

**ОП «Политология», 2018-19****Математика и статистика, часть 2****Выборки из случайных величин (16.04.2018)***А. А. Макаров, А. А. Тамбовцева, Н. А. Василёнок***Задача 1.** Дана выборка:

8; 1; -2; 3; 10; 7; 12; 0; 4

- (a) Запишите вариационный ряд.
- (b) Найдите медиану выборки. Проинтерпретируйте полученное значение.
- (c) Найдите нижний и верхний квартили выборки. Проинтерпретируйте полученные значения. Найдите межквартильный размах.
- (d) Проверьте, есть ли в выборке нетипичные наблюдения.

**Задача 2.** Дана выборка:

3; 2; 6; 21; 7; 0; 8; -0.5; 2; -11

- (a) Найдите медиану выборки. Найдите среднее значение выборки.
- (b) Найдите нижний и верхний квартили выборки и межквартильный размах.
- (c) Проверьте, есть ли в выборке нетипичные наблюдения.
- (d) Если в выборке присутствуют нетипичные наблюдения, уберите их из выборки. Посчитайте медиану и среднее значение для новой выборки. Сделайте вывод.

**Задача 3.** Дана выборка:

25, 10, 15, 20, 30, 25, 15

Найдите выборочное среднее и выборочную дисперсию (несмещенную оценку). Найдите выборочное стандартное отклонение.

**Задача 4.** В таблице представлены значения явки на выборах депутатов Государственной Думы 18 сентября 2016 года по некоторым районам Москвы<sup>1</sup>. Взяв в качестве стартового значения 32%, постройте гистограмму для значений явки на выборах депутатов Государственной Думы, приняв шаг гистограммы (интервал группировки): а) равным 1; б) равным 2; в) равным 4. Какая гистограмма кажется вам наиболее информативной?

---

<sup>1</sup>Явка посчитана на основе данных ЦИК Российской Федерации

Район	Явка (в %)	Район	Явка (в %)
Алексеевский	38.19	Измайлово	35.79
Арбат	44.2	Красносельский	40.2
Аэропорт	37.31	Мещанский	38.49
Басманный	37.58	Сокольники	41.47
Беговой	35.75	Тверской	41.1
Вешняки	36.62	Текстильщики	35.24
Восточный	46.16	Царицыно	32.71
Гагаринский	37.65	Черёмушки	32.69
Даниловский	35.02	Щукино	39.16
Зюзино	32.41	Якиманка	38.42

**Задача 5.** Используя условие задачи 4, постройте «ящик с усами» для показателя явки.

**Задача 6.** Выполняется в R.

- Загрузите файл `Chile.csv` в R (RStudio).
- Выведите описательные статистики для всех переменных в датафрейме. Найдите (выборочные) медиану, нижний и верхний квартили следующих переменных: `age`, `population`, `income`. Сильно ли медианное значение этих переменных отличается от их средних значений?
- Постройте гистограммы для переменных `age` и `income`. Похоже ли распределение этих переменных на нормальное?
- Постройте «ящики с усами» для переменных `age` и `income`. Есть ли среди значений этих переменных нетипичные? Обоснуйте ваш ответ.

**Решение задач в R см. в материалах по R.**