

Лекция 11.

Тема :

Методы сбора и статистической обработки информации

План лекции 11:

- **Типы отбора и требования к составлению выборки.**
- **Способы обработки выборки данных.**
- **Задачи геоэкологических исследований, решаемые с применением статистических анализов.**

Выборка

Совокупность всех возможных наблюдений за объектом называют **генеральной совокупностью (ГС)**

Подбор характерных ключей, ограничений, называют **выборкой** из ГС

Выборка должна отвечать критерию **репрезентативности** - наиболее точно отражать условия ГС

Тип отбора	Вид выборки
Направленный (типический)	Элементарная или «ключевых участков»
Случайный	Простая случайная, по схеме случайных чисел
	Простая систематическая, составленная регулярным способом
Смешанный	Типическая стратифицированная

Характеристика выборки данных

Показатели среднего положения	мода, медиана среднее арифметическое ошибка среднего арифметического
Показатели разнообразия признаков	дисперсия коэффициент вариации
Группировка вариационного ряда	границы и середина классов частота частость
Графическое представление группировки	полигон, гистограмма, огива

Задачи, решаемые с применением статистического анализа

Вид статистического анализа	Тип решаемой задачи
<i>Корреляционный</i>	показ формы и тесноты связи между двумя явлениями
<i>Регрессионный</i>	математическое описание зависимости между явлениями, прогнозирование значений признака
<i>Факторный</i>	выделение из множества характеристик сложного явления главных компонент
<i>Кластерный</i>	применяется для группировки вариантов по сходству
<i>Дисперсионный</i>	определяет степень влияния условий эксперимента на изучаемый объект

Ранговая корреляция

(цит. По Гагина Н.В. Методы геоэкологических исследований. Практикум. 2007., 47 с).

- расчет ранговой корреляции применяется в тех случаях, когда не удастся получить точных количественных значений;
- ранги - номера порядковых значений данных в возрастающей или убывающей последовательности;
- ранги показывают лишь положение вариант относительно друг друга;
- для расчета ранговой корреляции объем выборки не должен быть меньше 5;
- значения r менее 0,5 корреляционная зависимость между показателями слабая, 0,5-0,7 – средняя, 0,7-1,0 сильная.

$$r = 1 - \frac{6 \sum (x - y)^2}{N^3 - N}$$