

Выборочное наблюдение

Не всегда источником статистической информации служат данные сплошного наблюдения, чаще всего прибегают к выборочному методу наблюдения.

Генеральной совокупностью называется вся совокупность единиц изучаемого социально-экономического явления или процесса.

Выборочной совокупностью называется совокупность единиц, отобранных для выборочного наблюдения.

Выборочное наблюдение представляет собой одну из форм несплошного статистического наблюдения, при котором обобщающие характеристики изучаемой совокупности устанавливаются на основе отобранной случайным образом части единиц генеральной совокупности.

Преимущества проведения выборочного наблюдения по сравнению со сплошным наблюдением

- ❖ *быстрота получения результатов обследования.*
- ❖ *значительное снижение стоимости обследования.*
- ❖ *возможность лучшей организации проведения обследования, и, как следствие, – повышение достоверности получаемых результатов.*
- ❖ *возможность расширения программы наблюдения.*
- ❖ *возможность использования выборочного метода в тех случаях, когда проведение сплошного наблюдения методологически невозможно.*

Выборочный метод имеет определенный **недостаток**, связанный с наличием ошибок репрезентативности, который определенным образом компенсируется за счет снижения ошибок регистрации (возможность привлечения более квалифицированного персонала, более тщательный контроль исходных данных и т.п.).

По виду различают индивидуальный, групповой и комбинированный отбор.

При **индивидуальном отборе** в выборочную совокупность отбираются отдельные единицы генеральной совокупности, например, предприятия, при обследованиях населения – конкретные люди и т.д.

При **групповом отборе** – группы единиц, например, бригады, микрорайоны.

Комбинированный отбор предполагает сочетание индивидуального и группового отбора, например, сначала отбираются группы (групповой отбор), а в них случайным образом отдельные единицы (индивидуальный отбор).

При осуществлении этих видов отбора возможно использование одной из двух схем: применение бесповторного отбора или повторного.

Бесповторным называется такой отбор, при котором попавшая в выборку единица наблюдения не возвращается в совокупность и не может быть отобранной второй раз.

Повторным называется такой отбор, при котором попавшая в выборку единица наблюдения возвращается в совокупность и может быть отобранной во второй раз.

В зависимости от различных схем и способов отбора различают следующие виды выборок:

**Собственно-случайная выборка,
Механическая,
Типическая,
Серийная.**

Собственно случайная выборка

Отбор единиц при использовании собственно случайной выборки производится наугад, наудачу, путем жеребьевки или по таблице случайных чисел. При этом все единицы совокупности должны иметь равные шансы попасть в выборочную основу.

Механическая выборка

При механической выборке все единицы генеральной совокупности нумеруются числами от 1 до N , после чего отбираются каждые (N/n) -е объекты для выборки, находящиеся на равном расстоянии друг от друга.

Величина N/n называется **шагом** или интервалом отбора.

Например, если для 1500 единиц требуется создать 10-% выборку, соответственно объемом в 150 единиц, то в нее попадет каждый 10-й элемент, отобранный механически через определенный интервал совокупности ($150/1500=10$).

Существуют два способа формирования основы выборки: по *неранжированным* данным и по *ранжированной* генеральной совокупности.

В первом случае результаты механического отбора будут являться реализацией случайного бесповторного отбора, так как единицы наблюдения располагаются в случайном порядке. Усилить эту «случайность» можно, выбрав начальную точку отсчета случайным образом из интервала, соответствующему первому шагу отбора.

Во втором случае единицы наблюдения определенным образом упорядочиваются (ранжируются) по величине изучаемого или коррелирующего с ним признака и отбор осуществляется в соответствии с его шагом N/n , начиная с единицы, являющейся серединой первого интервала (шага отбора).

Типическая выборка

При значительной колеблемости признака в генеральной совокупности, например, при обследованиях предприятий, когда представители различных отраслей значительно отличаются друг от друга, ее целесообразно предварительно разбить на однородные типы или группы, а затем провести случайный (или механический) отбор единиц наблюдения внутри полученных групп. Извлеченная подобным образом выборка будет **типической** (в зарубежной и переводной литературе — «расслоенной», или «стратифицированной», выборкой).

Серийная (гнездовая) выборка

Если генеральную совокупность можно разделить на одинаковые по объему и однородные между собой группы, то производят отбор не единиц наблюдения, а их серий, после чего сами единицы обследуются сплошь.

Например, при оценке качества продукции можно отбирать партии товара, а затем на сплошной основе обследовать входящие в них изделия. Данная выборка будет являться **серийной**. Отбор серий может быть как неповторный, так и повторный.

В отличие от типического отбора, когда выделенные группы систематически различаются между собой, серийный отбор базируется на том предположении, что отличие единиц разных серийных групп является случайным, несистематическим и при изучении основных тенденций и закономерностей этим отличием можно пренебречь.

Порядок расчетов при проведении собственно случайной и механической выборки

1. Рассчитывается выборочное среднее (выборочная доля).
2. Рассчитывается средняя ошибка выборки:

$$\mu = \sqrt{\frac{\sigma^2}{n}} - \text{при повторной случайной выборке}$$

$$\mu = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)} - \text{при бесповторной случайной}$$

и механической выборке

3. Рассчитывается предельная ошибка выборки:

$$\Delta = t\mu$$

Значение t определяется по таблице нормального распределения.

Значение доверительной вероятности P	0,683	0,866	0,954	0,988	0,991	0,997	0,999
Значение коэффициента доверия t	1,0	1,5	2,0	2,5	2,6	3,0	4,0

4. Строятся доверительные интервалы:

при оценивании
среднего
значения

$$\tilde{x} - \Delta_{\tilde{x}} \leq \bar{x} \leq \tilde{x} + \Delta_{\tilde{x}}$$

при оценивании доли

$$w - \Delta_{\tilde{w}} \leq p \leq w + \Delta_{\tilde{w}}$$

Задача

Для изучения уровня ликвидности 225 предприятий-заемщиков было проведено выборочное обследование 90 предприятий методом случайной бесповторной выборки, в результате которого получены следующие данные:

Коэффициент ликвидности, x_i	Количество предприятий, f_i
До 1,4	13
1,4–1,6	15
1,6–1,8	17
1,8–2,0	15
2,0–2,2	16
2,2 и выше	14
Итого	90

1. Выборочная средняя:

$$\tilde{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{1,3 \cdot 13 + 1,5 \cdot 15 + \dots + 2,3 \cdot 14}{90} = 1,81$$

2. Выборочная дисперсия:

$$\begin{aligned} \sigma^2 &= \frac{\sum (x_i - \tilde{x})^2 f_i}{\sum f_i} = \\ &= \frac{(1,3 - 1,81)^2 \cdot 13 + (1,5 - 1,81)^2 \cdot 15 + \dots + (2,3 - 1,81)^2 \cdot 14}{90} = \\ &= 0,1095 \end{aligned}$$

3. Определяем среднюю ошибку бесповторной случ. выборки:

$$\mu_{\tilde{x}} = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{0,1095}{90} \left(1 - \frac{90}{225}\right)} = 0,027.$$

4. При вероятности равной 0,954 величина предельной ошибки выборки составит:

$$\Delta_{\tilde{x}} = t\mu_{\tilde{x}} = 2 \times 0,027 = 0,054.$$

5. Доверительные границы для среднего:

$$\begin{aligned} \tilde{x} - \Delta_{\tilde{x}} &\leq \bar{x} \leq \tilde{x} + \Delta_{\tilde{x}} \\ 1,81 - 0,054 &\leq \bar{x} \leq 1,81 + 0,054 \\ 1,7756 &\leq \bar{x} \leq 1,864 \end{aligned}$$

Определение численности выборочной совокупности

Формулы для определения численности выборки выводятся из формул предельных ошибок в соответствии со следующими исходными положениями:

- ❖ вида предполагаемой выборки;
- ❖ способа отбора (повторный или бесповторный);
- ❖ оцениваемого параметра (среднего значения или доли).

Собственно случайная выборка:

повторная случайная:

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

бесповторная случайная
и механическая:

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

Пример.

Определим, сколько объектов из совокупности 507 промышленных предприятий следует проверить налоговой инспекции, чтобы с вероятностью 0,997 определить долю предприятий с нарушениями в уплате налогов. По данным прошлого аналогичного обследования величина среднего квадратического отклонения составила 0,15; размер ошибки выборки предполагается получить не выше, чем 0,05.

При использовании повторного случайного отбора следует проверить:

$$n = \frac{t^2 w(1-w)}{\Delta_w^2} = \frac{3^2 \times 0,15^2}{0,05^2} = 81$$

При бесповторном случайном отборе потребуется проверить:

$$n = \frac{t^2 w(1-w)N}{\Delta_w^2 N + t^2 w(1-w)} = \frac{3^2 \times 0,15^2 \times 507}{0,05^2 \times 507 + 3^2 \times 0,15^2} = 70$$