



Методика социологических исследований

*Марат Хаджимухамедов,
Игорь Бокун,
Фархад Гарипов,
Иван Решетников*

СОВРЕМЕННЫЕ ПОДХОДЫ К ПОСТРОЕНИЮ РЕПРЕЗЕНТАТИВНОЙ НАЦИОНАЛЬНОЙ ТЕРРИТОРИАЛЬНОЙ ВЫБОРКИ

Введение

Процесс построения дизайна репрезентативной национальной территориальной выборки представляет собой нетривиальную научную задачу и нуждается в постоянном совершенствовании. Целью данной статьи является представление теоретического и практического обоснования существующих на сегодняшний день подходов к построению и реализации такого типа выборки при проведении опросов общественного мнения в Узбекистане.

В основу статьи положены материалы, описывающие наиболее распространенные подходы к построению национальных выборок за рубежом, а также результаты отечественных исследований в этой области, проведенных Центром изучения общественного мнения «Ижтимоий Фикр» в Узбекистане.

Современные концепции выборки

Вначале отметим еще раз, что данная статья ориентирована на рассмотрение территориальных аспектов репрезентативной выборки, связанной со спецификой проведения общенациональных опросов общественного мнения. Под репрезентативностью (представительностью) национальной выборки далее будет подразумеваться ее свойство обеспечивать возможность обобщения полученных результатов исследования на всю страну с приемлемым уровнем достоверности (точности и надежности).

В общем случае достижение репрезентативности обеспечивается сходством структуры и количества выбираемых объектов исследования со структурой и количеством объектов генеральной совокупности в определенном тематическом контексте. Следует отметить, что при обобщенном характере опросов, направленных на выявление консолидированного мнения значительных групп общества, как в случае проведения исследований общественного мнения, такое сходство может носить также достаточно обобщенный характер. При более конкретной ориентации опросов сходство должно носить более конкретный характер.

При оценке степени предметного сходства выборочных и генеральных объектов исследования целесообразно воспользоваться информацией о распределении ответов респондентов на ключевые вопросы-индикаторы по данной тематике иссле-



дования. Однако такая информация, основанная на материалах уже проведенных исследований, может быть использована только для косвенного прогноза, так как в новую выборку, как правило, попадают новые объекты. Поэтому при построении выборки ориентируются не на параметры непосредственно предметного характера, а на параметры статистического характера, значения которых априорно известны. Такие статистические параметры выборки должны позволять достоверно моделировать обобщенные предметные свойства генеральной совокупности.

С точки зрения концепций системного анализа изучение любого вопроса можно осуществить через призму понятий пространства, времени и группы. В случае параметров выборки по аналогии могут быть рассмотрены их территориальные, временные и групповые аспекты. Если не рассматривать временной аспект, то среди параметров выборки могут быть выделены социально-территориальные и социально-групповые параметры. Следовательно, обеспечение сходства выбираемых объектов исследования с объектами генеральной совокупности может быть рассмотрено в контексте соответствующих тематике исследования статистических социально-территориальных и социально-групповых параметров.

В качестве социально-территориальных параметров обычно используют такие параметры, которые имеют выраженную территориальную зависимость, например, к таким параметрам может быть отнесен некоторый уровень административной структуры страны или тип географической зоны. Хотя сам по себе тип географической зоны не выражает никаких социальных свойств, но, тем не менее, может оказывать значительное косвенное влияние на социальную картину в данной местности. В частности, этот параметр может определять направленность хозяйственной деятельности, влиять на культурные особенности населения.

В качестве социально-групповых параметров обычно используют такие параметры, которые более характеризуют не социально значимую специфику территории, а социально значимую специфику проживающего на ней населения, например, уровень образования или уровень доходов.

При проведении опросов обобщенного характера выбор социально-террито-

риальных и социально-групповых параметров также должен носить достаточно обобщенный характер, так как иначе может быть внесена излишняя конкретизация и нарушена обобщенность картины в масштабах страны.

А теперь, после рассмотрения структурных аспектов выборки, рассмотрим аспекты выборки, связанные с ее размером и размером ее частей.

В соответствии с положениями статистической теории под репрезентативностью понимается приемлемая степень различия между средним значением некоторого ключевого параметра в выборке и генеральной совокупности, называемой погрешностью, при приемлемой степени вероятности такого различия, называемой надежностью. Распределение значений ключевого параметра должно иметь нормальную форму.

При выполнении этого условия степень репрезентативности случайной равновесной выборки может быть оценена соотношением между диапазоном разброса значений ключевого параметра в выборке, который зависит от ее размера, и таким диапазоном в генеральной совокупности.

Желательно, чтобы ключевой параметр выборки обладал наиболее широким диапазоном разброса значений (дисперсией) среди всех используемых параметров, потому что в этом случае определяемый на его основе размер выборки будет гарантировать ее достоверность в разрезе параметров выборки, обладающих меньшими диапазонами разброса значений.

Необходимо также адаптировать к конкретной ситуации количество и размеры частей выборки, придерживаясь при этом некоторого баланса. Баланс необходим потому, что хотя мелкие и широко распределенные части выборки лучше моделируют генеральную совокупность в глобальном смысле, но более крупные части выборки могут лучше моделировать ее в локальном смысле. В частности, если предположить, что диапазон уровня доходов в городах значительно шире, чем в сельской местности, то на основании такого предположения можно определить большие объемы выборки для городских населенных пунктов, чем для сельских населенных пунктов.

В заключение следует также отметить, что на размер выборки сильное влияние могут оказывать требования, предъявляе-



мые к точности проведения внутреннего группового анализа, причем более сильное влияние, чем оказывают требования, предъявляемые к ее точности в целом¹.

Если использовать вариант расчета размера выборки для долевого признака с двумя градациями и максимально возможным разбросом значений 0.5/0.5, то, начиная с размера в 400 респондентов, выборка уже становится приемлемо репрезентативной с точностью 5% и надежностью 95% практически для населения любой страны.

Но если целями исследования предполагается проведение группового анализа, то размер выборки должен обеспечить как приемлемые масштабы представления этих групп, так и приемлемый уровень значимости различий между ними при сравнительном анализе. Например, размер выборки в 400 респондентов позволяет обеспечить статистически значимый анализ различий для двух равных групп по 200 респондентов, только начиная с уровня в 10%. Если необходимо проведение более точного анализа различий, то размер выборки должен быть соответственно увеличен.

Стратификация и кластеризация выборки

Наиболее привлекательной с точки зрения простоты реализации, адекватности структуры, а также возможности оценки точности и надежности является простая случайная выборка, при которой каждый объект генеральной совокупности имеет равные шансы попасть в эту выборку посредством применения специальных математических приемов. Такая случайная выборка имеет очевидные преимущества по сравнению с неслучайной выборкой, например, квотной (когда объекты выбираются в заданных пропорциях, но при этом допускается субъективный подход интервьюеров к выбору), так как лишь при использовании стандартных математических процедур может быть гарантирована объективность процесса выбора объектов.

Однако такая выборка обладает и существенными недостатками: необходим список всех представителей генеральной совокупности, стоимость исследования велика из-за удаленности объектов друг от друга, статистическая погрешность может возникать по всем параметрам выборки, даже по тем, для которых известны их истинные пропорции.

Поэтому при построении территориальной национальной выборки обычно применяют не простую случайную выборку, а случайную выборку с использованием методов стратификации и (или) кластеризации объектов генеральной совокупности². Стратификация позволяет обеспечить большую степень сходства структуры выборки со структурой генеральной совокупности, а кластеризация позволяет снизить уровень транспортных расходов посредством обеспечения компактного расположения объектов опроса.

Стратификация (называемая также ввиду территориальной ориентации – районированием) представляет собой разделение территориальных объектов на несколько непересекающихся частей одним из перечисленных ниже способов (Табл. 1).

Выбор определенного типа стратификации или же их комбинации зависит от конкретных целей и специфики исследования. Применение метода стратификации позволяет более гибко подходить к дизайну выборки. Следует также отметить, что применение метода стратификации в случае формирования достаточно однородных страт позволяет уменьшить погрешность по сравнению с погрешностью простой случайной выборки.

Наиболее распространенным на практике типом стратификации является пропорциональная стратификация, при которой сначала случайно, но с вероятностью, пропорциональной размеру, выбирается несколько страт, а затем уже из отобранных страт методом простой случайной выборки отбираются их объекты. Такой отбор может носить иерархический характер, когда одни страты входят в другие страты. При этом желательно, чтобы элементы, относящиеся к одному уровню иерархии, были как можно более однородными, а элементы, относящиеся к разным уровням иерархии, были как можно более разнородными.

С точки зрения территориальной непрерывности страт встречаются два подхода к стратификации: когда страты располагаются по принципу территориального единства составляющих их объектов и когда допускается разнообразие территориального расположения объектов страт.

Например, в первом случае вся страна может быть разделена на регионы, а каждый регион на городскую и сельскую его часть. Во втором случае страты могут, например, объединять объекты не обязательно тер-



Таблица 1

Описание распространенных типов стратификации

Тип стратификации	Описание	Преимущества	Недостатки
Пропорциональная стратификация	Выбираются пропорциональные размеры всех страт	Увеличивается точность выборки в целом	Усложняется процесс сравнения страт с разной точностью
Равномерная стратификация	Выбираются равные размеры всех страт	Позволяет сравнивать страты с одинаковой точностью	Уменьшается точность выборки в целом
Стратификация Неймана	Размеры страт выбираются в зависимости от сложности оценки некоторого параметра	Увеличивается точность выборки по некоторому важному параметру	Уменьшается точность выборки по другим менее важным параметрам
Оптимальная стратификация	Размеры страт выбираются в зависимости от сложности условий или стоимости опроса	Увеличивается размер выборки для «дешевых» страт	Уменьшается точность выборки для «дорогих» страт

риториально расположенные рядом, в частности, страты могут включать территории с примерно равным количеством населения, но входящие в разные регионы.

При выборе параметров стратификации чаще применяют социально-территориальные параметры, рассмотренные ранее.

К наиболее распространенным принципам стратификации, используемым в мировой практике при построении выборок, можно отнести:

- административно-территориальный;
- экономико-географический;
- социально-экономический;
- социально-культурный;
- политически ориентированный.

При этом среди наиболее часто используемых на практике параметров можно выделить:

- региональный тип;
- тип местности;
- категорию численности населения;
- тип географической зоны;
- тип хозяйственной деятельности.

В качестве примера возможного подхода к стратификации рассмотрим подход, примененный при проведении социологических исследований в Грузии. В 2006 г. Институтом социальных исследований и анализа была проведена пропорциональная численности населения стратификация в разрезе регионального типа – Тбилиси, Западная Грузия, Восточная Грузия, типа местности – город/село и категории численности населения – большие/малые города³. Требования к более подробной стратификации в разрезе типов географических зон и типов хозяйственной деятельности не

предъявлялись, предположительно, ввиду высокой степени однородности ситуации.

В России ввиду наличия большего разнообразия территориальных характеристик применяется и более сложный вариант стратификации. В частности, при проведении сотрудниками Красноярского государственного университета под руководством проф. В. Г. Немировского серии исследований 1992–1998 гг. применялись следующие подходы к стратификации выборки:

- на первой ступени выбирались:
 - в городской местности: большие, средние и малые города;
 - в сельской местности: сельские районы и районы с преобладанием лесной промышленности;
- на второй ступени выбирались:
 - в городской местности: административные, промышленные, «спальные» районы, места индивидуальной застройки;
 - в сельской местности: райцентры, села, деревни.

В этом примере уже присутствуют страты, выделяющие районы с таким типом хозяйственной деятельности, как лесная промышленность.

Кластеризация применяется для повышения экономичности выборки. Цель кластеризации заключается в снижении транспортных расходов на выборку посредством обеспечения приемлемо допустимой степени территориальной компактности объектов опроса. Наиболее часто кластеризация проводится последовательным выбором групп объектов в их естественной территориальной иерархии, начиная с некоторого верхнего уровня, например, уров-



ня районов, и заканчивая самым нижним уровнем, например, уровнем домохозяйств, в которых предполагается провести опрос. Реже встречается подход с объединением в кластеры территориально разрозненных объектов.

Для обеспечения представительности выборки при кластеризации необходимо стремиться к максимально возможному разнообразию объектов выборки внутри кластеров всех уровней.

В идеале желательно, чтобы кластеры были похожи на всю генеральную совокупность, так как это и означает, что выборка хорошо представлена своими кластерами. Если стратификация разделяет генеральную совокупность на страты однородных объектов, то кластеризация объединяет в кластеры разнородные объекты.

Возможны следующие сочетания кластеризации со стратификацией:

- только стратификация или только кластеризация;
- стратификация-кластеризация или кластеризация-стратификация;
- их последовательное многократное применение.

В качестве параметров кластеризации в зависимости от целей исследования могут быть определены как социально-территориальные, так и социально-групповые параметры.

Процесс проектирования кластеризации в общем случае включает в себя следующие этапы:

- определение типа сочетания кластеризации со стратификацией;
- определение тематического характера кластеризации;
- определение типа и состава параметров кластеризации;
- определение количества уровней кластеризации;
- определение количества и размера кластеров всех уровней;
- определение принципов формирования кластеров;
- определение принципов выбора кластеров и их объектов.

Следует отметить, что в целях обеспечения представительности выборки, с одной стороны, желательно определять как можно большее количество кластеров, но как можно меньших размеров, но, с другой стороны, слишком малые размеры кластеров или объектов опроса могут не позволить охватить дисперсию ключевого пара-

метра (или параметров) выборки на некотором важном для анализа уровне.

Также следует отметить, что более высокого уровня представительности при проведении кластеризации можно добиться, если использовать принцип случайного пропорционального отбора кластеров или объектов в кластерах, так как в этом случае обеспечивается равная вероятность попадания в выборку любого объекта генеральной совокупности.

В качестве примера комбинирования кластеризации со стратификацией рассмотрим подходы, примененные ВЦИОМ (Россия) в 2007–2008 гг.

Первая ступень выборки представляла собой региональный отбор субъектов РФ на основе параллельного проведения двух типов кластеризации: социально-экономической и политической.

1. Социально-экономическая кластеризация (17 кластеров):

- 1 группа параметров – характеристики уровня жизни;
- 2 группа параметров – характеристики инвестиционной активности;
- 3 группа параметров – характеристики экономического потенциала.

2. Политическая кластеризация (4 кластера – 4 партии).

Обе описанные выше кластеризации регионов накладывались друг на друга, и в результате были отобраны 33 группы регионов (обобщенных кластера), причем регионы из одной группы находились в одном кластере в каждой кластеризации. Затем, исходя из принципов максимальной представительности в разрезе предложенных типов кластеризации и пропорциональности численности населения, из общего числа 80 регионов (субъектов РФ) отбирались только 46 регионов.

После проведения комплексной кластеризации в каждом из отобранных регионов было выделено 7 типов страт:

- города с населением более 1 млн. чел.;
- города с населением от 500 тыс. до 1 млн. чел.;
- города с населением от 100 тыс. до 500 тыс. чел.;
- города с населением от 50 тыс. до 100 тыс. чел.;
- города с населением менее 50 тыс. чел.;
- поселки городского типа (ПГТ);
- сельские районы.

Что привлекает внимание в предложенной выборке?



• Тематическая ориентация данной выборки имела в совокупности социально-экономическую и социально-политическую направленность. При этом особенности территориального потенциала в разрезе каких-либо географических или хозяйственных характеристик не рассматривались.

• Достаточно высокий уровень территориального смещения свойств, характеризующих эти направления, предположительно вызвал необходимость применения на первом этапе выборки не четкого разделения всей территории на страты, а выделения обобщенных групп на основе комплексной многоуровневой кластеризации.

• Комплексная многоуровневая кластеризация включала в соответствие с тематической ориентацией выборки экономическую и политическую кластеризацию.

• Экономическая кластеризация была более ориентирована на определение текущего состояния наиболее социально значимых характеристик экономической ситуации (уровни доходов, доверия и занятости населения).

• Политическая кластеризация определялась результатами выборов в Государственную Думу и соответствовала четырем партиям.

• Обобщенная кластеризация позволила определить 33 специфически различных группы регионов на основе результатов комплексной многоуровневой экономико-политической кластеризации всех 80 субъектов РФ.

• Непосредственно в выборку попадали только 46 регионов из 80, но при этом обеспечивалась пропорциональная численности населения представительность в выборке всех 33 обобщенных кластеров.

• Каждый из 46 отобранных регионов далее разделялся на страты путем пропорциональной стратификации в разрезе типа местности проживания населения и его численности.

• Учитывая высокий уровень дифференциации и численности городского населения, были выделены 6 типов городских страт.

• Выборка носила комбинированный многоступенчатый характер.

Подходы к построению общенациональной территориальной выборки при проведении опросов общественного мнения Центром «Ижтимоий фикр» в Узбекистане

Многолетние исследования, проведенные Центром «Ижтимоий Фикр» в области построения, реализации и анализа выборов, позволили сформулировать следующие концепции⁶.

Общие концепции

• Ввиду значительной степени однородности основных социально-групповых характеристик всех регионов Узбекистана в качестве параметров территориального этапа выборки были определены параметры социально-территориальной ориентации.

• Социально-групповые (в данном случае социально-демографические) параметры используются лишь на этапе выбора респондентов в домохозяйствах, они включают в себя пол и возраст. Представительность по этим социально-демографическим параметрам обеспечивается путем применения метода выборки респондентов по картам Киша.

• Размер выборки определяется не на основе дисперсии некоторого конкретного параметра, а на основе максимально возможной дисперсии.

• Размер выборки, определенный на уровне 1100 респондентов, обеспечивает ее репрезентативность в национальных масштабах на уровне 3% точности и 95% надежности.

• Данный размер выборки позволяет проводить статистически значимый сравнительный анализ ответов основных социальных групп в региональных масштабах, например, мужчин и женщин или горожан и сельчан.

• Так как тематика изучения общественного мнения носит обобщенный характер, то выбираются и наиболее обобщенные социально-территориальные параметры выборки.

• Стратегия выборки заключается в проведении социально-территориальной стратификации, а ее тактика – в проведении последующей социально-территориальной иерархической кластеризации.

• Пропорциональная стратификация объема выборки обеспечивает репрезентативность в разрезе регионов и типа местности (городской и сельской).

Концепции стратификации

• Ввиду значительного своеобразия ситуации в каждом их регионов и необходимости проведения анализа в регио-



нальных масштабах территория страны была разделена на региональные страты.

• Так как социально-территориальные особенности значимы и хорошо различаемы по типу местности, то проводится дальнейшая стратификация территорий всех регионов на городскую и сельскую местность.

• Относительно небольшое количество городов и небольшая дифференциация численности их населения позволяют выделить только два типа страт: большие города и малые города.

• Вследствие относительной однородности внутренние городские территории стратифицируются по следующим крупным признакам:

– административные, промышленные и «спальные» районы;

– многоэтажные и одноэтажные районы;

– старые и новые районы.

• Ввиду значительной разнородности географических и хозяйственных характеристик сельской местности, но сочетающейся с возможностью ее территориального разделения, проводится комплексная стратификация по обоим соответствующим признакам:

– по типу географической зоны (горный тип, предгорный, степной, пустынный тип);

– по типу хозяйственной деятельности (растениеводство, животноводство, смешанный тип).

• Всего для сельских районов определяется 12 типов страт.

Концепции кластеризации

• С целью снижения транспортных расходов путем повышения территориальной компактности выборки применяется естественная территориальная иерархическая многоступенчатая кластеризация всех вышеописанных городских и сельских страт.

• Для городской местности определена следующая иерархия кластеров: страта города – район – квартал – махалля – домохозяйства.

• Для сельской местности определена следующая иерархия кластеров: страта села – район – сельский центр – махалля – домохозяйства.

• Выбор территориальных кластеров осуществляется случайным образом, но пропорционально численности, проживающего на этих территориях населения. С целью

повышения компактности в кластеризации не используются наиболее удаленные районы и сельские махалли.

• Размеры и количество кластеров, а также их объектов варьируются в зависимости от объема выборки, пропорционально определенному для каждой из страт, и особенностей численности населения для данного уровня кластеризации.

Заключение

Как следует из вышеизложенного материала, процесс построения выборки тесно связан с моделированием органической взаимосвязи между тремя группами параметров:

• социально-предметными;

• социально-территориальными;

• социально-групповыми.

Именно построение модели генеральной совокупности, достоверно отражающей такого рода взаимосвязи, может обеспечить необходимые условия для создания адекватной ей модели выборочной совокупности.

Ключевыми вопросами, возникающими при построении таких моделей, группе авторов данной статьи представляются три нижеследующих вопроса.

• Какие важные для целей исследования (или группы исследований) социальные формирования (группы) необходимо в первую очередь выделить из генеральной совокупности в контексте общности проявляемых ими предметноориентированных социальных свойств?

• Как обеспечить доступ к этим группам при сборе данных на основе построения связи их предметно ориентированных социальных свойств с имеющейся статистической информацией, отражающей их социально-территориальные и социально-групповые свойства?

• Как обеспечить эффективную реализацию дизайна выборки с учетом всего комплекса имеющихся ресурсных ограничений (методических, информационных, финансовых, технических, временных)?

В заключение отметим, что коллектив авторов имеет намерение изложить свое мнение как по этим, так и по другим, связанным с построением национальной выборки, вопросам и поэтому планирует продолжить публикацию статей по данной теме. ■



¹ Батыгин Г. С. Лекции по методологии социологических исследований. Глава 5. Проектирование выборки. М., 1995.

² Могильчак Е. Л. Выборочный метод в эмпирическом социологическом исследовании. Учебное пособие. Екатеринбург, 2015; Чуриков А. С. Случайные и неслучайные выборки в социологических исследованиях. ФОМ. Социальная реальность. 2007. №4; Ермолаев А. Выборочный метод в социологии. М., 2000.

³ Описание построения выборочной совокупности регулярного всероссийского опроса. Материалы исследований ВЦИОМ. М., 2007.

⁴ Результаты исследования «Отношение населения Грузии к социально-политическим процессам, происходящим в стране». (Материалы исследований ISSA, Грузия). Тбилиси, 2006.

⁵ Материалы исследований Центра изучения общественного мнения «Ижтимоий фикр». Т., 1997–2017.

РЕЗЮМЕ

Марат ХАДЖИМУХАМЕДОВ, Игорь БОКУН, Фархад ГАРИПОВ, Иван РЕШЕТНИКОВ

РЕПРЕЗЕНТАТИВ МИЛЛИЙ ХУДУДИЙ ТАНЛОВНИ ҚУРИШГА ДОИР ЗАМОНАВИЙ ЁНДАШУВЛАР

Маколада Ўзбекистонда жамоатчилик фикри сўровини ўтказишда фойдаланиладиган репрезентатив миллий ҳудудий танловга бугунги кунда мавжуд бўлган ёндашувлар назарий ҳамда амалий жиҳатдан асослаб берилган.

Калит сўзлар: социологик тадқиқот, миллий ҳудудий танлов, репрезентативлик, стратификация, кластеризация.

SUMMARY

Marat KHADJIMUKHAMEDOV, Igor BOKUN, Farkhad GARIPOV, Ivan RESHETNIKOV

MODERN APPROACHES TO CONSTRUCTION OF REPRESENTATIVE NATIONAL AREA SAMPLE

The article presents a theoretical and practical justification for the existing approaches to the construction of a representative national area sample used in conducting public opinion surveys in Uzbekistan.

Key words: sociological research, national area sample, representativeness, stratification, clusterization.

РЕЗЮМЕ

Марат ХАДЖИМУХАМЕДОВ, Игорь БОКУН, Фархад ГАРИПОВ, Иван РЕШЕТНИКОВ

СОВРЕМЕННЫЕ ПОДХОДЫ К ПОСТРОЕНИЮ РЕПРЕЗЕНТАТИВНОЙ НАЦИОНАЛЬНОЙ ТЕРРИТОРИАЛЬНОЙ ВЫБОРКИ

В статье представлено теоретическое и практическое обоснование существующих на сегодняшний день подходов к построению репрезентативной национальной территориальной выборки, используемой при проведении опросов общественного мнения в Узбекистане.

Ключевые слова: социологическое исследование, национальная территориальная выборка, репрезентативность, стратификация, кластеризация.