ГУАП

КАФЕДРА № 43

ОТЧЕТ   
ЗАЩИЩЕН С ОЦЕНКОЙ

ПРЕПОДАВАТЕЛЬ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Профессор |  |  |  | Колесникова С.А. |
| должность, уч. степень, звание |  | подпись, дата |  | инициалы, фамилия |

|  |
| --- |
| ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ №2  Генератор СВ. Имитация СМО. Сумма потоков. |
| **по дисциплине: Компьютерное моделирование** |
|  |
|  |

РАБОТУ ВЫПОЛНИЛ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| СТУДЕНТ ГР. | 4134К |  |  |  | Самарин Д. В. |
|  |  |  | подпись, дата |  | инициалы, фамилия |

Санкт-Петербург

2024

**Цель:**

Цель настоящей работы – освоить средства моделирования случайных величин (СВ) с произвольным распределением на основе равномерного распределения. Построить имитационную модель двух потоков, в котором длительность промежутков времени между поступлениями заявок имеет показательный закон с параметрами λ1,λ2. Осуществить проверку статистической гипотезы о соблюдении свойства аддитивности пуассоновского потока (сумма пуассоновских потоков есть поток пуассоновский).

**Ход работы:**

1. Ознакомиться со справочными сведениями; сформулировать особенности пуассоновского потока событий; указать связь (дискретного) пуассоновского потока и (непрерывного) показательного распределения.

2. Запрограммировать предложенный алгоритм генерации пуассоновского потока с использованием MatLab или Python.

3. Создать графическую интерпретацию потока событий.

4. Осуществить проверку гипотезы о виде распределения для суммарного потока.

5. Сравнить интенсивности выборочных и теоретических интенсивностей потоков.

6. Составить и представить преподавателю отчет о работе.

Исходные данные: промежуток наблюдения [𝑇1, 𝑇2], параметр λ. Значения параметра λ должны быть выбраны в зависимости от номера студента в списке группы N, где:

𝑇1 = 𝑁,

𝑇2 = 𝑁 + 100,

λ1 = (𝑁+8)/(𝑁+24),

λ2 = (𝑁+9)/(𝑁+25).

При этом N = 14.

Особенности Пуассоновского потока:

1. Однородность событий – количество событий за фиксированный промежуток времени распределено по закону Пуассона. При этом сам поток – последовательность событий, происходящих независимо друг от друга в случайные моменты времени.
2. Аддитивность – когда два независимых Пуассоновских потока с интенсивностями лямбда 1, 2 суммируются и при этом суммарный поток также является пуассоновским.
3. Независимость событий – события в потоке независимы, т.е знание о времени наступлении одного события не даёт нам информации о наступлении следующего.

Пуассоновский поток и показательное распределение:

Пуассоновский поток описывает дискретное количество событий, происходящих за фиксированные интервалы времени.

Показательное распределение описывает интервалы времени между событиями в пуассоновском процессе. Это распределение обладает важным свойством отсутствия "памяти", то есть вероятность наступления следующего события не зависит от того, сколько времени прошло с последнего события.

Пуассоновское распределение описывает количество событий, произошедших за фиксированный интервал времени. Если интервалы между событиями имеют показательное распределение, то число событий на фиксированном интервале будет распределено по закону Пуассона

**Заданные параметры:**

Время наблюдения T = 100 (от 14 до 114 секунд).

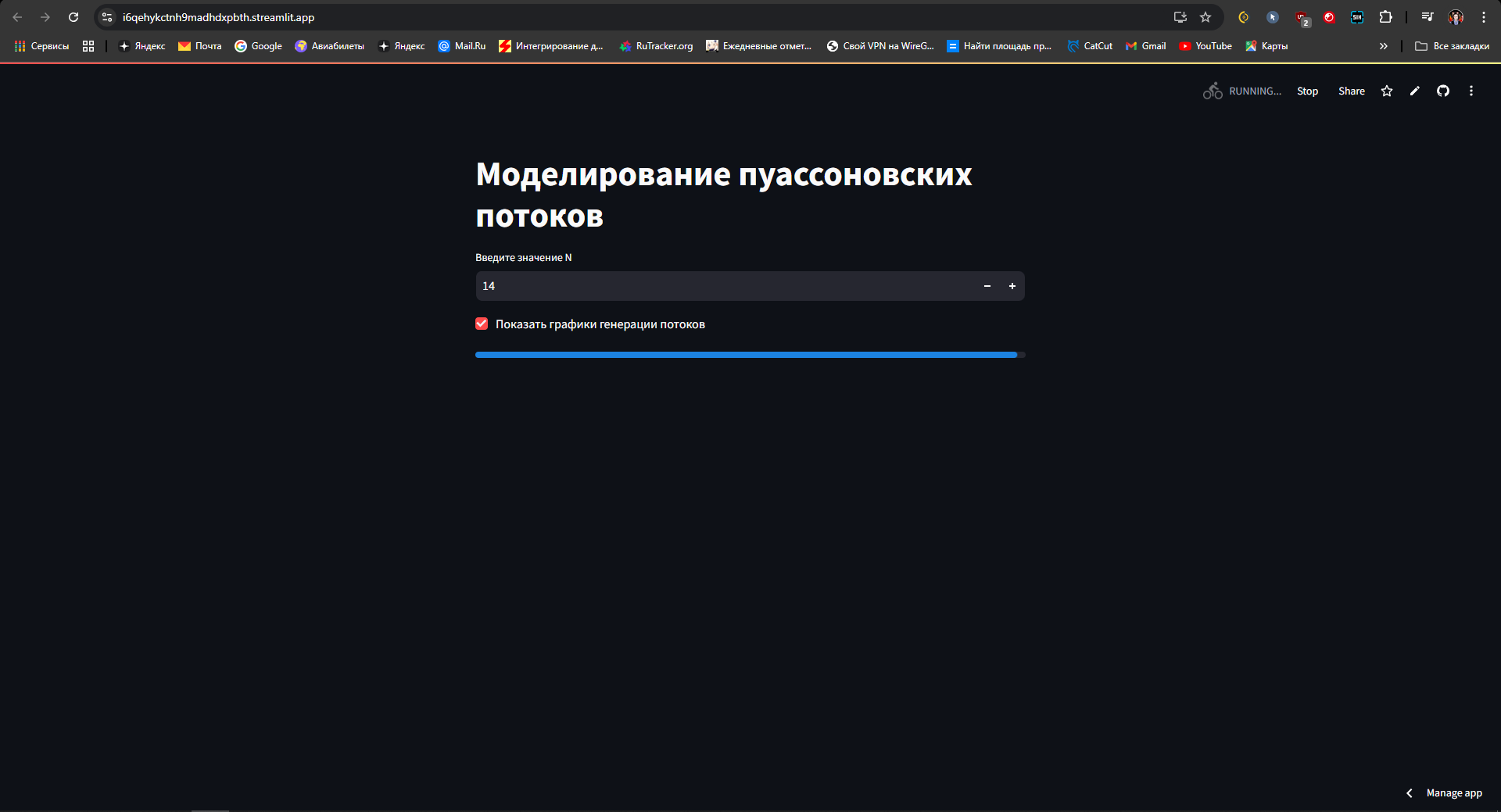
Пуассоновские потоки с интенсивностями λ1 и λ2​:

* λ1 = (N+8) / (N+24) = 0.5789
* λ2 = (N+9) / (N+25) = 0.5897

Где N=14 — значение параметра, связанное с номером студента.

**Результаты моделирования:**

В ходе работы был реализован код в виде WEB-приложения, на которое можно зайти по ссылке: <https://i6qehykctnh9madhdxpbth.streamlit.app/>



На графиках отображены результаты для 1000 реализаций потоков λ1​, λ2 ​, а также их суммы.

**Алгоритм генерации:**

1. **Генерация случайного числа:**  
   Сгенерировать случайное число из равномерного распределения на интервале (0,1).
2. **Преобразование в экспоненциальный интервал:**  
   Преобразовать сгенерированное число U в интервал времени между событиями t с помощью формулы:

где λ — интенсивность потока.

1. **Накопление времени:**  
   Суммировать интервалы ti​ для получения времени каждого следующего события. Процесс продолжается, пока сумма не превысит заданный временной отрезок [T1, T2].

**Рассмотрим статистические показатели:**

Статистика для процесса с λ1=0.58

* Эмпирическая интенсивность λ: 0.5774
* Теоретическая дисперсия: 57.8947
* Эмпирическая дисперсия: 40.9124
* Статистика Колмогорова-Смирнова: 0.0141
* p-значение Колмогорова-Смирнова: 0.6185
* Статистика χ²: 28.1719
* p-значение χ²: 1.0000
* Теория подтверждена: Да
* Теория подтверждена по χ²: Да

Статистика для процесса с λ2=0.59

* Эмпирическая интенсивность λ: 0.5832
* Теоретическая дисперсия: 58.9744
* Эмпирическая дисперсия: 57.2576
* Статистика Колмогорова-Смирнова: 0.0115
* p-значение Колмогорова-Смирнова: 0.8376
* Статистика χ²: 52.8252
* p-значение χ²: 0.9918
* Теория подтверждена: Да
* Теория подтверждена по χ²: Да

Статистика для суммы процессов (λ1 + λ2)

* Эмпирическая интенсивность λ: 1.1606
* Теоретическая дисперсия: 116.8691
* Эмпирическая дисперсия: 97.5764
* Статистика Колмогорова-Смирнова: 0.0062
* p-значение Колмогорова-Смирнова: 0.9797
* Статистика χ²: 128.9863
* p-значение χ²: 0.9024
* Теория подтверждена: Да
* Теория подтверждена по χ²: Да

**На основе полученных данных сделаем небольшой анализ:**

**Сопоставление интенсивностей**.

* Эмпирические интенсивности для обоих потоков близки к теоретическим значениям. Это подтверждает корректность генерации пуассоновских процессов.

**Дисперсия:**

Несмотря на небольшие расхождения, эмпирические значения достаточно близки к теоретическим, чтобы гипотеза о аддитивности дисперсии оставалась правдивой.

**Статистика:**

Малые значения статистики и высокие p-значения подтверждают гипотезу о том, что эмпирическое распределение соответствует теоретическому. Следовательно, теоретическая модель адекватно описывает процессы и их сумму.

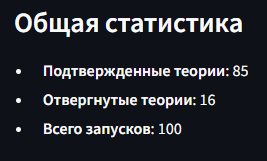
**Критерий хи-квадрат**

Статистические значения χ² и высокие p-значения подтверждают гипотезу, распределение потоков согласуется с теорией.

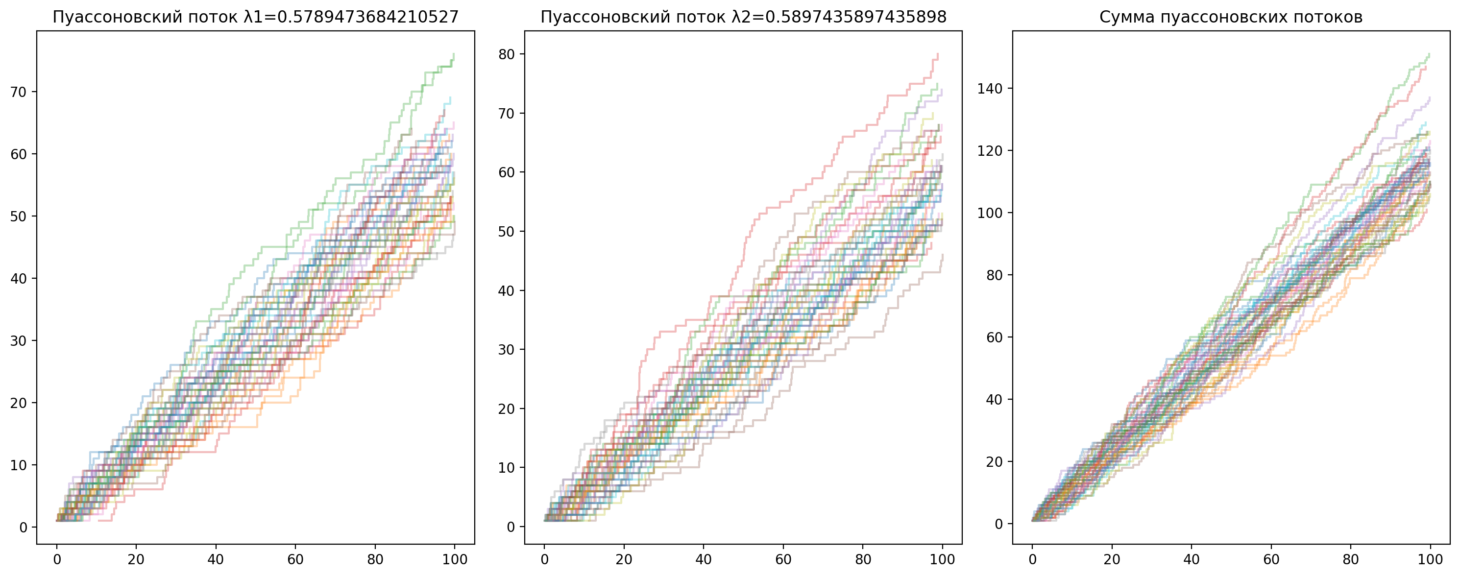
Все статистические проверки (интенсивность, дисперсия, K-S и χ²) подтверждают свойство аддитивности пуассоновского потока:

λсумма​=λ1​+λ2​

В конце объявлена общая статистика, которая отображает количество подтвержденных и отвергнутых теорий:



**Рассмотрим полученные графики:**



**Листинг кода:**

|  |
| --- |
| import numpy as np  import matplotlib.pyplot as plt  from scipy.stats import expon, poisson, kstest, chisquare  import streamlit as st  # Функция для генерации пуассоновских потоков  def generate\_poisson\_process(lmbda, T, size=500):      inter\_arrival\_times = -np.log(np.random.rand(size)) / lmbda      event\_times = np.cumsum(inter\_arrival\_times)      return event\_times[event\_times <= T]  # Функция для построения графиков  def plot\_processes(lmbda1, lmbda2, T, num\_realizations=50):      fig, axes = plt.subplots(1, 3, figsize=(15, 6))      # Генерация пуассоновских потоков      events1 = [generate\_poisson\_process(lmbda1, T) for \_ in range(num\_realizations)]      events2 = [generate\_poisson\_process(lmbda2, T) for \_ in range(num\_realizations)]        # Суммируем два пуассоновских потока      sum\_events = [np.sort(np.concatenate([e1, e2])) for e1, e2 in zip(events1, events2)]      # Построение графиков потоков      for e in events1:          axes[0].step(e, np.arange(1, len(e) + 1), where='post', alpha=0.3)      axes[0].set\_title(f'Пуассоновский поток λ1={lmbda1}')        for e in events2:          axes[1].step(e, np.arange(1, len(e) + 1), where='post', alpha=0.3)      axes[1].set\_title(f'Пуассоновский поток λ2={lmbda2}')        for e in sum\_events:          axes[2].step(e, np.arange(1, len(e) + 1), where='post', alpha=0.3)      axes[2].set\_title('Сумма пуассоновских потоков')      plt.tight\_layout()      return fig, events1, events2, sum\_events  # Функция для расчета статистических характеристик  def compute\_statistics(events, lmbda, T):      counts = np.array([len(e) for e in events])        # Вычисление теоретических и эмпирических параметров      empirical\_lambda = np.mean(counts) / T      empirical\_var = np.var(counts)      theoretical\_var = lmbda \* T      # KS-тест      inter\_arrival\_times = np.concatenate([np.diff(e) for e in events if len(e) > 1])      ks\_stat, ks\_pvalue = kstest(inter\_arrival\_times, 'expon', args=(0, 1/lmbda))        # χ²-тест      max\_count = max(counts)      expected\_counts = poisson(lmbda \* T).pmf(np.arange(0, max\_count + 1)) \* len(events)        # Приведение массивов к одинаковой длине      observed\_counts = np.bincount(counts, minlength=len(expected\_counts))      expected\_counts = expected\_counts[:len(observed\_counts)]        # Нормализация      expected\_counts \*= observed\_counts.sum() / expected\_counts.sum()      chi2\_stat, chi2\_pvalue = chisquare(observed\_counts, f\_exp=expected\_counts)      return {          'empirical\_lambda': empirical\_lambda,          'theoretical\_var': theoretical\_var,          'empirical\_var': empirical\_var,          'ks\_stat': ks\_stat,          'ks\_pvalue': ks\_pvalue,          'chi2\_stat': chi2\_stat,          'chi2\_pvalue': chi2\_pvalue      }  def main():        # Пользовательский ввод      N = st.number\_input("Введите значение N", value=14)      # Промежуток наблюдения      T1 = N      T2 = N + 100      T = T2 - T1  # Время наблюдения      # Параметры потоков      lmbda1 = (N + 8) / (N + 24)  # Интенсивность первого потока      lmbda2 = (N + 9) / (N + 25)  # Интенсивность второго потока      num\_realizations = 50  # Количество реализаций      # Инициализация счётчиков в состоянии сессии      if 'confirmed\_count' not in st.session\_state:          st.session\_state.confirmed\_count = 0      if 'rejected\_count' not in st.session\_state:          st.session\_state.rejected\_count = 0      total\_runs = 100      # Чекбокс для отображения графиков      show\_plots = st.checkbox("Показать графики генерации потоков", value=True)      # Инициализация статус-бара      progress\_bar = st.progress(0)      # Инициализация состояния сессии для хранения графиков      if 'plots' not in st.session\_state:          st.session\_state.plots = []      if 'stats' not in st.session\_state:          st.session\_state.stats = []      # Генерация графиков и статистики только при первом запуске      if len(st.session\_state.plots) == 0:          for run in range(total\_runs):              # Визуализация потоков              if show\_plots:                  fig, events1, events2, sum\_events = plot\_processes(lmbda1, lmbda2, T, num\_realizations)                  st.session\_state.plots.append(fig)              # Расчет статистических характеристик для каждого потока              stats1 = compute\_statistics(events1, lmbda1, T)              stats2 = compute\_statistics(events2, lmbda2, T)              stats\_sum = compute\_statistics(sum\_events, lmbda1 + lmbda2, T)              st.session\_state.stats.append((stats1, stats2, stats\_sum))              # Определение подтверждённых и отвергнутых теорий              if stats1['ks\_pvalue'] > 0.05 and stats2['ks\_pvalue'] > 0.05:                  st.session\_state.confirmed\_count += 1              else:                  st.session\_state.rejected\_count += 1              # Обновление статус-бара              progress\_bar.progress((run + 1) / total\_runs)      # Выбор цикла для отображения графика      selected\_run = st.selectbox("Выберите цикл для отображения графика", range(total\_runs))        if show\_plots and selected\_run < len(st.session\_state.plots):          st.pyplot(st.session\_state.plots[selected\_run])  # Отображение выбранного графика      # Получаем статистику для выбранного цикла      if selected\_run < len(st.session\_state.stats):          stats1, stats2, stats\_sum = st.session\_state.stats[selected\_run]      # Добавляем определение подтверждённых и отвергнутых теорий для выбранного        is\_theory\_confirmed1 = stats1['ks\_pvalue'] > 0.05      is\_theory\_confirmed2 = stats2['ks\_pvalue'] > 0.05      is\_theory\_confirmed\_sum = stats\_sum['ks\_pvalue'] > 0.05      # Определение подтверждённых и отвергнутых теорий      if stats1['ks\_pvalue'] > 0.05 and stats2['ks\_pvalue'] > 0.05 and stats1['chi2\_pvalue'] > 0.05 and stats2['chi2\_pvalue'] > 0.05:          st.session\_state.confirmed\_count += 1      else:          st.session\_state.rejected\_count += 1      # Вывод результатов      stats1, stats2, stats\_sum = st.session\_state.stats[selected\_run]      st.subheader("Статистика для процесса с λ1={:.2f}".format(lmbda1))      st.markdown(f"""      - \*\*Эмпирическая интенсивность λ\*\*: {stats1['empirical\_lambda']:.4f}      - \*\*Теоретическая дисперсия\*\*: {stats1['theoretical\_var']:.4f}      - \*\*Эмпирическая дисперсия\*\*: {stats1['empirical\_var']:.4f}      - \*\*Статистика Колмогорова-Смирнова\*\*: {stats1['ks\_stat']:.4f}      - \*\*p-значение Колмогорова-Смирнова\*\*: {stats1['ks\_pvalue']:.4f}      - \*\*Статистика χ²\*\*: {stats1['chi2\_stat']:.4f}      - \*\*p-значение χ²\*\*: {stats1['chi2\_pvalue']:.4f}      - \*\*Теория подтверждена\*\*: {"Да" if is\_theory\_confirmed1 else "Нет"}      - \*\*Теория подтверждена по χ²\*\*: {"Да" if stats1['chi2\_pvalue'] > 0.05 else "Нет"}      """)      st.subheader("Статистика для процесса с λ2={:.2f}".format(lmbda2))      st.markdown(f"""      - \*\*Эмпирическая интенсивность λ\*\*: {stats2['empirical\_lambda']:.4f}      - \*\*Теоретическая дисперсия\*\*: {stats2['theoretical\_var']:.4f}      - \*\*Эмпирическая дисперсия\*\*: {stats2['empirical\_var']:.4f}      - \*\*Статистика Колмогорова-Смирнова\*\*: {stats2['ks\_stat']:.4f}      - \*\*p-значение Колмогорова-Смирнова\*\*: {stats2['ks\_pvalue']:.4f}      - \*\*Статистика χ²\*\*: {stats2['chi2\_stat']:.4f}      - \*\*p-значение χ²\*\*: {stats2['chi2\_pvalue']:.4f}      - \*\*Теория подтверждена\*\*: {"Да" if is\_theory\_confirmed2 else "Нет"}      - \*\*Теория подтверждена по χ²\*\*: {"Да" if stats1['chi2\_pvalue'] > 0.05 else "Нет"}      """)      st.subheader("Статистика для суммы процессов (λ1 + λ2)")      st.markdown(f"""      - \*\*Эмпирическая интенсивность λ\*\*: {stats\_sum['empirical\_lambda']:.4f}      - \*\*Теоретическая дисперсия\*\*: {stats\_sum['theoretical\_var']:.4f}      - \*\*Эмпирическая дисперсия\*\*: {stats\_sum['empirical\_var']:.4f}      - \*\*Статистика Колмогорова-Смирнова\*\*: {stats\_sum['ks\_stat']:.4f}      - \*\*p-значение Колмогорова-Смирнова\*\*: {stats\_sum['ks\_pvalue']:.4f}      - \*\*Статистика χ²\*\*: {stats\_sum['chi2\_stat']:.4f}      - \*\*p-значение χ²\*\*: {stats\_sum['chi2\_pvalue']:.4f}      - \*\*Теория подтверждена\*\*: {"Да" if is\_theory\_confirmed\_sum else "Нет"}      - \*\*Теория подтверждена по χ²\*\*: {"Да" if stats1['chi2\_pvalue'] > 0.05 else "Нет"}      """)      # Вывод общей статистики      st.subheader("Общая статистика")      st.markdown(f"""      - \*\*Подтвержденные теории\*\*: {st.session\_state.confirmed\_count}      - \*\*Отвергнутые теории\*\*: {st.session\_state.rejected\_count}      - \*\*Всего запусков\*\*: {total\_runs}      """)  if \_\_name\_\_ == "\_\_main\_\_":      st.title("Моделирование пуассоновских потоков")      main()  # Как генерируются пуассоновские потоки?  # Пуассоновский поток — это последовательность случайных событий, происходящих в некоторый момент времени. Время между событиями (интервалы) имеет экспоненциальное распределение с параметром λ — интенсивностью потока. Для генерации пуассоновских потоков:  # 1. Генерируются интервалы между событиями: время ожидания до следующего события экспоненциально распределено (Expon(λ)). Для генерации этих интервалов используется формула:  #     interval = -1 / λ \* ln(U),  #     где U — равномерно распределённая случайная величина на отрезке [0, 1].  # 2. Накопление событий во времени: суммируются интервалы между событиями для получения момента каждого события. Генерированные интервалы добавляются к текущему времени, пока не достигнут конца наблюдаемого промежутка.  # Определение и свойства пуассоновских потоков  # Пуассоновский поток описывает случайный процесс, при котором события происходят независимо друг от друга с постоянной интенсивностью λ.  # Основные свойства:  # 1. Отсутствие памяти: вероятность того, что событие произойдет в ближайший момент времени, не зависит от того, когда произошло последнее событие.  # 2. Независимость событий: количество событий в непересекающихся промежутках времени — независимые случайные величины.  # 3. Гомогенность: вероятность появления события в малом интервале времени Δt примерно равна λΔt, где λ — интенсивность потока.  # Распределение величины потока  # Количество событий N(T), произошедших за фиксированное время T, имеет пуассоновское распределение с параметром λT:  #     P(N(T) = k) = (λT)^k \* e^(-λT) / k! ,  k = 0, 1, 2, ...  # где:  # - λ — интенсивность потока (среднее число событий в единицу времени),  # - T — период времени наблюдения.  # Это означает, что вероятность появления точно k событий за время T описывается этой формулой.  # Математическое ожидание и дисперсия пуассоновского потока  # Математическое ожидание (среднее количество событий) для пуассоновского потока равно:  #     E[N(T)] = λT.  # Дисперсия числа событий за время T также равна λT:  #     Var(N(T)) = λT.  # Это важное свойство пуассоновского процесса: среднее значение и дисперсия числа событий одинаковы.  # Оценка интенсивности  # Чтобы оценить интенсивность λ по наблюдениям, можно использовать эмпирическую оценку интенсивности:  #     λ^ = Σ N(T\_i) / (nT),  # где N(T\_i) — число событий за время T\_i, n — количество наблюдений.  # Вероятность в пуассоновском потоке  # Вероятность того, что произойдет k событий за время T, определяется пуассоновским распределением. Вероятность появления хотя бы одного события на малом промежутке времени Δt приблизительно равна λΔt. Для небольших Δt вероятность более чем одного события очень мала.  # Смысл χ²-статистики  # Статистика χ² используется для проверки гипотезы о том, что наблюдаемые частоты событий соответствуют ожидаемым (теоретическим) частотам для пуассоновского распределения.  # Шаги проведения теста:  # 1. Наблюдаемые частоты: подсчитываются фактические частоты количества событий в разных интервалах времени.  # 2. Ожидаемые частоты: вычисляются по формуле пуассоновского распределения с параметром λ.  # 3. χ²-статистика:  #     χ² = Σ ((O\_i - E\_i)² / E\_i),  # где O\_i — наблюдаемые частоты, E\_i — ожидаемые частоты.  # Если χ²-статистика мала, то гипотеза о том, что наблюдения соответствуют пуассоновскому распределению, подтверждается.  # Основная задача лабораторной работы  # Основная задача ЛР — исследовать пуассоновские потоки и их характеристики, проверить гипотезы о распределении времени между событиями (экспоненциальное) и числе событий (пуассоновское распределение), используя статистические методы (тесты Колмогорова-Смирнова и χ²).  # Практические частоты  # Практические частоты — это фактически наблюдаемые числа событий за время T. Они сравниваются с ожидаемыми частотами, вычисленными с использованием пуассоновского распределения для проверки соответствия модели и реальных данных.    # ИНФА ПО КОДУ:  # Генерация Пуассоновского Процесса  # Для генерации пуассоновского процесса используется формула интервалов между событиями, которая определяется как:  # Интервал между событиями: T = -ln(U) / λ, где:  # - U — случайная величина, равномерно распределенная на [0, 1]  # - λ — интенсивность потока (среднее количество событий на единицу времени)  # Визуализация Пуассоновских Потоков  # Пуассоновский поток отображается на графике с помощью лестничной диаграммы (step plot), где ось X — это время событий, а ось Y — это число накопленных событий.  # Для отображения двух независимых потоков с различными интенсивностями и их суммы используется функция step(), которая строит графики по каждому потоку.  # Сумма Пуассоновских Потоков  # Суммирование пуассоновских потоков с интенсивностями λ₁ и λ₂ даёт новый поток с интенсивностью:  # λ\_sum = λ₁ + λ₂  # Статистика Пуассоновских Потоков  # Эмпирическая интенсивность пуассоновского потока определяется как:  # λ\_empirical = N\_events / T, где:  # - N\_events — количество событий за период наблюдения  # - T — время наблюдения  # Теоретическая дисперсия пуассоновского потока: Var\_theoretical = λ \* T  # Эмпирическая дисперсия вычисляется на основе числа событий в разных реализациях процесса.  # Тест Колмогорова-Смирнова  # Тест Колмогорова-Смирнова используется для проверки гипотезы о том, что интервалы между событиями распределены экспоненциально с параметром λ.  # Проверяемая гипотеза: интервалы между событиями следуют распределению Exp(λ), где Exp(λ) — экспоненциальное распределение с параметром λ.  # Тест χ²  # Тест χ² используется для сравнения распределения числа событий с теоретическим распределением Пуассона.  # Статистика χ² определяется как сумма отклонений фактических и ожидаемых значений по формуле:  # χ² = Σ (O\_i - E\_i)² / E\_i, где:  # - O\_i — наблюдаемое количество событий  # - E\_i — ожидаемое количество событий согласно распределению Пуассона  # Общая Статистика  # На основе статистических тестов определяется, подтверждаются ли теории о том, что процесс является пуассоновским, и проводятся оценки соответствия данных теоретическим предположениям. |

**Вывод:**

Выполнив лабораторную работу, я освоил моделирование СВ с произвольным распределением на основе равномерного распределения. А также построил имитационную модель двух потоков, осуществив проверку статической гипотезы о соблюдении свойства аддитивности пуассоновского потока.