**Лекция 13-14. Понятие ошибки выборки. Случайные ошибки.**

Итак, репрезентативность – это свойство выборки отражать характеристики генеральной совокупности. Но если совпадений нет, говорят об ошибки выборки, т.е. отклонении статистической структуры выборки от структуры генеральной совокупности.

Например, в среднем студенты КГЭУ тратят на подготовку к занятиям 5 часов в неделю, а в выборке -8 часов, это означает, что социологи опросили более добросовестных студентов, произошел перекос и можно говорить об ошибки репрезентативности.

Ошибка репрезентативности – это расхождение между значениями генеральной совокупности и выборочной.

В отечественной литературе наряду с термином «ошибка репрезентативности» встречается и другой – «ошибка выборки». Обычно они используются как синонимы, но понятие «ошибка выборки» количественно более точное, чем «ошибка репрезентативности».

Ошибка выборки – это «отклонение средних характеристик выборочной совокупности от средних характеристик генеральной совокупности. На практике она определяется путём сравнения известных характеристик генеральной совокупности с выборочными средними». Средние характеристики чаще берут из данных переписей населений, статистических данных, результатов предшествующих исследований.

Так специалисты ВЦИОМ используют для сравнений такие показатели как пол, возраст, образование, семейное положение, которые заимствуют в Госкомстате и на этапе полевой работы проводят тщательный ремонт выборки, чтобы минимизировать ошибку.

Ошибка выборки обусловлена двумя факторами: методом формирования выборки и ее объемом.

Ошибку выборки установить невозможно, если неизвестны значения переменных в выборочной и генеральной совокупностях.

Репрезентативность выборки определяется двумя компонентами: систематическими и случайными ошибками.

Случайные ошибки связаны «со статистическими погрешностями (зависят от динамики исследуемых признаков) и непредвиденными нарушениями процедуры сбора информации (процедурные ошибки, допущенные при регистрации признаков)». Случайные ошибки уменьшаются с увеличением объема выборочной совокупности (См. Табл.).

*Таблица.* Зависимость объема выборки от ее ошибки(размер генеральной совокупности составляет 20 тыс. ед.)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Ошибка**выборки,**% | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 14 | 17 | 20 |
| **Объем выборки, ед.** | 2500 | 1100 | 620 | 400 | 280 | 200 | 160 | 110 | 100 | 67 | 50 | 30 | 25 |

Случайную ошибку можно измерить методами математической статистики, если при формировании выборочной совокупности соблюдался принцип случайности, обеспечивающийся строго определенными правилами, которые составляют метод формирования выборочной совокупности, и устранить.

На практике принцип случайности соблюсти очень сложно, а иногда просто невозможно, что приводит к появлению систематической ошибки, которые возникают «из-за неполной объективности выборки генеральной совокупности (недостаток информации о генеральной совокупности, отбор наиболее «удобных» для исследования элементов генеральной совокупности), а так же из-за несоответствия выборки целям и задачам исследования». Иногда такие ошибки называют ошибками смещения. Они возникают при различных телевизионных опросах, когда телеведущий предлагает телезрителям позвонить по определённым номерам телефонов, послать смс-сообщение и высказать своё мнение по какой-то проблеме. Естественно мы не можем утверждать, что эти люди отражают мнение всего населения страны, и даже телеаудитории. Вероятнее всего в таких опросах участвуют более образованные и активные люди, чем вся генеральная совокупность, поэтому любой телевизионный опрос содержит в себе систематическое искажение и носит поверхностный характер.

Но систематические ошибки возникают и в ходе корректно организованного опроса. Например, на улице на вопросы интервьюера отвечают только те, кто никуда не спешит. Искажения можно избежать, если соблюдать принципы случайного отбора и опрашивать, к примеру, каждого десятого прохожего.

Причины возникновения систематических ошибок:

1. в ходе исследования была неправильно составлена основа выборки (использовались устаревшие, неполные данные либо отсутствовала статистика по некоторым важным для формирования выборки признакам),

2. неудачно выбран способ отбора единиц наблюдения,

3. часть респондентов по разным причинам «выпала» из опроса (отсутствовала, отказалась отвечать) и так далее.

При помощи математических средств такие ошибки устранить невозможно, поэтому необходимо осуществить логический анализ причин появления систематических ошибок и разработать меры, которые смогли бы их устранить. «Величину ошибок смещения определить при помощи математических формул практически не возможно, поэтому они автоматически переходят на результаты и выводы исследования.

Ошибки смещения бывают обычно следствием:

– неверных исходных статистических данных о параметрах контрольных признаков генеральной совокупности;

– слишком малого (статистически не значимого) объёма выборочной совокупности;

– неверного применения способа отбора единиц анализа (например, отбор из неверно составленного списка, неудачный выбор места и времени проведения опроса.