

ВЫБОРКА КЛАСТЕРНАЯ (ГНЕЗДОВАЯ) - метод извлечения выборки (см.), основанный на предварительном разделении генеральной совокупности (см.) на относительно компактные структурные части (кластеры, гнезда). Главным требованием является более широкая вариация основных изучаемых показателей внутри кластера по сравнению с их вариацией между кластерами [в отличие от выборки стратифицированной (см.), цель которой - выделение страт, в которых вариация основных показателей была бы минимальной].

В.К. может осуществляться в несколько этапов. На первом этапе основа выборки представляет собой полный список кластеров; из этого списка тем или иным способом извлекается выборка кластеров. Далее в исследовании участвуют только выбранные кластеры. Они, в свою очередь, могут обследоваться полностью или выборочно. Если отобранные кластеры обследуются полностью, мы получаем выборку серийную (см.).

Примером серийной выборки является опрос студентов целыми академическими группами, когда на первом этапе отбираются группы, а на втором - опрашиваются все студенты из отобранных групп. Если отобранные кластеры обследуются выборочно, выборка является двухступенчатой или многоступенчатой. Примером двухступенчатой выборки является исследование населения, на первом этапе которого в качестве кластеров отбираются отдельные населенные пункты; на втором - в каждом из отобранных населенных пунктов извлекается простая случайная выборка его жителей (в качестве основы выборки могут использоваться, например, данные адресного стола). Для многоступенчатой выборки процесс кластеризации может продолжаться "вглубь", посредством деления отобранных кластеров на все более мелкие.

Если на всех этапах В.К. применяются процедуры случайного (простого или стратифицированного) отбора либо сплошное обследование всех объектов, составляющих кластер, выборка называется случайной кластерной. Существуют две основные стратегии, обеспечивающие при кластерном отборе всем элементам из генеральной совокупности одинаковую вероятность попадания в выборку. Согласно первой из них, кластеры отбирают из списка с равными вероятностями; затем предполагаемый объем выборки делят между выбранными кластерами пропорционально их размеру. При использовании второй стратегии, кластеры отбирают с вероятностями, пропорциональными их размеру, предполагаемый объем выборки делится между отобранными кластерами в равных частях.

Главным преимуществом В.К. является минимизация трудовых и финансовых затрат; недостатком - более высокая ошибка выборки по сравнению с простым или стратифицированным случайным отбором.

Ошибка выборки тем выше, чем больше средний размер кластера и, соответственно, меньше число обследуемых кластеров. Кластерный отбор рекомендуется применять:

- 1) если основа выборки не может быть получена для всей генеральной совокупности, но может быть получена для ее отдельных структурных частей;
- 2) если простой или расслоенный случайный отбор становится неприемлемо дорогостоящим, например, из-за того, что генеральная совокупность рассеяна по слишком большой территории.

В некоторых случаях отбор кластеров производится нестатистическими методами. Например, целевым образом выбираются населенные пункты, в которых проживают интервьюеры. В этом случае выборка не будет случайной (даже если в населенных пунктах производится случайный отбор), и статистическое оценивание ошибки выборки не будет корректным.

О.В. Терещенко