Санкт-Петербургский Национальный Исследовательский Университет Информационных технологий, механики и оптики

**Лабораторная работа №5**

**Генерация случайных чисел и анализ выборки данных**

Выполнил: Долматов

Дмитрий Алексеевич

Проверила: Казанова

Полина Петровна

Санкт-Петербург

2021

**Цель работы:**

Изучить средства программы Microsoft Excel для генерации случайных чисел с требуемыми законами распределения, для построения и анализа выборок данных.

**Задачи:**

1. Построение выборки данных;
2. Анализ данных;
3. Составление гистограммы;
4. Изучение показателей уровня;
5. Изучение показателей рассеяния;
6. Изучение показателей асимметрии и элементов описательной статистики.

**Ход работы:**

Решим задачи, в которых необходимо использовать генерацию случайных чисел с требуемыми законами распределения и анализа выборки данных.

**Упражнение 1:**

Сформируем в столбцах массивы случайных чисел с числом переменных, равных 1, число случайных чисел = 100, с распределением Бернулли, если p = 0.3. Вторая таблица представляет биномиальное распределение с p = 0.85, число испытаний = 25. Третья таблица представляет нормальное распределение со средним – 100, а стандартным отклонением – 20. Данная таблица приведена на рисунке 5.1.1.

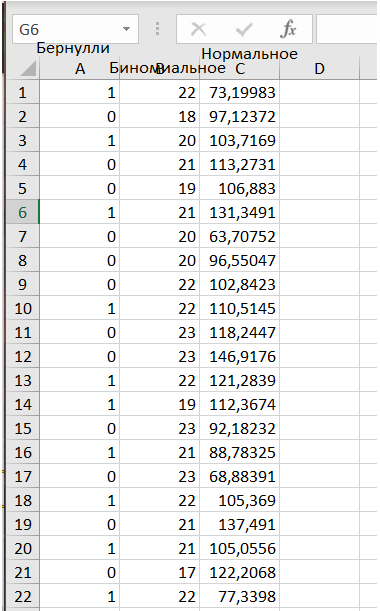


Рисунок 5.1.1 – Три столбца распределений (Бернулли, Биномиальное, Нормальное соответственно)

Затем на основе данных массивов получим случайные и систематические выборки размером 20 (ну или шагом в 5), представленные на рисунке 5.1.2.

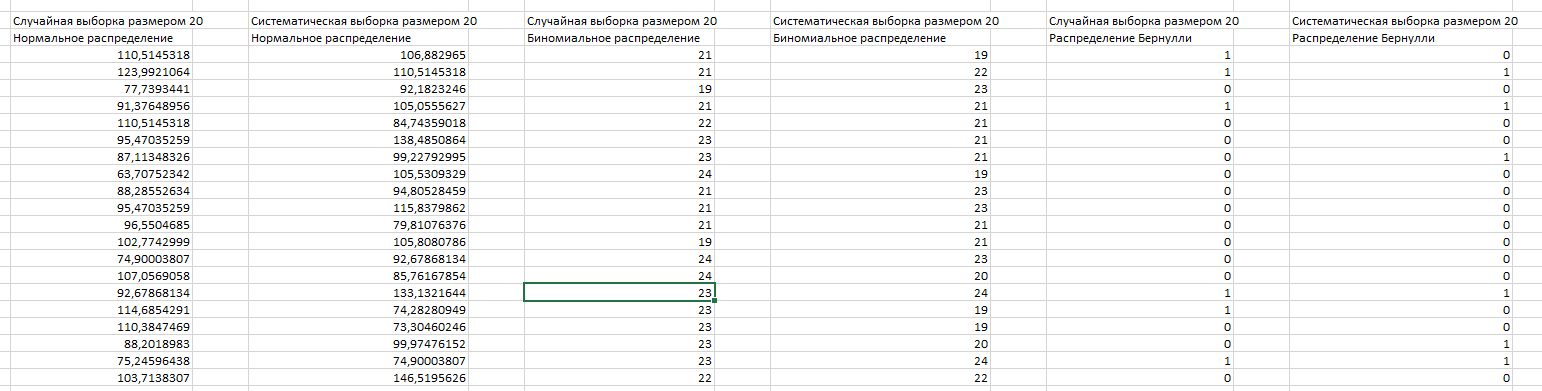


Рисунок 5.1.2 – Различные выборки размера 20

Следующим шагом построим гистограммы (столбчатая диаграмма частот), на оси абсцисс которой откладываются значения интервалов, а по оси ординат – частоты в виде столбиков, высота которых соответствует частоте попадания случайной величины в интервал.

Построим гистограмму для нормального распределения и выборок, полученных в п.2, представленных на рисунке 5.1.3.

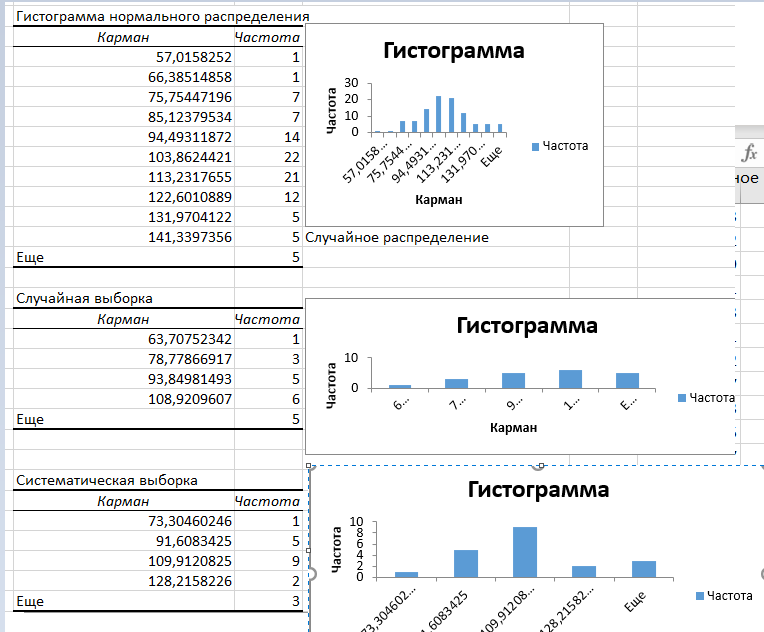


Рисунок 5.1.3 – Гистограммы нормального распределения и выборок

Самый корректный результат для нормального распределения показала систематическая выборка, для которой характерен пик на середине области определения (примерно).

Рассчитаем для случайных выборок, полученных ранее, показатели анализа данных: гистограмма, показатели уровня, показатели рассеивания и асимметрии.

Результаты приведены на рисунке 5.1.4 (перцентиль и ранг во вложении).

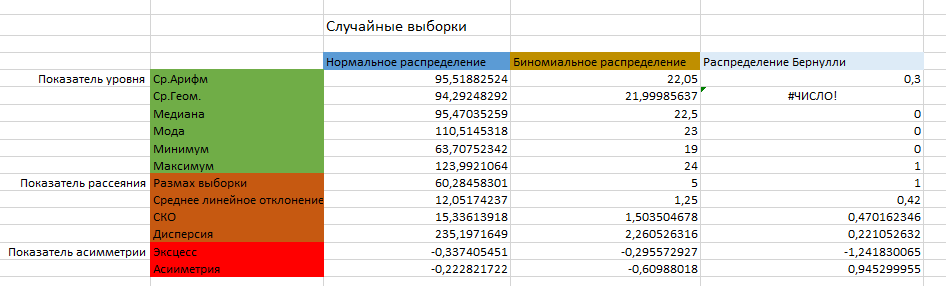


Рисунок 5.1.4 – Результаты нахождения аналитических данных из случайных выборок

Последней частью данной лабораторной работы является нахождение данных с помощью инструмента “Описательная статистика” для случайных выборок.

Результаты приведены на рисунке 5.1.5.

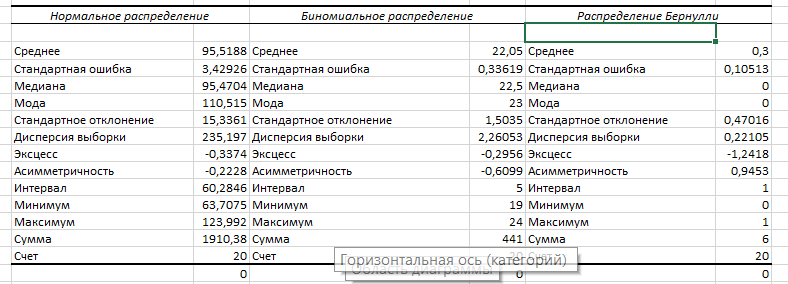


Рисунок 5.1.5 – Описательная статистика случайный выборок

Вывод заключается в том, что с помощью описательной статистики мы можем получить информацию, которая представлена на рисунке 5.1.4, более быстрым способом.

**Вывод:**

В результате выполнения лабораторной работы были изучены элементы генерации случайных чисел и анализ выборки данных, которая представлена показателями распределения данных, показателями уровней, показателями рассеивания и асимметрии.

**Контрольные вопросы**:

1. Закон распределения случайной величины – это любое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями.
2. Нормальный закон распределения случайных событий имеет место быть в природе – это погрешности измерений, отклонения при стрельбе, показатели живых популяций.
3. Показатель асимметрии характеризуем меру скошенности распределения. Если этот коэффициент больше нуля, то она правосторонняя, если меньше – левосторонняя. Показатель эксцесса характеризуем островершнинность распределения. Если он больше нуля, то распределение островершинное, если меньше – плосковершинное.
4. Этапы построения гистограммы состоят из входного интервала, когда мы вводим диапазон ячеек (массива). Затем указываем место, где будет указана таблица частот с помеченной ячейкой галочкой, которая покажет графическое представление гистограммы.
5. Отличие систематической выборки от случайной заключается в том, что у систематической выборки имеется определенный интервал (период) отбора элементов массива, однако случайная выборка избирается хаотично двумя способами: или с возможностью выбрать тот элемент, который уже был выбран ранее, либо без этой возможности.