

Выборочный метод (Sampling method)

Разделы: [Алгоритмы](#)

Выборочными называются методы математической статистики, в которых свойства генеральной совокупности изучаются на основе исследования свойств ее части, полученной с помощью определенного алгоритма выборки (сэмплинга), который обычно носит случайный характер.

Необходимость применения таких методов обусловлена либо невозможностью, либо нецелесообразностью исследования всей генеральной совокупности, что может быть связано с очень большим ее объемом, сложностью сбора данных и т.д.

Корректное применение выборочных методов является важным условием качественного анализа на основе любого статистического исследования и во многом определяется особенностями предмета анализа.

Среди оцениваемых с помощью выборочных методов характеристик чаще всего фигурируют:

- дисперсия признака;
- математическое ожидание признака;
- коэффициент корреляции;
- доля объектов совокупности с тем или иным значением признака (например, доля клиентов, готовых приобрести новый продукт компании);
- средняя величина признака (например, уровень продаж) и др.

Главной проблемой выборочного метода является оценка того, насколько уверенно можно по свойствам выборки судить о свойствах генеральной совокупности в целом. Всякая подобная оценка неизбежно имеет вероятностный характер, и задача сводится к тому, чтобы вероятность правильных выводов (точности статистических оценок) была бы максимальной.

Особенно широко выборочные методы применяются в демографических и социологических исследованиях, выборочном контроле качества изделий и других предметных областях, где специалисту приходится иметь дело с огромным количеством объектов или наблюдений.

Большинство методов Data Mining (нейронные сети, деревья решений и др.) используют этот подход, поскольку построение и обучение аналитических моделей всегда производится на ограниченной выборке данных, в то время как закономерности, обнаруженные моделью, должны распространяться на всю генеральную совокупность.

Выборочный метод реализуется в Logipom с помощью обработчика сэмплинг. Сэмплинг является одним из этапов выборочного метода и осуществляет различные способы отбора записей в выборку из исходного набора данных для дальнейшего наблюдения за поведением тех или иных характеристик.