

Задание ПЗ 1.

1. Найти выборку $n \geq 40$ данных любого количественного показателя. (Используя белстат, нацбанк, мировой банк и др.) Выборка должна быть из пространственных данных. Если временной ряд- уточняем можно или нет (курсы валют - сразу нет).

2. Провести статистический анализ данных.

- мода, медиана, среднее, дисперсия, коэффициент вариации;
- оценить однородность выборки используя коэффициент вариации;
- убрать выбросы если есть (сравниваем с медианой – разница более чем в 3 раза);
- разбить выборку на интервалы, количество интервалов $s = 1 + [\log_2 n]$ или $s = 1 + [3,322 \lg n]$ (формула Стерджесса)
- построить полигон и гистограмму, по их виду сделать предположение о виде распределения выборки;
- посчитать числовые характеристики: среднее, дисперсию, стандартное квадратическое отклонение;

3. Проверить соответствие выборки закону распределения, которое предположили в п.2 (используя критерии Пирсона, Колмогорова, ...)

4. Посчитать групповые дисперсии, внутригрупповую дисперсию, межгрупповую дисперсию. Проверить правило сложения дисперсий. (для выборки разбитой на s интервалов)

- разбить выборку на $s-1$ и $s+1$ интервалов. Посчитать дисперсии из предыдущего пункта. Оценить какое разбиение лучше (доля межгрупповой дисперсии в общей)