом значения признаков, если бы жребий снова падал на те же единицы совокупности. Кроме того, в основе расчетов при этом отборе лежит абсолютная численность выборки,

в то время как при других схемах отбора, например при так называемой бесповторной выборке, приобретает значение также и доля выборки в изучаемой совокупности. Наконец, и это пожалуй самое главное, все попытки применить принципы случайной выборки в чистом виде постоянно встречаются с почти непреодолимыми трудностями, заставляющими отступать от теоретической схемы и производить отбор с существенными его модификациями. Практика применения выборок показывает, что в ряде случаев отступления от точного соблюдения принципов случайного отбора не только не ухудшают, но, наоборот, повышают точность результатов наблюдения.

Уже бесповторная случайная выборка, которую в учебниках теории вероятностей называют статистической выборкой, тем самым подчеркивая происхождение ее из практики статистических исследований, означает известное ограничение «свободной игры» случая и нарушение принципа равной вероятности попасть в выборку. После того, как отобрана первая единица из генеральной совокупности, вероятность попасть в выборку у остальных единиц несколько изменяется. После отбора второй единицы она опять изменяется, и так продолжается до конца отбора. Вариация возможных значений выборочной средней заметно суживается по сравнению с их вариацией при повторном отборе. Это означает ограничение «свободной игры случайностей», ограничение действия закона больших чисел и подчинение его воле исследователя. При бесповторной выборке это ограничение, правда, еще очень не велико и наиболее полно осуществляется только в планомерно организованном отборе.

Такие же ограничения принципов случайного отбора наблюдаются и при других видах выборки. Например, при районированной или «расслоенной» выборке вероятности попасть в состав выборки по отдельным районам могут быть различны и зависят от принятой пропорции отбора. Последний организуется так, чтобы в выборку обязательно попали представители всех районов («слоев»), т. е. признаки отбора связаны с изучаемыми признаками.

Обобщение практики выборочных исследований показывает, что более или менее точное соблюдение принципов случайного отбора встречается исключительно при изучении природных явлений и в технике, т. е. в области математической статистики. Что касается практики экономических исследований, то здесь применение повторного случайного отбора совершенно нецелесообразно. Осуществление принципов случайного отбора требует, чтобы совокупность, из которой производится отбор, была такого большого объема, из которого можно отобрать многочисленную выборку. В такой выборке должно проявиться действие закона больших чисел, сокращающее размеры ошибок репрезентативности. В противном случае выборка нецелесообразна и необходимо исчерпывающее сплошное наблюдение. Мало того,

случайная выборка, легко организуемая для изучения небольшого числа независимых друг от друга признаков, становится практически неосуществимой, когда число наблюдаемых признаков возрастает и они связаны между собой. Увеличение числа признаков приводит к сочетанию признаков и требует такого огромного объема выборки, вместо которого выгоднее произвести сплошное наблюдение.

Практика выборов в социально-экономических исследованиях выдвинула ряд приемов, которые объединяются под общим названием планомерно организованного отбора. В отличие от случайной выборки этот метод отбора теоретически разработан пока еще недостаточно, особенно в оценке возможных ошибок выборки. Однако практика применения его в русской и особенно советской статистике показывает, что он обладает большими преимуществами перед случайной выборкой.

В основе планомерно организованного отбора, как уже упоминалось, также лежат два принципа: принцип дифференциации изучаемых явлений и принцип пропорционального представительства. Эти принципы исходят из того, что изучаемые статистикой общественные явления неоднородны, объективно разбиваются на типические группы, каждая из которых обладает своими специфическими, присущими только ей свойствами. Статистическое наблюдение, в том числе и выборочное наблюдение, обязано собирать цифровые данные о каждой такой группе. Большое значение имеет и систематизация объектов изучаемой совокупности. Чем теснее связь между изучаемыми признаками и признаками отбора, тем лучше результаты выборки, тем точнее она воспроизводит генеральную совокупность.

Для обеспечения в выборочной совокупности пропорционального представительства всех типов изучаемой совокупности требуется, чтобы предварительные данные о генеральной совокупности включали в себя сведения о том, с какими группами и признаками мы встретимся при наблюдении. Планомерно организованный отбор немыслим без предварительной оценки репрезентативности выборки, обеспечиваемой на основе всестороннего анализа собранных ранее данных. Случайная выборка эту оценку производит, как известно, на основе положений теории вероятностей.

В практике социально-экономических исследований выборочное наблюдение так или иначе увязывается с имеющимися данными обо всей совокупности, что и обеспечивает необходимую репрезентативность планомерно организованного отбора. Основное отличие этого метода отбора от метода случайной выборки заключается в том, что он принимает все доступные меры к максимально возможному сокращению случайных отклонений и вытекающих из них ошибок выборки как основному средству обеспечения репрезентативности отбора. Некоторое

ограничение «игры случайностей» имеет место и в отдельных видах случайной выборки, но только планомерно организованный отбор доводит его до максимально возможных пределов.

3. Основные виды выборочного метода

В зависимости от способа отбора отдельных единиц или групп генеральной совокупности различают следующие основные виды выборочного метода: собственно случайную выборку — повторную и бесповторную, типическую (районированную), механическую и гнездовую.

Собственно случайная повторная выборка наиболее последовательно осуществляет принципы случайного отбора. Для того чтобы обеспечить равную для всех единиц генеральной совокупности вероятность попасть в выборку, отбор производится последовательно по одной единице, но так, что каждая отобранная единица продолжает принимать участие в дальнейшей жеребьевке и может повторно и даже несколько раз попасть в выборку. Собственно случайная повторная выборка равносильна многократным измерениям одной и той же величины. Точность этой выборки обеспечивается положениями теории вероятностей, в частности законом больших чисел, который для данного случая может быть сформулирован следующим образом: при достаточно большом объеме выборки можно ожидать с вероятностью, как угодно близкой к единице, что выборочная средняя будет как угодно мало отличаться от генеральной средней. Мерой точности этой выборки служит так называемая квадратическая ошибка, т. е. средняя квадратическая из всех возможных при данном объеме выборки отклонений выборочной средней от генеральной, которая, как доказывается в математической статистике, может определяться из выражения:

$$\mu = \sqrt{\frac{\sigma^2}{n}},$$

где σ^2 — дисперсия (средний квадрат отклонения) значений признака в генеральной совокупности,

n — объем (численность) выборки.

Для практических расчетов обычно пользуются не средней квадратической ошибкой, а ее произведением на «коэффициент доверия» t , связанный с вероятностью того, что фактическая ошибка выборки окажется меньшей, чем исчисленная по этому выражению величина. Связь t с величиной указанной вероятности довольно сложна и выражается так называемым интегралом вероятностей Ляпунова—Лапласа—Гаусса. Для практических целей используют специальные таблицы, указывающие значе-

ния вероятностей и соответствующих им значений «коэффициентов доверия». Эти таблицы имеются во всех учебниках теории вероятностей и математической статистики.

Если произведение $t\mu$ назвать «доверительной ошибкой», то ее размеры определяются из выражения $\Delta = t \sqrt{\frac{\sigma^2}{n}}$, откуда видно, что при заданной вероятности, от которой зависит t , и определенной дисперсии признака σ^2 точность выборки зависит исключительно от ее объема n . На основании приведенного выражения нетрудно получить и выражение для исчисления объема выборки, достаточного для того, чтобы обеспечить заданную точность с предполагаемой вероятностью, а именно:

$$n = \frac{t^2 \sigma^2}{\Delta^2}.$$

Из приведенных выражений видно, что ни точность выборки, ни ее объем не зависят от объема генеральной совокупности, а численность выборки в ряде случаев может даже превышать этот объем, т. е. для достижения точности выборка должна будет обследовать больше единиц, чем сплошное наблюдение, и отдельные единицы будут учитываться в выборке не один, а два и более раза. Поэтому собственно случайная выборка практического значения не имеет и фактически никогда не применяется. В теории выборочного метода она рассматривается только потому, что математическая схема ее наиболее проста и бесспорна.

Собственно случайная бесповторная выборка производится также по жребию, наудачу, но в процессе отбора принимаются меры к тому, чтобы единица, однажды попавшая в выборку, вторично в нее попасть не могла. Для этого уже отобранные единицы в жребии участия не принимают. Это приводит к нарушению принципа равной вероятности попасть в выборку, но оно не ухудшает, а улучшает результаты отбора. Чтобы убедиться в этом, достаточно сравнить выражения для квадратической ошибки и определения достаточного объема бесповторной выборки с приведенными выше формулами для повторной.

Квадратическая ошибка при бесповторной выборке равна:

$$\mu = \sqrt{\frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right)},$$

а объем выборки определяется из выражения

$$n = \frac{N t^2 \sigma^2}{N \Delta^2 + t^2 \sigma^2},$$

где N — объем генеральной совокупности, а все остальные обозначения те же, что и выше.