|  |  |  |
| --- | --- | --- |
| ОТЧЕТ ЗАЩИЩЕН С ОЦЕНКОЙ: |  |  |

ПРЕПОДАВАТЕЛЬ:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| доцент, к.т.н., доцент | / |  | / |  | / | В. В. Мышко |
| (должность, учёная степень, звание) |  | (подпись) |  | (дата защиты) |  | (инициалы, фамилия) |

ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ №1

«СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ ЧИСЛОВЫХ ХАРАКТЕРИСТИК ЗАКОНОВ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН»

ПО КУРСУ: «ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ»

|  |  |  |  |
| --- | --- | --- | --- |
| РАБОТУ ВЫПОЛНИЛ СТУДЕНТ: | 4134К | / | Самарин Д. В. |
|  | (номер группы) |  | (инициалы, фамилия) |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | / |  | / | 18.02.2025 |
|  |  | (подпись студента) |  | (дата отчета) |

Цель

Целью лабораторной работы является получение оценки математического ожидания случайной величины по заданному массиву экспериментальных данных, построение 95%-го доверительного интервала, отсеивание аномальных наблюдений, уточнение оценки математического ожидания и проверка качества оценивания как по заданной доверительной вероятности, так и по максимальной вероятной погрешности.

Задание на лабораторную работу

На основе массива экспериментальных данных:

* найти оценку математического ожидания случайной величины;
* проверить качество оценивания по заданной доверительной вероятности;
* проверить качество оценивания по заданной максимальной вероятной погрешности.

Порядок выполнения задания:

1. Найти оценку математического ожидания по массиву экспериментальных данных (таблица 1.1)

2. Построить 95-ти процентный доверительный интервал для исследуемой случайной величины

3. Выполнить отсеивание аномальных наблюдений, не попавших в 95-ти процентный доверительный интервал

4. Найти уточнённую оценку математического ожидания после отсеивания аномальных наблюдений

5. Проверить качество оценивания математического ожидания:

1. по заданной доверительной вероятности (таблица 1.2) построить доверительный интервал для математического ожидания;
2. по заданной максимальной вероятной погрешности (таблица 1.2) найти доверительную вероятность попадания математического ожидания в интервал, определяемой указанной погрешностью.

Ход работы

Вариант 96

Массив экспериментальных данных:



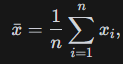
Доверительная и максимальная вероятности:



Решение:

**Оценка математического ожидания**

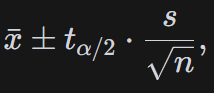
На основе экспериментальных данных было вычислено математическое ожидание:



Математическое ожидание: 3.7083333333333335

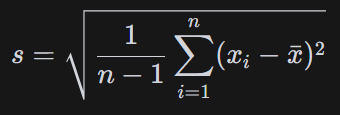
**Построение 95%-го доверительного интервала**

Используя стандартное отклонение, был рассчитан 95%-й доверительный интервал.



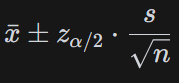
95%-й доверительный интервал: [0.6640213014491598, 6.752645365217507]

**Вычисление выборочного стандартного отклонения:**



Выборочное стандартное отклонение: 5.380513380935339

**Доверительный интервал:**



Доверительный интервал для уточнённого матожидания: [1.1729326044166368, 3.712781681297649]

**Доверительная вероятность для заданной погрешности**:



Доверительная вероятность для погрешности 0.38: 0.3995

**Результаты выполнения работы**

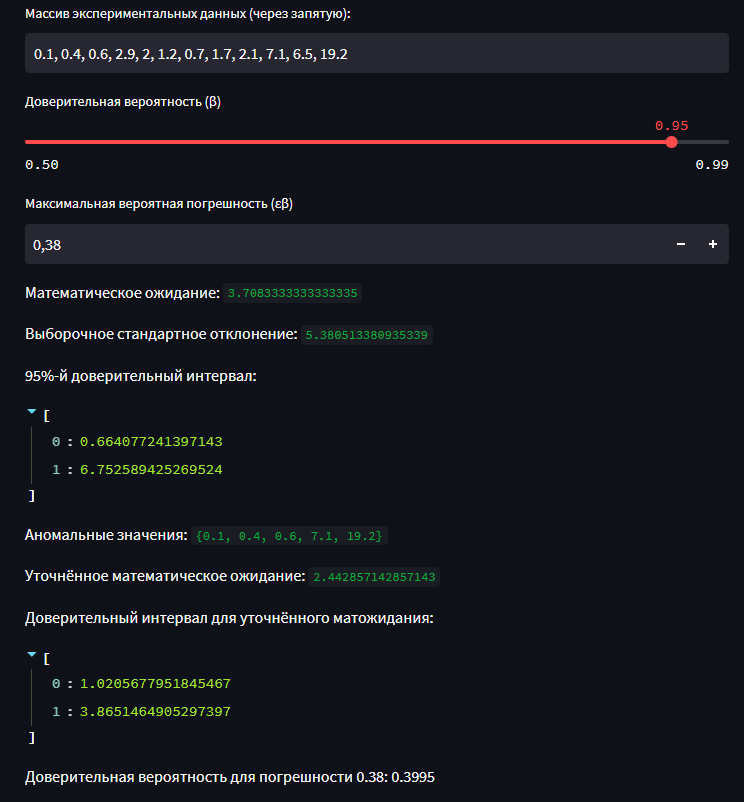
В ходе работы был реализован алгоритм на основе массива экспериментальных данных (WEB – page формат).

С работоспособностью алгоритма можно ознакомиться по ссылке ниже:

<https://firstlar.streamlit.app/>

Полученные результаты в ходе вычислений:

* Математическое ожидание: 3.7083333333333335
* Выборочное стандартное отклонение: 5.380513380935339
* 95%-й доверительный интервал: [0.6640213014491598, 6.752645365217507]
* Аномальные значения: {0.1, 0.4, 0.6, 7.1, 19.2}
* Уточнённое математическое ожидание: 2.442857142857143
* Доверительный интервал для уточнённого матожидания: [1.1729326044166368, 3.712781681297649]
* Доверительная вероятность для погрешности 0.38: 0.3995



**Анализ полученных результатов:**

В рамках данной задачи была проведена статистическая обработка массива экспериментальных данных с целью оценки математического ожидания случайной величины, построения доверительных интервалов и проверки качества оценивания. Рассмотрим полученные результаты более детально.

1. **Оценка математического ожидания:** математическое ожидание для исходных данных было оценено как 3.71. Это значение представляет собой среднее арифметическое всех наблюдений в выборке, что является классической процедурой для получения оценки центральной тенденции. Оно отражает средний уровень значений случайной величины в экспериментальных данных, однако следует помнить, что для более точной оценки важно учитывать степень отклонения значений от среднего.
2. **Выборочное стандартное отклонение:** стандартное отклонение составило 5.38, что является достаточно высоким показателем. Это свидетельствует о большой вариативности данных относительно их среднего значения. Такая большая дисперсия может указывать на наличие выбросов или сильных колебаний в данных, что важно учитывать при дальнейших анализах и при построении доверительных интервалов.
3. **95%-й доверительный интервал:** для оценки математического ожидания был построен 95%-й доверительный интервал, который составил [0.664, 6.753]. Это означает, что с вероятностью 95% истинное значение математического ожидания случайной величины лежит в этом диапазоне. Однако следует отметить, что данный интервал довольно широкий, что подтверждает высокое стандартное отклонение и большую вариативность данных.
4. **Отсеивание аномальных значений:** в ходе анализа было выявлено 5 значений, выходящих за пределы 95%-го доверительного интервала: {0.1, 0.4, 0.6, 7.1, 19.2}. Эти данные были исключены из дальнейшего анализа, так как они значительно отклоняются от основной массы наблюдений. Отсев аномальных значений позволяет улучшить точность оценки математического ожидания, так как выбросы могут существенно исказить результаты.
5. **Уточнённая оценка математического ожидания:** после исключения аномальных значений уточнённая оценка математического ожидания составила 2.44. Это значение отражает более точную центральную тенденцию данных, что стало возможным благодаря удалению искажённых данных. Уточнённое математическое ожидание значительно отличается от первоначальной оценки, что подтверждает важность отсеивания аномалий.
6. **Доверительный интервал для уточнённого математического ожидания:** для уточнённого математического ожидания был построен доверительный интервал, который составил [1.173, 3.713]. Этот интервал уже по ширине и более точно отражает диапазон истинного значения математического ожидания. Уменьшение ширины интервала по сравнению с первоначальным интервалом подтверждает улучшение точности оценки после удаления аномальных значений.
7. **Проверка качества оценивания по заданной максимальной вероятной погрешности:** для максимальной вероятной погрешности 0.38 была вычислена доверительная вероятность, которая составила 0.3995. Это означает, что вероятность того, что погрешность не превысит заданное значение, составляет примерно 40%. Этот результат указывает на определённую неопределённость в оценке, что может быть связано с высоким стандартным отклонением исходных данных и необходимостью улучшения точности дальнейших измерений.

**Вывод:**

В ходе выполнения данной лабораторной работы были проведены статистические расчёты по оцениванию числовых характеристик случайной величины.

Процесс оценки математического ожидания с использованием доверительных интервалов и фильтрации аномальных значений показал, что исходные данные имеют высокую вариативность и содержат выбросы, которые существенно влияют на результаты. Удаление аномальных значений позволило улучшить точность оценки математического ожидания и уменьшить ширину доверительного интервала. Однако при дальнейшем анализе следует учитывать влияние малых выборок и возможные отклонения от нормальности распределения данных, что может потребовать применения других методов статистической обработки, например, использования t-распределения или тестов на нормальность.

**Листинг кода:**

|  |
| --- |
| import math  from scipy.stats import norm  import streamlit as st  # Заголовок приложения  st.title("Анализ данных с использованием доверительных интервалов")  # Ввод данных пользователем  st.header("Введите данные")  data\_input = st.text\_input(      "Массив экспериментальных данных (через запятую):",      "0.1, 0.4, 0.6, 2.9, 2, 1.2, 0.7, 1.7, 2.1, 7.1, 6.5, 19.2"  )  confidence\_level = st.slider(      "Доверительная вероятность (β)",      min\_value=0.5, max\_value=0.99, value=0.95, step=0.01  )  epsilon\_beta = st.number\_input(      "Максимальная вероятная погрешность (εβ)",      min\_value=0.01, value=0.38, step=0.01  )  # Преобразуем введенные данные в массив чисел  try:      arr = [float(x.strip()) for x in data\_input.split(",")]  except ValueError:      st.error("Ошибка: Пожалуйста, введите корректные числовые данные через запятую.")      st.stop()  # 1. Оценка математического ожидания  avg = sum(arr) / len(arr)  st.write("Математическое ожидание:", avg)  # 2. Выборочное стандартное отклонение  std = math.sqrt(sum((x - avg)\*\*2 for x in arr) / (len(arr) - 1))  st.write("Выборочное стандартное отклонение:", std)  # 3. Построение доверительного интервала для заданной доверительной вероятности  n = len(arr)  z\_score\_confidence = norm.ppf(1 - (1 - confidence\_level) / 2)  # Z-значение для β  margin\_of\_error = z\_score\_confidence \* (std / math.sqrt(n))  confidence\_interval = [avg - margin\_of\_error, avg + margin\_of\_error]  st.write(f"{confidence\_level \* 100:.0f}%-й доверительный интервал:", confidence\_interval)  # 4. Отсеивание аномальных значений  filtered\_values = [x for x in arr if confidence\_interval[0] <= x <= confidence\_interval[1]]  if len(filtered\_values) < len(arr):      st.write("Аномальные значения:", set(arr) - set(filtered\_values))  else:      st.write("Аномальных значений не найдено")  # 5. Уточнённая оценка математического ожидания  filtered\_avg = sum(filtered\_values) / len(filtered\_values)  st.write("Уточнённое математическое ожидание:", filtered\_avg)  # 6a. Доверительный интервал для уточнённого математического ожидания  filtered\_std = math.sqrt(sum((x - filtered\_avg)\*\*2 for x in filtered\_values) / (len(filtered\_values) - 1))  n\_filtered = len(filtered\_values)  z\_score\_filtered = norm.ppf(1 - (1 - confidence\_level) / 2)  quality\_error = z\_score\_filtered \* (filtered\_std / math.sqrt(n\_filtered))  interval\_filtered = [filtered\_avg - quality\_error, filtered\_avg + quality\_error]  st.write("Доверительный интервал для уточнённого матожидания:", interval\_filtered)  # 6b. Доверительная вероятность для заданной максимальной вероятной погрешности  z\_score\_epsilon = (epsilon\_beta \* math.sqrt(n\_filtered)) / filtered\_std  p\_confidence = norm.cdf(z\_score\_epsilon) - norm.cdf(-z\_score\_epsilon)  st.write(f"Доверительная вероятность для погрешности {epsilon\_beta}: {p\_confidence:.4f}") |