

1. ОБЩИЕ ПОЛОЖЕНИЯ

1.1. ВИДЫ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ В ЗАВИСИМОСТИ ОТ ОХВАТА ЕДИНИЦ ИССЛЕДУЕМОЙ СОВОКУПНОСТИ

Главной функцией органов государственной статистики является количественное и качественное описание важнейших социально-экономических процессов, происходящих в стране. Это достигается с помощью системы статистических показателей, рассчитываемых на основе информации, собираемой посредством проведения статистических обследований. Адекватность описания происходящих процессов во многом зависит от используемых методов сбора и обработки первичных данных.

С развитием новых экономических отношений, при переходе от экономики централизованного планирования к рыночной, традиционно применявшийся государственной статистикой сплошной метод наблюдения, носивший учетный характер, стал неэффективен в силу ряда объективных причин, в том числе его затратности. В условиях возрастания роли негосударственного сектора экономики и бурного роста числа организаций, подлежащих статистическому учёту, обеспечить полноту, точность и надёжность информации сплошным методом статистического наблюдения стало практически невозможно.

Это обусловило объективную необходимость освоения и широкого применения в статистической практике выборочных методов наблюдения во всех отраслях статистики и новых методов обработки информации.

1.1.1. СПЛОШНОЕ И НЕСПЛОШНОЕ НАБЛЮДЕНИЕ

Генеральная совокупность и статистическая основа наблюдения. Фундаментальным понятием в статистической теории является понятие совокупности, представляющей собой множество объектов или явлений, объединенных общими признаками и подвергающихся статистическому исследованию. При проведении выборочного наблюдения (обследования) различают генеральную и выборочную совокупности.

Генеральная совокупность представляет собой множество реально или гипотетически существующих объектов или явлений, из которых тем или иным способом формируется совокупность выборочная. В качестве примера можно привести генеральные совокупности населения, домашних хозяйств, предприятий и т.п.

Статистика изучает прежде всего свойства генеральных совокупностей, а не отдельных составляющих их единиц. Так, при статистическом исследовании генеральной совокупности предприятий розничной торговли представляют интерес суммарные по отрасли значения розничного товарооборота, численности занятых работников, размера торговых площадей и т.д., а также вариации и взаимосвязи данных показателей. Необходимо отметить, что генеральная совокупность на практике очень редко в точности подчиняется простым математическим законам, и именно этим, в первую очередь, обусловлена необходимость ее статистического исследования.

В математической статистике генеральная совокупность - абстрактное понятие, представляющее собой множество результатов всех мыслимых наблюдений, которые могли бы быть получены при данном комплексе условий. Так, например, обследовав все предприятия отрасли, можно рассматривать полученную совокупность только как представителя гипотетически возможной более широкой совокупности предприятий, которые могли бы функционировать при тех же самых факторах воздействия внутренней и внешней среды.

Рассматриваемые в статистике генеральные совокупности могут реально существовать, а могут и нет. Различение реальных и гипотетических генеральных совокупно-

стей на практике важно, когда делаются выводы о генеральной совокупности по извлеченной из нее выборке.

Генеральная совокупность может быть как конечной, так и бесконечной. Однако для органов государственной статистики практический интерес представляет конечная генеральная совокупность. Поэтому в сборнике под термином "**совокупность**" понимается конечная совокупность.

Конечная совокупность имеет место, например, при обследовании семейных бюджетов, когда выборка берется из совокупности семей, фактически имеющих в стране, а затем осуществляются наблюдения за доходами и расходами отобранных семей. Конечная совокупность также имеет место при экономических обследованиях организаций - резидентов рынка страны.

Государственное статистическое наблюдение базируется на создании и поддержании в актуальном состоянии статистических основ – перечней (списков) единиц реальных совокупностей с набором признаков, соответствующих целям статистических исследований.

Статистическое исследование представляет собой процесс изучения социально-экономических объектов и явлений на основе статистических методов и систем статистических показателей. К наиболее распространенным обобщающим статистическим характеристикам, получаемым в результате исследования, относятся: доля элементов генеральной совокупности, отвечающих некоторому критерию; средние и суммарные значения признаков элементов; отношение средних или суммарных значений и др.

Статистическое наблюдение – первая стадия статистического исследования, представляющая собой организованный по единой программе сбор данных о социально-экономических явлениях и процессах путем регистрации их существенных признаков с целью получения первичной статистической информации.

С точки зрения полноты охвата элементов генеральной совокупности статистическое наблюдение может быть сплошным или несплошным.

Сплошное наблюдение представляет собой обследование всех без исключения элементов изучаемой совокупности и, следовательно, получение исчерпывающей статистической информации.

Важнейшую роль для статистики играют специальные сплошные обследования - переписи населения и предприятий. В ходе проведения переписи осуществляется выяснение структуры общей совокупности и создание статистической основы (списка элементов общей совокупности с их основными характеристиками) для организации последующих обследований, в том числе выборочных.

Сплошное наблюдение используется в случае, когда обследуемая совокупность состоит из сравнительно небольшого числа элементов, либо в силу необходимости получения точной информации, в том числе по каждому элементу. Сплошные обследования необходимы для актуализации статистической основы и получения структурных данных в региональной и тематической группировке. Эта информация используется для уточнения результатов выборочных обследований.

В случае, когда проведение сплошного наблюдения не имеет смысла, невозможно по организационным причинам или экономически нецелесообразно в силу требующихся на его осуществление ресурсных затрат, используется несплошное наблюдение посредством применения выборочных методов.

Несплошное наблюдение представляет собой учёт только подмножества элементов общей совокупности, на основе которого можно получить обобщающие характеристики всей совокупности с некоторой степенью точности.

Основной причиной организации обследования выборочным методом является экономия средств, а предпосылкой - возможность судить с достаточной степенью точности (т.е. обеспечение приемлемого уровня ошибок) о характеристиках генеральной совокупности по некоторому подмножеству ее элементов, называемому выборочной совокупностью или просто выборкой.

Вследствие этого, несплошное наблюдение имеет следующие преимущества перед сплошным.

1. Затраты на получение данных лишь от части элементов общей совокупности меньше, следовательно, выборочное обследование является более экономически выгодным, чем сплошное.

2. Объем работы по сбору и обобщению результатов обследования значительно меньше, поэтому результаты выборочного обследования можно получить значительно быстрее, чем при сплошном наблюдении.

3. Так как наблюдению подвергается лишь часть элементов общей совокупности, появляется возможность расширения программы обследования, т.е. более широкого и детального наблюдения каждой единицы в отдельности.

4. При проведении обследования выборочным методом общий объем работы меньше, поэтому можно лучше подготовить и более тщательно контролировать его проведение и обработку результатов. Следовательно, выборочное обследование может дать более достоверные результаты, чем соответствующее сплошное.

При решении ряда прикладных статистических задач выборочное наблюдение является единственно возможным. Так, в технологической статистике контроль качества отдельных видов продукции невозможен без нарушения ее потребительских свойств, поэтому он может осуществляться только на основе низкопроцентной выборки. В маркетинговых обследованиях в большинстве случаев также невозможно полностью обследовать всю совокупность потенциальных потребителей того или иного товара.

Следует отметить, что, исходя из нужд практической статистики, выборочные методы широко используются для проверки результатов сплошного наблюдения (контроля качества), например, данных переписи населения, когда из-за трудностей обработки огромного объема первичной информации возможны ошибки в результатах.

1.1.2. ВЫБОРОЧНЫЙ МЕТОД НАБЛЮДЕНИЯ СПОСОБЫ ФОРМИРОВАНИЯ ВЫБОРОЧНОЙ СОВОКУПНОСТИ

Применение выборочного метода наблюдения включает следующие этапы:

определение генеральной совокупности и единиц наблюдения, обладающих первичной информацией необходимой для решения задач обследования;

создание статистической основы выборки;

формирование выборочной совокупности путем отбора элементов основы;

распространение собранных по выборке данных на генеральную совокупность.

Последний этап зависит от примененного способа отбора элементов в выборку и используемой формулы оценивания характеристик генеральной совокупности по данным выборки.

В статистической практике выборки извлекаются из конечных списочных основ. Однако единица основы, единица отбора и единица наблюдения могут отличаться. Например, это обычная ситуация при обследованиях населения, и сельскохозяйственного сектора.

При рассмотрении любой схемы извлечения выборки должны быть учтены два фактора:

а) использовалась или нет вероятностная процедура;

б) наличие или отсутствие объективности в действиях специалиста, формирующего выборку.

Смысл объективности ясен и однозначен: любой специалист, производящий отбор, получил бы ту же самую выборку, т.е. выборку с теми же самыми свойствами. Субъективность означает, что специалисту, производящему отбор, позволено опираться на собственное суждение или интуицию относительно того, что является "хорошей" выборкой.

Рассматривая каждый из этих факторов на двух уровнях, можно выделить четыре типа выборок:

Таблица 1

ТИПЫ ВЫБОРОК

Роль, которую играет специалист, осуществляющий отбор	Процедура отбора	
	Вероятностная	Невероятностная
Объективная	Выборки, построенные вероятностным (случайным) образом.	Выборки, построенные с определенной целью.
Субъективная	Выборки, построенные квазислучайным образом.	Выборки, построенные на основе суждения эксперта.

В статистической практике используются все четыре типа выборок. Однако в органах государственной статистики обычно отдают предпочтение вероятностным (случайным) выборкам как наиболее объективным, поскольку имеется хорошо обоснованная теория, позволяющая понимать поведение таких выборок и оценивать их свойства (качество) отображения характеристик всей совокупности. Свойства и объективная ценность других выборок известны в меньшей мере.

Имеется два типа выборок, основывающихся на вероятностном способе отбора: выборки, отбираемые по объективным правилам вероятностного (случайного) отбора, и выборки, отбираемые, строго говоря, не по этим правилам (квазислучайные). Материалы сборника содержат значительное число примеров использования в статистической практике объективных вероятностных выборок. Одно из наиболее ценных качеств вероятностных выборок состоит в том, что можно оценить точность получаемых результатов по данным самой выборки.

Вероятностные выборки

В теории выборочных обследований рассматриваются выборки, извлеченные из совокупностей (статистических основ), содержащих некоторое конечное число единиц N . Эти единицы различимы между собой и число различных выборок объема n , которые могут быть извлечены из списка N единиц, равно числу сочетаний C_N^n .

В выборочных обследованиях в целях расчета параметров совокупности основное внимание направлено на изучение определенных свойств единиц, которые измеряются и фиксируются в процессе наблюдения для каждой единицы, включенной в выборку. Эти свойства называют характеристиками или просто признаками.

Хотя выборка используется для многих целей, обычно представляют интерес четыре характеристики совокупности:

среднее значение признака \bar{Y} (например, среднее число занятых на одном предприятии);

суммарное значение признака Y (например, выпуск продукции предприятиями промышленности);

отношение двух суммарных или средних значений (например, отношение стоимости ликвидных активов к общей стоимости активов);

доля единиц в совокупности, относящихся к некоторой определенной группе (например, доля промышленных предприятий, оказывающих платные услуги населению) или обладающих определенным значением признака.

Главным вопросом методологии выборочного наблюдения является обеспечение приемлемого уровня ошибок получаемых значений показателей (характеристик совокупности), в том числе по требуемым разрезам, например, отраслям экономики, формам собственности и регионам России.

Полученные в результате выборочного наблюдения характеристики практически всегда несколько отличаются от характеристик генеральной совокупности. Эти отличия называются **ошибками выборки (или репрезентативности)**, которые могут быть систематическими или случайными.

Систематические ошибки имеют место в том случае, когда нарушен принцип случайности отбора и в выборку попали единицы, обладающие какими-либо свойствами, не характерными для всех единиц генеральной совокупности. Случайные ошибки обусловлены тем обстоятельством, что даже при тщательной организации выборка не может в точности воспроизвести генеральную совокупность. В отличие от ошибок систематических, случайные ошибки являются вполне допустимыми и могут быть оценены статистически.

Для измерения ошибки выборки, а также сравнения двух оценок, т.е. выявления более эффективной оценки, используют средний квадрат ошибки оценки (СКО), который измеряет ошибку относительно оцениваемого параметра совокупности:

$$\begin{aligned} \text{СКО}(\hat{\theta}) &= E(\theta - \hat{\theta})^2 = E[(\theta - \hat{\theta}) + (\bar{\theta} - \hat{\theta})]^2 = \\ &= E(\hat{\theta} - \bar{\theta})^2 + 2(\bar{\theta} - \theta)E(\hat{\theta} - \bar{\theta}) + (\bar{\theta} - \theta)^2 = \\ &= V(\hat{\theta}) + B^2, \quad \text{т.к. } E(\hat{\theta} - \bar{\theta}) = 0, E(\hat{\theta}) = \bar{\theta} \end{aligned}$$

где $\hat{\theta}$ – оценка некоторой характеристики совокупности θ , получаемая согласно некоторой схеме отбора и примененной формуле оценивания;

$\bar{\theta}$ – математическое ожидание $\hat{\theta}$ – среднее значение, взятое по всем возможным выборкам;

$B = \bar{\theta} - \theta$ – смещение оценки.

Таким образом, СКО является критерием **достоверности** оценки, который характеризует величину отклонений от истинного значения характеристики совокупности θ .

Поскольку на практике трудно проследить, чтобы оценки не давали никаких смещений, для характеристики оценки используется понятие «**точности**», относящееся к величине отклонений от среднего значения $\bar{\theta}$.

Степень точности оценки обычно характеризуется ее дисперсией, стандартной ошибкой, коэффициентом вариации (относительной стандартной ошибкой) и доверительным интервалом.

Точность какой-либо оценки, полученной по выборке, зависит от двух факторов: от способа, которым оценка вычисляется по данным выборки, и от способа формирования самой выборки.

Например, в выборочных обследованиях способ оценивания называется состоятельным, если оценка становится в точности равной оцениваемому параметру для совокупности при $n = N$, т.е. когда выборку составляет вся совокупность. Очевидно, что при простом случайном отборе выборочное среднее \bar{y} и $N\bar{y}$ представляют собой состоятельные оценки соответственно среднего и суммарного значений для совокупности.

В данном контексте способ оценивания называется **несмещенным**, если среднее значение оценки, взятое по всем возможным выборкам данного объема n , в точности равно истинному значению для совокупности и если это утверждение справедливо для любой совокупности конечных значений y_i и для любого n . Например, при простом случайном отборе выборочное среднее \bar{y} - несмещенная оценка среднего значения признака, $N\bar{y}$ - несмещенная оценка суммарного значения Y для совокупности, где \bar{y} - среднее признака y_i по выборке.

В теории и практике выборочных обследований приходится рассматривать смещенные оценки. Это обусловлено следующими причинами. В некоторых, часто встречающихся в статистической практике случаях, особенно при оценивании отношений двух величин, оценки оказываются смещенными. Даже в случае использования теоретически несмещенных оценок ошибки наблюдения и неполучение ответов от респондентов могут привести к смещениям в распространенных результатах.

Кратко опишем некоторые, наиболее часто используемые в статистической практике способы формирования вероятностной выборки.

Простой случайный отбор. Простым случайным отбором называется способ, при котором извлечение единиц из совокупности для обследования осуществляется методом жеребьевки или с использованием таблиц или генератора случайных чисел без деления этой совокупности на какие-либо классы или группы.

Простую случайную выборку получают, отбирая последовательно единицу за единицей. Единицы в совокупности нумеруются числами от 1 до N , после чего выбирается последовательность n случайных чисел, заключенных между 1 и N . Единицы совокупности, имеющие эти номера, составляют выборку. На каждом этапе отбора такой процесс обеспечивает для всех еще не выбранных номеров равную вероятность быть отобранными. Легко проверить, что равную вероятность быть отобранными имеют все C_N^n возможных выборов.

Уже отобранные номера исключаются из списка, иначе одна и та же единица могла бы попасть в выборку более одного раза. Поэтому такой отбор называется **отбором без возвращения**. Отбор с возвращением легко осуществим, но им, за исключением особых случаев, пользуются редко, поскольку нет особых оснований допускать, чтобы одна и та же единица встречалась в выборке дважды (также известно, что выборка без возвращения более эффективна).

При простом случайном отборе для получения выводов о параметрах совокупности используют выборочное среднее в качестве оценки среднего значения признака совокупности, а дисперсию выборки - для оценки дисперсии признака совокупности. Для простой случайной выборки усредненные выборочные средние и дисперсии точно равны среднему и дисперсии признака совокупности.

Таблица 2

ФОРМУЛЫ ОЦЕНИВАНИЯ ПРИ ПРОСТОМ СЛУЧАЙНОМ ОТБОРЕ

Статистические показатели	Истинное значение	Оценка
Суммарное значение признака	$Y = \sum_{i=1}^N y_i$	$\hat{Y} = \frac{N}{n} \sum_{i=1}^n y_i$
Среднее значение признака	$\bar{Y} = \frac{1}{N} \sum_{i=1}^N y_i$	$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$
Дисперсия признака	$S^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y})^2$	$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$
Дисперсия оценки суммарного значения признака	$V(\hat{Y}) = \frac{N^2 S^2}{n} \left(\frac{N-n}{N} \right)$	$v(\hat{Y}) = \frac{N^2 s^2}{n} \left(\frac{N-n}{N} \right)$
Дисперсия оценки среднего значения признака	$V(\bar{y}) = \frac{S^2}{n} \left(\frac{N-n}{N} \right)$	$v(\bar{y}) = \frac{s^2}{n} \left(\frac{N-n}{N} \right)$
Стандартная ошибка оценки суммарного значения признака	$S(\hat{Y}) = \frac{NS}{\sqrt{n}} \sqrt{\left(\frac{N-n}{N} \right)}$	$s(\hat{Y}) = \frac{Ns}{\sqrt{n}} \sqrt{\left(\frac{N-n}{N} \right)}$
Стандартная ошибка оценки среднего значения признака	$S(\bar{y}) = \frac{S}{\sqrt{n}} \sqrt{\left(\frac{N-n}{N} \right)}$	$s(\bar{y}) = \frac{s}{\sqrt{n}} \sqrt{\left(\frac{N-n}{N} \right)}$
Коэффициент вариации оценки	$CV = \frac{S(\hat{Y})}{\hat{Y}} 100\% = \frac{S(\bar{y})}{\bar{Y}} 100\%$	$cv = \frac{s(\hat{Y})}{\hat{Y}} 100\% = \frac{s(\bar{y})}{\bar{y}} 100\%$

В таблице 2 использованы следующие обозначения:

- N** - объем (количество элементов) генеральной совокупности;
- n** - объем (количество элементов) выборочной совокупности;
- i** - номер элемента;
- y_i** - значение признака у *i*-го элемента; **i = 1, 2... N** или **i = 1, 2... n**.

Замечание. Формулы для долей получаются, если в качестве признака **y** взять индикатор со значениями 0 или 1 принадлежности элемента к интересующему классу совокупности.

Другие методы отбора часто оказываются предпочтительнее простого случайного отбора по соображениям удобства или повышения точности. Однако простая случайная выборка - наипростейший вид объективной вероятностной выборки, она служит основой для многих более сложных ее видов.

Расслоенный (типизированный) случайный отбор. Расслоенный случайный отбор - это отбор, предусматривающий предварительное разделение совокупности, содержащей N единиц, на слои и проведение простого случайного отбора в каждом слое.

При расслоенном случайном отборе совокупность, содержащая N единиц, сначала подразделяется на подсовокупности, состоящие соответственно из N_1, N_2, \dots, N_L единиц. Эти подсовокупности не содержат общих единиц и вместе исчерпывают всю совокупность:

$$N_1 + N_2 + \dots + N_L = N.$$

Такие подсовокупности называются **слоями**. Для того, чтобы можно было полностью воспользоваться преимуществами этого метода отбора, значения N_L должны быть известны. Когда слои определены, из каждого слоя извлекается простая случайная выборка, причем отбор в разных слоях производится независимо. Объемы выборок внутри слоев обозначаются соответственно через n_1, n_2, \dots, n_L и, следовательно:

$$n_1 + n_2 + \dots + n_L = n.$$

Объем выборки из каждого слоя может быть пропорционален объему (размеру) этого слоя или определяется степенью дифференциации признака в данном слое или устанавливается в соответствии с некоторым составным критерием, учитывающим оба названных фактора. Однако реализация второго и третьего вариантов на практике затруднена, так как они предполагают наличие информации о вариации признаков еще до проведения обследования, получаемой, например, по результатам предшествующих обследований.

Расслоение - довольно распространенный прием, что обусловлено многими причинами. Перечислим некоторые из них.

Расслоение можно рассматривать как процедуру извлечения выборок, в которой на простой случайный отбор наложены некоторые ограничения или условия. При выполнении определенных условий и наложении правильных ограничений можно получить значительный выигрыш в точности и, как правило, с малыми дополнительными затратами либо вовсе без них. В другом, но близком смысле, расслоение - это способ включения знаний об общей совокупности и ее совокупностях по признакам в процедуру отбора таким образом, чтобы повысить ее эффективность.

При расслоенном случайном отборе управление обследованием значительно упрощено. Однако сама процедура предполагает знание объемов слоев, общего числа единиц в выборке, а также определение долей отбора в каждом слое.

Расслоение может дать выигрыш в точности при оценивании характеристик всей совокупности. Иногда неоднородную совокупность удается расслоить на подсовокупности (слои), каждый из которых внутренне однороден. Если каждый слой однороден в том смысле, что результаты измерений в нем очень мало изменяются от единицы к единице, то можно получить точную оценку среднего значения для любого слоя по небольшой выборке в этом слое. Затем эти оценки можно объединить в одну точную оценку для всей совокупности.

Проблемы, связанные с отбором в разных слоях совокупности, могут сильно различаться. Например, при обследовании, предпринятом с целью изучения деловой активности, можно составить список крупных предприятий, выделив их в отдельный, полно-

стью наблюдаемый слой. Для более мелких предприятий можно применить один из видов территориального отбора.

Таблица 3

ФОРМУЛЫ ОЦЕНИВАНИЯ ПРИ РАССЛОЕННОМ СЛУЧАЙНОМ ОТБОРЕ

Статистические показатели	Истинное значение	Оценка
Суммарное значение признака	$Y = \sum_{h=1}^L Y_h$	$\hat{Y}_{st} = \sum_{h=1}^L N_h \bar{y}_h$
Среднее значение признака	$\bar{Y} = \frac{1}{N} \sum_{h=1}^L N_h \bar{Y}_h$	$\bar{y}_{st} = \frac{1}{N} \sum_{h=1}^L N_h \bar{y}_h$
Дисперсия оценки суммарного значения признака	$V(\hat{Y}_{st}) = \sum_{h=1}^L N_h (N_h - n_h) \frac{S_h^2}{n_h}$	$v(\hat{Y}_{st}) = \sum_{h=1}^L N_h (N_h - n_h) \frac{s_h^2}{n_h}$
Дисперсия оценки среднего значения признака	$V(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L N_h (N_h - n_h) \frac{S_h^2}{n_h}$	$v(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L N_h (N_h - n_h) \frac{s_h^2}{n_h}$
Стандартная ошибка оценки суммарного значения признака	$S(\hat{Y}_{st}) = NS(\bar{y}_{st}) = \sqrt{\sum_{h=1}^L N_h (N_h - n_h) \frac{S_h^2}{n_h}}$	$s(\hat{Y}_{st}) = Ns(\bar{y}_{st}) = \sqrt{\sum_{h=1}^L N_h (N_h - n_h) \frac{s_h^2}{n_h}}$
Стандартная ошибка оценки среднего значения признака	$S(\bar{y}_{st}) = \sqrt{\frac{1}{N^2} \sum_{h=1}^L N_h (N_h - n_h) \frac{S_h^2}{n_h}}$	$s(\bar{y}_{st}) = \sqrt{\frac{1}{N^2} \sum_{h=1}^L N_h (N_h - n_h) \frac{s_h^2}{n_h}}$
Коэффициент вариации оценки	$CV = \frac{S(\hat{Y}_{st})}{Y} 100\% = \frac{S(\bar{y}_{st})}{\bar{Y}} 100\%$	$cv = \frac{s(\hat{Y}_{st})}{\hat{Y}_{st}} 100\% = \frac{s(\bar{y}_{st})}{\bar{y}_{st}} 100\%$

В таблице 3 использованы следующие обозначения:

- L** - число слоев;
- h** - номер слоя;
- Y_h** - суммарное значение признака **y** в **h**-м слое генеральной совокупности;
- N_h** - объем **h**-го слоя генеральной совокупности;

- \bar{y}_h - среднее значение признака y в h -м слое выборки;
 N - объем генеральной совокупности;
 \bar{Y}_h - среднее значение признака y в h -м слое генеральной совокупности;
 n_h - объем h -го слоя выборки;
 S_h^2 - истинное значение дисперсии для h -го слоя:

$$S_h^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2$$

- i - номер элемента внутри слоя;
 y_{hi} - значение признака y i -го элемента слоя h ;
 s_h^2 - несмещенная оценка дисперсии для h -го слоя:

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$$

Гнездовой (кластерный) отбор. Гнездовой отбор – способ формирования выборки, при котором единица отбора состоит из группы или гнезда более мелких единиц, которые называются элементами. Таким образом, **гнездовая выборочная единица** – группа элементов, которая в процессе извлечения выборки рассматривается как одна единица. В простейшем случае элементы, составляющие гнездо, либо входят в выборку как группа, либо не входят в нее вообще.

Например, если гнездом являются все квартиры в жилом квартале города, то квартиры из этого жилого квартала либо входят в выборку, либо нет - в зависимости от того, оказался ли отобранным этот квартал или нет. Ни в коем случае не может быть, чтобы одна часть квартир из жилого квартала попала в выборку, а другая из этого же квартала не попала. Это было бы возможно, если бы извлекалась, например, случайная выборка из списка квартир города.

Гнездовой метод отбора единиц наблюдения наиболее характерен для выборочных обследований в таких отраслях статистики, как статистика сельского хозяйства и статистика населения. Широкое применение гнездового отбора в статистической практике обусловлено двумя основными причинами.

Первая из них заключается в том, что для обследования может не существовать достоверной статистической основы (списка элементов совокупности), а ее составление или невозможно, или обошлось бы очень дорого. Например, эта ситуация имеет место, когда для обследований населения нет полных и неустаревших его списков. Однако по картам подлежащие обследованию районы могут быть разделены на территориальные участки с легко идентифицируемыми границами. Относясь к таким участкам как к гнездам, возможно решить задачу построения списка единиц отбора.

Вторая причина состоит в том, что, даже если имеется списочная основа элементов, экономические соображения могут диктовать выбор более крупных единиц отбора.

В статистической практике определение единицы отбора в значительной степени зависит от природы статистического исследования, от того, какого рода имеется основа, а также от ряда других факторов, которые не всегда поддаются количественной оценке. Этот выбор может быть сделан также на основании правила, согласно которому выбирается единица, обеспечивающая наибольшую точность оценок при заданной стоимости или наименьшую стоимость обследования при заданной точности.

Эффективность гнездового отбора можно повысить объединением в гнезда непохожих элементов. А затраты обычно уменьшаются, если в гнездах свести вместе территориально близкие элементы.

Таблица 4

ФОРМУЛЫ ОЦЕНИВАНИЯ ПРИ ГНЕЗДОВОМ ОТБОРЕ

Статистические показатели	Истинное значение	Оценка
Суммарное значение признака	$Y = \sum_{i=1}^M T_i$; $\bar{T} = \frac{1}{M} \sum_{i=1}^M T_i$	$\hat{Y} = \frac{M}{m} \sum_{i=1}^m T_i$
Среднее значение признака	$\bar{Y} = \frac{Y}{N}$	$\hat{\bar{Y}} = \frac{\hat{Y}}{\hat{N}}$
Дисперсия оценки суммарного значения признака	$V(\hat{Y}) = M^2 \left(1 - \frac{m}{M}\right) \frac{1}{m} \frac{1}{M-1} \sum_{i=1}^M (T_i - \bar{T})^2$	$v(\hat{Y}) = M^2 \left(1 - \frac{m}{M}\right) \frac{1}{m} \frac{1}{m-1} \sum_{i=1}^m (T_i - \hat{\bar{T}})^2$
Дисперсия оценки среднего значения признака	$V(\hat{\bar{Y}}) = \frac{1}{N^2} \left[M^2 \left(1 - \frac{m}{M}\right) \frac{1}{m} \frac{1}{M-1} \sum_{i=1}^M (T_i - \bar{T})^2 \right]$	$v(\hat{\bar{Y}}) = \frac{1}{N^2} \left[M^2 \left(1 - \frac{m}{M}\right) \frac{1}{m} \frac{1}{m-1} \sum_{i=1}^m (T_i - \hat{\bar{T}})^2 \right]$
Стандартная ошибка оценки суммарного значения признака	$S(\hat{Y}) = \sqrt{M^2 \left(1 - \frac{m}{M}\right) \frac{1}{m} \frac{1}{M-1} \sum_{i=1}^M (T_i - \bar{T})^2}$	$s(\hat{Y}) = \sqrt{M^2 \left(1 - \frac{m}{M}\right) \frac{1}{m} \frac{1}{m-1} \sum_{i=1}^m (T_i - \hat{\bar{T}})^2}$
Стандартная ошибка оценки среднего значения признака	$S(\hat{\bar{Y}}) = \frac{1}{N} \sqrt{M^2 \left(1 - \frac{m}{M}\right) \frac{1}{m} \frac{1}{M-1} \sum_{i=1}^M (T_i - \bar{T})^2}$	$s(\hat{\bar{Y}}) = \frac{1}{N} \sqrt{M^2 \left(1 - \frac{m}{M}\right) \frac{1}{m} \frac{1}{m-1} \sum_{i=1}^m (T_i - \hat{\bar{T}})^2}$
Кoeffициент вариации оценки	$CV = \frac{S(\hat{Y})}{Y} 100\% = \frac{S(\hat{\bar{Y}})}{\bar{Y}} 100\%$	$cv = \frac{s(\hat{Y})}{\hat{Y}} 100\% = \frac{s(\hat{\bar{Y}})}{\hat{\bar{Y}}} 100\%$

В таблице 4 использованы следующие обозначения:

- i - номер гнезда (кластера);
- M - количество гнезд (кластеров);
- T_i - суммарное значение признака в i -м гнезде (кластере).
- \bar{T} - среднее значение признака Y по гнездам (кластерам);
- m - количество выбранных гнезд (кластеров);
- N - объем генеральной совокупности;
- \hat{N} - оценка количества элементов генеральной совокупности;
- \hat{T} - оценка среднего значения признака Y по гнездам (кластерам):

$$\hat{T} = \frac{1}{m} \sum_{i=1}^m T_i$$

Систематический (механический) отбор. Для осуществления систематического отбора все единицы совокупности нумеруются в некотором порядке числами от 1 до N . Для получения выборки объемом n единиц сначала извлекается, например, случайным образом какая-либо единица из первых $k = Nm$ единиц совокупности. После этого в выборку включается каждая k -я единица, начиная с уже извлеченной. Извлечение первой единицы определяет всю выборку. Такая выборка называется **систематической выборкой** каждой k -й единицы. Отношение Nm называется интервалом отбора.

В отдельных случаях при наличии соответствующей информационной базы для повышения точности выборочных результатов единицы генеральной совокупности предварительно ранжируются по какому-либо существенному признаку. При таком подходе отбор единиц в выборочную совокупность начинается с единицы, находящейся в середине первого интервала.

В теории систематический отбор считается более эффективным, чем простая случайная выборка. Также его легче осуществлять при работе вручную, что потеряло актуальность с широким распространением персональных компьютеров следует отметить два недостатка этого способа отбора:

затруднено получение несмещенной оценки выборочной дисперсии;

существуют такие совокупности и значения n , при которых простая случайная выборка дает более точные оценки показателей (дисперсия оценки среднего систематической выборки для этих совокупностей может даже расти при увеличении объема выборки). Наличие периодического или циклического изменения в значениях признака, период которого равен интервалу отбора, – наихудшая из возможных ситуаций при систематическом отборе.

При систематическом отборе обычно применяются оценочные формулы простого случайного отбора, так как систематический отбор можно рассматривать как простой случайный, содержащий одну гнездовую единицу из совокупности k гнездовых единиц.

Дисперсия среднего значения при систематическом отборе определяется по формуле:

$$V(\bar{y}_{sy}) = \frac{N-1}{N} S^2 - \frac{k(n-1)}{N} S_{wsy}^2,$$

- где N - объем генеральной совокупности;
 S^2 - истинное значение дисперсии признака;
 k - шаг отбора;
 n - объем выборочной совокупности;
 S_{wsy}^2 - дисперсия единиц, принадлежащих одной и той же систематической выборке (wsy - от английского "within" - внутри и "systematic" - систематический):

$$S_{wsy}^2 = \frac{1}{k(n-1)} \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2;$$

- y_{ij} - j -й член i -й систематической выборки; $j = 1, 2, \dots, n, i = 1, 2, \dots, k$;
 \bar{y}_i - среднее значение i -й выборки.

Многоступенчатый или многошаговый отбор (подвыборки). При организации статистических выборочных обследований широко применяется метод многоступенчатого отбора. Если исследуемая совокупность содержит некоторые группы и имеется информация о принадлежности элементов к той или иной группе, то в этом случае при выборочных обследованиях может быть удобным вначале осуществить случайную выборку из этих групп, а затем в целях экономии средств и времени не проводить обследование всех единиц отобранных групп, как при гнездовом отборе, а отобрать лишь часть элементов в каждой выбранной группе, т.е. осуществить двухступенчатый отбор. Таким образом, при многоступенчатом отборе извлечение единиц наблюдения осуществляется после нескольких последовательных случайных отборов групп.

Для примера многоступенчатого отбора рассмотрим его специальный случай - двухэтапный групповой отбор, в котором элементы второго этапа - обследуемые единицы, случайным образом отобранные из элементарных единиц выбранных групп.

Таблица 5

**ФОРМУЛЫ ОЦЕНИВАНИЯ ПРИ МНОГОСТУПЕНЧАТОМ
(ДВУХЭТАПНОМ ГРУППОВОМ) ОТБОРЕ**

Статистические показатели	Истинное значение	Оценка
Суммарное значение признака	$Y = \sum_{i=1}^N Y_i;$ $Y_i = \sum_{j=1}^{M_i} Y_{ij}$	$\hat{Y} = \frac{N}{n} \sum_{i=1}^n \frac{M_i}{m_i} y_i;$ $\hat{Y}_i = \frac{M_i}{m_i} y_i$
Количество элементов в генеральной совокупности	$M = \sum_{i=1}^N M_i$	$\hat{M} = \frac{N}{n} \sum_{i=1}^n M_i$
Среднее значение признака	$\bar{Y} = \frac{Y}{M};$ $\bar{Y}_c = \frac{1}{N} \sum_{i=1}^N Y_i;$ $\bar{Y}_i = \frac{1}{M_i} \sum_{j=1}^{M_i} Y_{ij}$	$\bar{y} = \frac{\hat{Y}}{\hat{M}};$ $\bar{y}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} y_{ij}$
Количество элементов в среднем на группу	$\bar{M} = \frac{1}{N} \sum_{i=1}^N M_i = \frac{M}{N}$	$\bar{m} = \frac{1}{n} \sum_{i=1}^n m_i$

Продолжение таблицы 5

Статистические показатели	Истинное значение	Оценка
Дисперсия оценки суммарного значения признака	$V(\hat{Y}) = K_1 S_b^2 + K_2 S_{wi}^2$, где $K_1 = \frac{N^2(N-n)}{Nn}$; $K_2 = \frac{N}{n} \sum_{i=1}^N \frac{M_i^2(M_i - m_i)}{M_i m_i}$	$v(\hat{Y}) = K_1 s_b^2 + K_2 s_{wi}^2$, где $K_1 = \frac{N^2(N-n)}{Nn}$; $K_2 = \frac{N}{n} \sum_{i=1}^n \frac{M_i^2(M_i - m_i)}{M_i m_i}$
Дисперсия оценки среднего значения признака	$V(\bar{y}) = \frac{1}{M^2} [K_1 S_b^2 + K_2 S_{wi}^2]$	$v(\bar{y}) = \frac{1}{M^2} [K_1 s_b^2 + K_2 s_{wi}^2]$
Стандартная ошибка оценки суммарного значения признака	$S(\hat{Y}) = \sqrt{K_1 S_b^2 + K_2 S_{wi}^2}$	$s(\hat{Y}) = \sqrt{K_1 s_b^2 + K_2 s_{wi}^2}$
Стандартная ошибка оценки среднего значения признака	$S(\bar{y}) = \frac{1}{M} \sqrt{K_1 S_b^2 + K_2 S_{wi}^2}$	$s(\bar{y}) = \frac{1}{M} \sqrt{K_1 s_b^2 + K_2 s_{wi}^2}$
Коэффициент вариации оценки	$CV = \frac{S(\hat{Y})}{\hat{Y}} 100\% = \frac{S(\bar{y})}{\bar{Y}} 100\%$	$cv = \frac{s(\hat{Y})}{\hat{Y}} 100\% = \frac{s(\bar{y})}{\bar{y}} 100\%$

В таблице 5 использованы следующие обозначения:

- N - количество групп в генеральной совокупности;
- i - номер группы;
- Y_i - сумма признака в i -й группе совокупности;
- M_i - объем в i -й группе, $i=1, \dots, N$;
- j - номер элемента в группе;
- Y_{ij} - величина признака для j -го элемента в i -й группе совокупности;
- n - количество выбранных групп;
- m_i - количество выбранных элементов в i -й выбранной группе, $i=1, \dots, n$;

- $y_i = \sum_{j=1}^{m_i} y_{ij}$ - суммарное значение признака по выборке в i -й выбранной группе;
- y_{ij} - величина признака для j -го выбранного элемента в i -й группе;
- \bar{Y}_c - среднее значение признака по совокупности;
- \bar{Y}_i - средняя величина характеристики признака в i -й группе совокупности;
- \hat{Y}_i - несмещенная оценка для Y_i ;
- \bar{y}_i - среднее значение характеристики выбранного элемента в i -й выбранной группе;
- S_b^2 - дисперсия совокупности между группами:
- $$S_b^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y}_c)^2 ;$$
- S_{wi}^2 - дисперсия внутри i -й выбранной группы:
- $$S_{wi}^2 = \frac{1}{M_i-1} \sum_{j=1}^{M_i} (Y_{ij} - \bar{Y}_c)^2 ;$$
- s_b^2 - несмещенная оценка дисперсии совокупности между группами:
- $$s_b^2 = \frac{1}{n-1} \sum_{i=1}^n (M_i y_i - \frac{1}{n} \sum_{i=1}^n M_i \bar{y}_i)^2 ;$$
- s_{wi}^2 - несмещенная оценка дисперсии внутри i -й выбранной группы:
- $$s_{wi}^2 = \frac{1}{m_i-1} \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2 .$$

В случае, если на каждой ступени сохраняется одна и та же единица отбора, говорят о **многофазном отборе**. Многофазный отбор широко применяется в выборочных переписях населения, когда одна и та же совокупность обследуется на различных фазах отбора по разным, обычно расширяющимся от фазы к фазе, программам наблюдения.

Квазислучайные выборки

В выборках квазислучайного типа предполагается наличие вероятностного отбора на том основании, что специалист, рассматривающий выборку, считает это допустимым (т.е. предполагается, обстоятельства таковы, что возможно рассматривать выборку как вероятностную). Обоснованность этого решения всецело зависит от обстоятельств, поэтому делать обобщения сложно.

Примером использования квазислучайной выборки в статистической практике является "Выборочное обследование малых предприятий по изучению социальных процессов в малом предпринимательстве", проведенное в 1996 г. в некоторых регионах России. Единицы наблюдения (малые предприятия) отбирались экспертно с учетом представительства отраслей экономики из уже сформированной выборки обследования финансово-хозяйственной деятельности малых предприятий (форма № МП "Сведения об основных показателях финансово-хозяйственной деятельности малого предприятия"). При обобщении выборочных данных предполагалось, что выборочная совокупность сформирована методом простого случайного отбора.

В случае отбора невероятностным способом также имеются два типа выборок: отобранных в соответствии с определенной целью и отобранных на основе суждения эксперта. Отнесение выборки к тому или иному типу зависит от того, является ли процедура отбора объективной или нет.

Выборки, построенные на основе суждения эксперта

Прямое использование суждения, вероятно, является наиболее общим методом намеренного включения единиц в выборку. Примером такого способа отбора является монографический метод, предполагающий получение информации только от одной единицы наблюдения, являющейся типичной по мнению организатора обследования - эксперта.

По сравнению с вероятностными, выборки, построенные на основе суждения, наилучшим образом проявляют себя там, где:

- а) выборка мала;
- б) исследуемая совокупность весьма невелика и обзрима, или известна организатору наблюдения;
- в) исследуемое свойство элементов общей совокупности существенно варьирует;
- г) специалист, формирующий выборку, является большим и признанным мастером своего дела.

Выборки, построенные с определенной целью

Выборки, построенные с определенной целью, извлекаются с помощью объективной процедуры, но без использования вероятностного механизма. Существует значительное число разнообразных способов отбора, приводящих к целевым выборкам.

Широко известен метод основного массива, при котором в выборку включаются наиболее крупные (существенные) единицы наблюдения, обеспечивающие основной вклад в показатель, например, суммарное значение признака, представляющего основной интерес обследования.

В заключение можно отметить, что целевые выборки несколько разочаровывают своими результатами по сравнению с вероятностными выборками, однако при работе с малыми выборками имеет смысл рассматривать виды отбора, приводящие к формированию целевой выборки. Также в мировой статистической практике существуют факты, свидетельствующие о том, что при очень небольших объемах выборки на основе суждения могут иметь преимущества перед, например, простыми случайными выборками. Возможно при некоторых условиях это справедливо и для целевых выборок.

"Квазислучайные" выборки нельзя сравнить с остальными тремя типами выборок каким-либо простым и удовлетворительным способом. Каждая квазислучайная выборка может оказаться уникальной. Можно сделать только единственное обобщение – такими выборками следует пользоваться крайне осторожно и обоснованно.

1.1.3. КОМБИНИРОВАННЫЙ МЕТОД НАБЛЮДЕНИЯ

Под **комбинированным методом статистического наблюдения** понимается сочетание использования сплошного и выборочного методов. Комбинирование обоих методов наблюдения имеет два аспекта: чередование во времени и одновременное их использование (часть совокупности наблюдается на сплошной основе, а часть - выборочно).

Чередование периодических выборочных обследований со сравнительно редкими сплошными или переписями необходимо для уточнения состава исследуемой совокупности и получения значений основных классифицирующих и количественных признаков для всех единиц. В дальнейшем эта информация используется как статистическая основа выборочного наблюдения. От наличия такой информации зависит качество создаваемых методик, а в итоге – точность и достоверность результатов, получаемых в будущем на выборочной основе. Примером здесь служат переписи населения и выборочные обследования домашних хозяйств в промежутках времени между ними.

В данном случае предмет исследования состоит в установлении состава признаков сплошного наблюдения, обеспечивающих организацию репрезентативной выборки, и оптимальных периодов чередования, т.е. когда сплошные данные теряют актуальность и требуются затраты на их обновление.

Одновременное использование в рамках одного обследования сплошного и выборочного наблюдения обусловлено неоднородностью встречающихся в статистической практике совокупностей.

В особенности это справедливо для обследований экономической деятельности совокупности предприятий, для которой характерны скошенные распределения изучаемых признаков, когда некоторое число единиц имеет характеристики, сильно отличающиеся от основной массы значений. В этом случае такие единицы наблюдаются на сплошной основе, а другая часть совокупности - выборочно.

Предмет исследования в случае одновременного использования сплошного и несплошного наблюдения - установление их оптимальной пропорции и разработка способов оценки точности результатов.

Типичным примером, иллюстрирующим данный аспект применения комбинированного метода статистического наблюдения, является общий принцип проведения обследований совокупности предприятий Российским статистическим агентством, в соответствии с которым обследования совокупности крупных и средних предприятий проводятся преимущественно сплошным методом, а обследования малых предприятий – выборочным.

Таким образом, использование Российским статистическим агентством комбинированного метода статистического наблюдения во взаимосвязи с современной системой статистических показателей даёт возможность комплексно оценить происходящие социально-экономические процессы, тенденции и перспективы их развития в России и ее регионах.

1.2. ОСНОВНЫЕ ЭТАПЫ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ И ВЫБОР МЕТОДА ЕГО ПРОВЕДЕНИЯ

Можно выделить два основных класса проблем в планировании обследований, которые влияют на выбор метода проведения обследования и состав этапов наблюдения:

1. определение качества результатов, которые будут получены при применении той или иной из альтернативных схем наблюдения;
2. определение стоимости каждой схемы.

В целом этапы статистического наблюдения и метод его проведения определяются исходя из цели и задач обследования, характера изучаемой совокупности, наличия и состояния статистической основы, возможных затрат на его организацию (включая методологическое исследование, разработку программно-технологических средств, обучение персонала), на сбор и обработку первичной информации.

Статистическое наблюдение имеет следующие основные методологические и практические этапы его проведения:

- конкретизация цели и формулировка задач статистического исследования;
- составление программы наблюдения и программы разработки полученной статистической информации;
- ограничение обследуемой совокупности (установление ее рамок);
- создание статистической основы, определение ее единицы и единицы наблюдения;
- выбор метода наблюдения, для выборочного наблюдения – установление оптимального процента отбора;
- формирование выборочной совокупности;
- сбор и редактирование первичной информации, включающее логический и арифметический способы контроля, основанные на соотношениях между связанными пунктами вопросника и допустимыми значениями статистических признаков, представленных респондентами;
- восстановление неполноты первичных данных, связанное с частичными и полными неответами респондентов, т.е. с незаполнением некоторых пунктов вопросника или отсутствием его как такового;
- в случае выборочного наблюдения распространение его результатов на обследуемую совокупность и определение ошибок выборки;
- сводка и группировка данных, расчет обобщающих статистических показателей;
- коррекция итоговых показателей обследования на несоответствие использованной статистической основы наблюдения исследуемой совокупности (например, досчет на неучтенную экономическую деятельность);
- анализ и публикация результатов.

Основные этапы статистического наблюдения содержат не только методологические и практические аспекты по непосредственному сбору и обработке данных. Знание общего решения проблемы является недостаточным. Необходимо принимать во внимание стоимость различных операций и их возможную эффективность, так как в практической деятельности всегда нужно учитывать ограниченность бюджета обследования, времени проведения и вполне возможную недостаточную надежность персонала.

Выбор метода проведения статистического наблюдения неоднозначен. Основными факторами, которые необходимо учитывать при его определении, являются следующие:

- материальные и трудовые ресурсы, которыми ограничены органы статистики;
- состав решаемых статистических задач;
- количество единиц наблюдения;

- необходимая точность получаемых результатов.

Первый вопрос, на который надо дать ответ, – возможно ли проведение наблюдения на сплошной основе? Это, в свою очередь, предполагает решение следующих вопросов:

- какое количество единиц наблюдения необходимо обследовать, чтобы решить поставленные задачи;

- обладает ли организатор необходимыми ресурсами для этого и целесообразно ли их затратить для достижения цели обследования.

Если проведение **сплошного** наблюдения невозможно, например, из-за высокого уровня затрат или нецелесообразно, например, по причине цели обследования, возможности достижения приемлемого результата при обследовании лишь некоторой части исследуемой совокупности, нужно выяснить возможность применения квазисплошного метода – ценза (основного массива).

При применении **цензового** метода наблюдения нужно сосредоточить основное внимание на качестве используемой статистической основы – ее актуальности и достаточно полном охвате исследуемой совокупности.

При наличии актуальных данных по интересующему признаку(кам) (или некоторому связанному с ним(и)) по единицам наблюдения, например, данных переписи, может оказаться, что погрешность оценок показателей, представляющих главный интерес обследования, рассчитанных на основе данных ограниченного количества единиц наблюдения (например, 10% наиболее крупных единиц), пренебрежимо мала либо легко корректируется на ненаблюдаемую часть исследуемой совокупности. В этом случае применение ценза оправдано.

Если применение ценза нецелесообразно, то необходимо проработать возможность применения методологически сложного решения - **выборочного** метода.

В первую очередь, требуется определение типа выборки. Если организатор обследования, исходя из цели и задач обследования, применяет не вероятностную выборку, т.е. экспертно включает (объективно или нет) единицы наблюдения в выборочную совокупность, то он принимает на себя ответственность за способ расчета и качество конечных результатов, количественное описание которого затруднено.

При принятии решения об использовании вероятностной выборки необходимо учитывать, что существуют два подхода к решению этой задачи - модельный и плано-обоснованный.

В первом случае по имеющейся базовой информации и собранным выборочным данным строится модель распределения наблюдаемых признаков, в соответствии с которой определяются интересующие статистические показатели.

Во втором случае рассматриваются всевозможные выборки, извлеченные из перечней, содержащих некоторое конечное число N единиц. Число различных выборок объема n , которые могут быть извлечены из списка N единиц, равно C_N^n . Фактически на практике оперируют вероятностями отбора единиц и вероятностями извлечения их пар из основы. Для распространения выборочных данных и расчета соответствующих характеристик точности используются выборочные веса.

И, наконец, следует отметить, что в статистической практике часто используется комбинация сплошного и несплошного метода наблюдения, когда часть совокупности наблюдается полностью, а часть - выборочно.