

**Выборочный  
метод  
социологического  
исследования**

**1. Основные понятия  
выборочного метода**

**2. Этапы построения  
выборки**

**3. Основные техники  
выборки**

С точки зрения **представительности** все исследования можно разделить на:

- **сплошные** – обеспечивают полную представительность информации, выборочная совокупность = генеральной совокупности; проводятся, если выборочная совокупность невелика (не более 100 единиц);
- **выборочные**

**Генеральная совокупность** – вся совокупность единиц наблюдения, имеющих отношение к проблеме исследования.

**Выборочная совокупность** – часть генеральной совокупности, отобранная с помощью специальных методов для всестороннего изучения.

Главная проблема выборочного исследования - отбор из генеральной совокупности объектов такой подсовкупности (выборки), которая сделала бы исследование одновременно и представительным (репрезентативным), и экономичным.

# Репрезентативность

- это свойство выборочной совокупности представлять основные, интересующие исследователя, характеристики генеральной совокупности.

Проблема репрезентативности - это проблема **правомерности экстраполяции выводов**, полученных при анализе выборочной совокупности, на всю генеральную совокупность.

**Эффективной** называется выборка, которая позволяет получить наиболее точные результаты при заданной стоимости исследования либо обеспечить заданную точность результатов при минимальных затратах.

Репрезентативность и эффективность зависят от **дизайна** выборки - стратегии и конкретных процедур ее формирования.

Дизайн выборки определяется его целями, задачами и гипотезами, а также характеристиками генеральной совокупности.

# **Основные этапы построения выборки**

- 1. Определение генеральной совокупности**
- 2. Выбор представительной основы выборки**
- 3. Подбор методов формирования выборки**
- 4. Определение объема выборочной совокупности**
- 5. Отбор единиц анализа**

# Определение генеральной совокупности

- **какие именно элементы составляют генеральную совокупность** (отдельные люди, семьи, академические группы, предприятия, населенные пункты или целые государства)?
- **какими признаками обладают элементы генеральной совокупности** (пол, возраст, образование, род занятий и т.п.); насколько они доступны для определения?
- **какова численность генеральной совокупности** (конкретная или гипотетическая)?

# Выбор представительной основы выборки

- совокупность, из которой непосредственно формируется выборка.

Различают 2 вида основы:

- **Территориальная** - документы, регистрирующие население по месту жительства. Основное достоинство - возможность охвата всех слоев населения и экономичность.
- **Производственная** - документы, регистрирующие население по месту работы или учебы. Применение ограничено обследованием отдельных предприятий, организаций или учреждений.

# Подбор методов формирования выборки

## Выборочные методы

```
graph TD; A[Выборочные методы] --> B[Случайные (вероятностные)]; A --> C[Неслучайные (детерминированные)]; B --> D[Простой случайный отбор]; B --> E[Стратифицированный отбор]; B --> F[Кластерный (гнездовой) отбор]; B --> G[Систематический отбор]; C --> H[Стихийный отбор]; C --> I[Направленный отбор]; H --> J[Отбор первого встречного]; H --> K[Отбор себе подобных]; H --> L[Интерактивные и прессовые опросы]; I --> M[Квотная выборка]; I --> N[Целевая выборка]; I --> O[Метод основного массива]; I --> P[Метод «снежного кома»];
```

### Случайные (вероятностные)

- Простой случайный отбор
- Стратифицированный отбор
- Кластерный (гнездовой) отбор
- Систематический отбор

### Неслучайные (детерминированные)

#### Стихийный отбор

- Отбор первого встречного
- Отбор себе подобных
- Интерактивные и прессовые опросы

#### Направленный отбор

- Квотная выборка
- Целевая выборка
- Метод основного массива
- Метод «снежного кома»

# Простой случайный отбор

Предполагает, что:

- генеральная совокупность **однородна**
- все ее элементы доступны для исследования в одинаковой степени
- имеется полный список элементов, составляющих генеральную совокупность (репрезентативная основа выборки).

К этому списку применяются процедуры случайного отбора, с использованием таблиц или компьютерных генераторов случайных чисел (все элементы генеральной совокупности имеют одинаковую вероятность попасть в выборку ).

# Стратифицированный случайный отбор

Применяется в случаях, когда генеральная совокупность не является однородной, слишком велика или имеет сложную структуру (основу выборки значительно проще получить для отдельных ее частей, чем для генеральной совокупности в целом).

**Генеральную совокупность разделяют на относительно однородные части или слои (страты), для каждой страты определяют собственную основу выборки, из которой производят простой случайный отбор.**

# Разновидности стратифицированного отбора:

**1. С пропорциональным размещением выборки** – предполагает, что из каждого слоя отбирается определенный процент.

Предполагаемый объем выборки при этом делится между стратами пропорционально их численности (это позволяет обеспечить для всех элементов генеральной совокупности одинаковую вероятность быть отобранным).

**2. С равномерным размещением выборки** - из каждого слоя отбирается одинаковое число единиц. Применяют тогда, когда неизвестны объемы слоев исходной совокупности.

# **Кластерный (гнездовой) отбор**

**Применение возможно, если генеральная совокупность состоит из относительно мелких групп элементов (кластеров, гнезд).**

**Основа выборки представляет собой список кластеров, к которому применяется процедура простого случайного отбора. Затем отобранные кластеры обследуются полностью или выборочно.**

**Метод применяется в тех случаях, когда трудно получить репрезентативную основу выборки, а также при ограниченных материальных и временных ресурсах.**

## **Систематический отбор**

**Предполагает, что основа выборки упорядочивается по какому-либо критерию, а затем из упорядоченного списка, с заданным шагом, извлекаются элементы.**

Лучшим критерием для списков людей считается алфавитный порядок. Более крупные объекты (населенные пункты, организации, фирмы) могут быть упорядочены по размеру, объему товарооборота и т.п.

# Стихийный отбор

Используют при ограниченности ресурсов исследования, сложности доступа к респондентам. Предполагает, что **сформированная выборочная совокупность состоит из наиболее доступных для интервьюера респондентов.**

Разновидности:

- **отбор «первого встречного»** (тех, кто попадается исследователю на улице, в учреждении и т.д.)
- **отбор «себе подобных»** (тех, кого исследователь подбирает в своем окружении, профессиональной, деловой, досуговой, дружеской среде)
- **интерактивные опросы на телевидении, радио, а также прессовые, Internet-опросы и т.д.**

Результаты стихийных опросов в лучшем случае могут быть использованы для разведывательных исследований, с целью углубления информации об изучаемом объекте.

# Квотный отбор

Наиболее распространенный среди неслучайных методов, конкурирует по точности с вероятностными видами отбора.

Представляет собой целенаправленное формирование выборочной совокупности на основе имеющихся статистических данных о структуре объекта (генеральной совокупности). По основным параметрам выборочная совокупность конструируется пропорционально генеральной, а для единиц отбора задается система **КВОТ**.

Число характеристик, выбранных в качестве квот, как правило, не превышает четырех, поскольку в противном случае процедура отбора становится очень сложной.

# Целевая выборка

Применяется для описательных качественных исследований.

Основная задача - выявить информационно богатые случаи для глубокого многостороннего изучения. На отбор решающее воздействие оказывают цели исследования.

**Выборочная совокупность проектируется как состоящая из ряда целевых групп, выделяемых на основе значимых для исследователя характеристик.**

**Метод основного массива** применяется на **небольших генеральных совокупностях**, для которых нет смысла проводить выборочное исследование. Предполагает **обследование всех, оказавшихся в момент проведения исследования на месте.**

**Метод «снежного кома» (цепная выборка)** применяется к **малочисленным гипотетическим генеральным совокупностям.** Каждого найденного члена такой совокупности спрашивают, кого из своих коллег он мог бы назвать. Полученный список принимается за основу выборки. Опрос продолжается до тех пор, пока имена в списке не начнут повторяться.

# Определение объема выборочной совокупности

Результаты выборочных исследований всегда являются отчасти **неопределенными**. Это происходит потому, что **изучается только часть генеральной совокупности**, и измерения производятся с ошибками.

Однако при отсутствии грубых просчетов в планировании и реализации выборки эти **ошибки можно контролировать**.

Виды ошибок:

- Случайные
- Систематические
- Ошибки регистрации

**Систематическая ошибка** - некоторое смещение выборочного среднего значения признака по отношению к генеральному среднему, не уменьшающееся с увеличением объема выборки. Такие ошибки обычно связаны с ошибками проектирования выборки и ошибками инструментария исследования.

**Случайные ошибки** связаны с вероятностным характером процедур извлечения выборки из генеральной совокупности и ошибками измерения, не имеющими систематического характера. Такие ошибки неустранимы, но подчиняются статистическим законам и, соответственно, поддаются контролю - они уменьшаются с увеличением объема выборки.

**Ошибки регистрации** свойственны и выборочному методу исследования.

2 вида **отклонений от плана выборки**:

1. **Замена намеченных при планировании выборки единиц наблюдения другими, более доступными.** Способ избежать - контроль деятельности анкетеров и интервьюеров.

2. **Неполный охват выборочной совокупности** - неполучение информации от части единиц наблюдения, включенных в выборку. Эти ошибки устанавливаются путем сравнения реально сформированной выборки с ее планом. Способ «снятия» - **«корректировка» выборки** - специальный пересчет значений изучаемого признака с учетом того, какая именно часть выборочной совокупности выпала из обследования.

**Степень точности** для каждого показателя, измеряемого в процессе обследования, задается  $2$  количественными характеристиками:

- **предельно допустимой величиной ошибки**
- **доверительной вероятностью** - вероятностью того, что эта величина не будет превышена.

Оба эти значения существенным образом зависят от объема выборки и способа ее извлечения. Стремление повысить точность приводит к быстрому росту необходимого объема выборки и, соответственно, стоимости исследования.

**Объем выборки** - это общее число единиц исследования, включенных в выборочную совокупность.

Объем выборки **зависит от множества факторов**, связанных с проведением исследования:

- его целей и задач
- используемых методик сбора первичной информации
- уровня однородности генеральной совокупности
- требующейся точности получаемой информации.

## Подходы к определению объема выборки:

- **Применение «правила большого пальца»** - бездоказательно принимается, что для получения точных результатов ВС должна составлять 5 % от ГС. (простой, доступный в исполнении, дешевый метод).
- **Установление объема ВС исходя из заранее оговоренных условий** (заказчик).
- **Установление объема ВС исходя из стоимости проведения опроса** (ценность и достоверность получаемой информации не принимается в расчет).
- **Определение объема выборки на основе статистического анализа** (с учетом требований к надежности и достоверности получаемых результатов).
- **Теория малых выборок** (количество градаций вопросов умножают на 25 – минимальный статистически значимый размер группы) .

	Вид выборки	Отбор	
		повторный	бесповторный
		2	3
1			
Количественный признак			
1	Собственно-случайная	$n = \frac{t^2 s^2}{\Delta_{\bar{x}}^2}$	$n = \frac{t^2 s^2 N}{\Delta_{\bar{x}}^2 N + t^2 s^2}$
2	Механическая ( $n \rightarrow \infty$ )	-	"-
3	Типическая (стратифицированная)	$n = \frac{t^2 \bar{s}^2}{\Delta_{\bar{x}}^2}$	$n = \frac{t^2 \bar{s}^2 N}{\Delta_{\bar{x}}^2 N + t^2 \bar{s}^2}$
4	Серийная	$r = \frac{t^2 \delta^2}{\Delta_{\bar{x}}^2}$	$r = \frac{t^2 \delta^2 R}{\Delta_{\bar{x}}^2 R + t^2 \delta^2}$
Альтернативный признак			
5	Собственно-случайная	$n = \frac{t^2 w(1-w)}{\Delta_w^2}$	$n = \frac{t^2 w(1-w)N}{\Delta_w^2 N + t^2 w(1-w)}$
6	Механическая ( $n \rightarrow \infty$ )	-	"-
7	Типическая (стратифицированная)	$n = \frac{t^2 \bar{w}(1-\bar{w})}{\Delta_w^2}$	$n = \frac{t^2 \bar{w}(1-\bar{w})\bar{N}}{\Delta_w^2 N + t^2 \bar{w}(1-\bar{w})}$
8	Серийная	$r = \frac{t^2 w_r(1-w_r)}{\Delta_w^2}$	$r = \frac{t^2 w_r(1-w_r)R}{\Delta_w^2 R + t^2 w_r(1-w_r)}$

Ядов В.А.: Требуемый **объем выборочной совокупности** при заданном объеме генеральной совокупности (при 5% ошибке выборки и доверительном интервале – 95%):

<b>Объем ГС</b>	<b>500</b>	<b>1000</b>	<b>2000</b>	<b>3000</b>
<b>Объем ВС</b>	<b>222</b>	<b>286</b>	<b>333</b>	<b>350</b>

<b>4000</b>	<b>5000</b>	<b>10 000</b>	<b>100 000</b>	<b>∞</b>
<b>360</b>	<b>370</b>	<b>385</b>	<b>398</b>	<b>400</b>

Расчет объема выборки для количественных вопросов строится исходя из **коэффициента вариации** – среднее квадратическое отклонение от средней арифметической.

Для пятичленной шкалы – от 27 до 62%.

Для семичленной – от 78 до 113%.

*Чем длиннее шкала, тем выше коэффициент вариации и больше объем выборки.*

Зависимость объема выборки от коэффициента вариации

К вариации	10	20	30	40	50	60	70	80	90	100	110	120
V выборки	15	61	138	246	384	553	753	984	1245	1537	1860	2213

В.И. Добренъков и А.И. Кравченко:  
*исследовательская практика позволяет  
сформулировать следующее правило:*

**Объем выборки должен обеспечивать не менее 100 наблюдений для каждой первостепенной и не менее 20-50 наблюдений для каждой классификационной составляющей.**

*Первостепенные классификационные составляющие – наиболее критичные, второстепенные – наименее критичные ячейки перекрестных классификаций, принятых в исследовании.*

# Отбор единиц анализа

2 метода отбора единиц в выборочную совокупность:

- **повторный** (каждая попавшая в выборку единица или серия возвращается в генеральную совокупность и имеет шанс вторично попасть в выборку)
- **бесповторный** (каждая отобранная единица (или серия) не возвращается в генеральную совокупность и не может подвергнуться вторичной регистрации, а потому для остальных единиц вероятность попасть в выборку увеличивается).

# Отбор единиц анализа

В зависимости от количества ступеней отбора выделяют 2 вида выборки:

- **одноступенчатая выборка** - предполагает, что из генеральной совокупности сразу выбираются респонденты для опросов;
- **многоступенчатая выборка** - включает в себя ряд этапов, на каждом из которых единица отбора меняется.