Отчет о выполнении лабораторной работы 1.1.4

Изучение статистических закономерностей на примерее фона космического излучения

Содержание

1. Введение

Цель:

Познакомиться с основными понятиями стастистики, изучить статистические закономерности однородного во времени случайного процесса, проверить возможность описания исследуемого процесса статистическими законами Пуассона и Гаусса,
измерить среднеее число регистрируемых космических лучей в секунду и определить погрешность результата

Оборудование:

Счётчик Гейгера-Мюллера СТС-6, компьютер с интерфейсом для связи с счетчиком

2. Теоретическая часть

Природа космических лучей

Первичные космические лучи - это поток стабильных частиц, имеющих большую $\text{Е}_{\text{к}}$, которая составляет от 10^9 до 10^{21} эВ (1 эВ $= 1,6 * 10^{-19}$ Дж)

В состав первичных космических лучей входят: протоны (92%), ядра гелия (6,6%) и более тяжелые ядра вплоть до никеля (0,8%). Электронов и позитронов набирается примерно 1%, причем позитронов в 10 раз меньше, чем электронов. Число γ-квантов с энергиями больше 10^8 эВ составляет всего 0,01%.

Помимо первичных, существуют и вторичные космические лучи, которые образуются из первичных взаимодействуя с ядрами атомов атмосферных газов и образуют вторичные космические лучи. Именно вторичные протоны, мюоны и нейтроны образуют *жесткую* компоненту вторичных космических лучей, которая проходит через свинцовую пластинку толщиной в 10 см.

Обратным образом определяется и *мягкая* компонента вторичных космических лучей, состоящая из электронов, позитронов и фотонов. *Мягкая* компонента существует в атмосфере вблизи поверхности Земли лишь благодаря тому, что она генерируется *жесткой*.

Космические лучи и естественная радиоактивность Земли и воздуха являются основными источниками ионов в нижней части атмосферы земли (до высот порядка 60 км). Поэтому ионизация в атмосфере сначала падает, а выше 1 км начинает возрастать.

Устройство счетчика Гейгера-Мюллера

Счетчик СТС-6 представляет собой тонкостенный металлический цилиндр (катод), который заполнен газом, с натянутой вдоль его оси нити (анод). На электроды подается напряжение в 400В. Проходящие сквозь счетчик частицы космических лучей ионизируют газ в счетчике, а также выбивает электроны из его стенок. Все эти электроны ускоряются в сильном электрическом поле между электродами, соударяются с другими молекулами газа, выбивая из них вторичные электроны, которые так же ускоряются в электрическом поле и могут ионизировать другие молекулы газа.

Таким образом создается лавина электронов и через счетчик резко увеличивается ток.

Компьютер подключен к счетчику согласно следующей схеме:

Схема включения счетчика

Заряд на счетчике и C_1 обеспечивает развитие лавины электронов на короткое время, а через C_2 в компьютер передается короткий импульс.

Статистические понятия и описание эксперимента

Процесс называется *Пуассоновским*, если случайные события однородны по времени и каждое последующее событие не зависит от предыдущего. Соответственно говорят, что результаты такого процесса подчиняются *распределению Пуассона*.

Из теории вероятностей выводится формула для такого распределения:

$$P_n = \frac{\nu t}{n!} e^{-\nu t}$$

Для распределения Пуассона справедливо: $\sigma = \sqrt{n_o}$

где σ - квадратичная ошибка числа отсчетов, а n_o - среднее число отсчетов за тот же интервал.

Но поскольку мы не знаем истинное среднее значение измеряемой величины, поэтому вместо n_o

мы используем измеренное значение $\bar{n}$, то есть: $\sigma = \sqrt{\bar{n}}$

Тогда результат измерений записывается в виде: $n_o = \bar{n} \pm \sqrt{\bar{n}}$

Пусть было проведено N измерений, в результате которых были получены числа частиц $n_1, n_2, \ldots, n_N$. Тогда:

$$\bar{n} = \frac{1}{N} \sum_{i=1}^{N} n_i$$

$$\sigma_{\text{отд}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (n_i - \bar{n})^2}$$

Полученное из первой формулы значение $\bar{n}$ путем усреднения результатов по серии из N опытов не вполне точно совпадает с истинным средним значение n_o и сама является случайной величиной. Теория вероятностей показывает, что стаднартная ошибка отклонения $\bar{n}$ от n_o может быть определена следующим образом:

$$\sigma_{\bar{n}} = \frac{1}{N} \sqrt{\sum_{i=1}^{N} (n_i - \bar{n})^2} = \frac{\sigma_{\text{отд}}}{\sqrt{N}}$$

Однако наибольший интерес представляет не абсолютная, а относительная точность измерений. Для той же серии из N измерений относительная ошибка отдельного измерения и относительная ошибка в определении среднего по всем измерениям значения $\bar{n}$ равны соответственно:

$$\epsilon_{\text{о т д}} = \frac{\sigma_{\text{о т д}}}{n_i} \approx \frac{1}{\sqrt{n_i}}$$

$$\epsilon_{\bar{n}} = \frac{\sigma_{\bar{n}}}{\bar{n}} = \frac{\sigma_{\text{о т д}}}{\bar{n}\sqrt{(N)}} \approx \frac{1}{\sqrt{n_i N}}$$

Для оценки было использовано соотношение $\sigma_{\text{о т д}} \approx \sqrt{\bar{n}}$, которое следует из: $\sigma_{\text{о т д}} \approx \sigma_i = \sqrt{n_i}$

Таким образом, относительная точность измерения $\bar{n}$ определяется только полным числом отсчетов $\bar{n}N$ и не зависит от интервалов разбиения серии. И как видно, относительная точность измерения постепенно улучшается с увеличением числа отсчетов.

3. Ход работы

Демонстрационный эксперимент

При проведении демонстрационного эксперимента, данные для обработки получались от генератора случайных чисел. Но на основании полученных данных были сделаны следующие выводы:

1. Измеряемая величина флуктуирует
2. Флуктуации среднего значения измеряемой величины уменьшаются, и среднее значение выходит на постоянную величину
3. Флуктуации величины погрешности отдельного измерения уменьшаются, и погрешность отдельного измерения выходит на постоянную величину
4. Флуктуации величины погрешности среднего значения уменьшаются, а сама величина убывает

Построение гистограмм

In [8]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

Строим таблицу по данным эксперимента (t = 20c)

Processing math: 100%

In [6]:
```python
df_20 = pd.read_csv("data/data_20.csv")
display(df_20)
```

	1	2	3	4	5	6	7	8	9	10
0	24	29	28	27	29	23	23	22	17	36
1	20	23	19	27	20	17	24	20	24	28
2	19	32	20	26	23	20	30	17	24	19
3	31	23	28	26	37	28	30	21	28	22
4	28	22	28	29	19	25	24	18	21	22
5	23	21	23	19	21	25	13	28	25	33
6	21	25	23	26	25	36	22	19	21	29
7	23	22	26	20	22	31	25	27	22	25
8	36	24	23	29	25	25	27	24	28	25
9	25	26	35	22	20	31	23	23	22	30
10	26	28	31	22	35	31	27	25	28	20
11	20	27	30	29	33	21	25	23	18	16
12	23	22	19	31	14	31	33	19	23	31
13	32	27	21	28	27	29	32	36	33	28
14	22	33	32	29	32	24	25	29	23	25
15	24	30	27	24	21	24	25	27	19	14
16	24	24	20	15	33	26	30	27	36	29
17	25	31	22	22	19	27	26	16	19	25
18	29	27	21	28	22	28	33	18	29	25
19	20	32	33	35	33	30	28	30	26	25

In [7]:
```python
data_20 = df_20.to_numpy()
```

Строим таблицу для t = 40c

In [9]:
```python
data_40 = list()
for i in range(20):
    for j in range(0, 10, 2):
        data_40.append(data_20[i][j] + data_20[i][j+1])

data_40 = np.array(data_40).reshape((20,5))
df_40 = pd.DataFrame(data_40, columns=['1','2','3','4','5'])
display(df_40)
```

Processing math: 100%

	1	2	3	4	5
0	53	55	52	45	53
1	43	46	37	44	52
2	51	46	43	47	43
3	54	54	65	51	50
4	50	57	44	42	43
5	44	42	46	41	58
6	46	49	61	41	50
7	45	46	53	52	47
8	60	52	50	51	53
9	51	57	51	46	52
10	54	53	66	52	48
11	47	59	54	48	34
12	45	50	45	52	54
13	59	49	56	68	61
14	55	61	56	54	48
15	54	51	45	52	33
16	48	35	59	57	65
17	56	44	46	42	44
18	56	49	50	51	54
19	52	68	63	58	51

Делаем таблицу, удобную для построения гистограмм при t = 40с

In [11]:
```
data_40_values = np.unique(data_40, return_counts=True)[0]
data_40_counts = np.unique(data_40, return_counts=True)[1]
data_40_rates = list()

for i in range(len(data_40_counts.tolist())):
    data_40_rates.append(data_40_counts[i] / sum(data_40_counts))

data_40_rates = np.array(data_40_rates)
data_h40 = np.dstack((data_40_values, data_40_counts, data_40_rates)).reshape((29,3
```

In [13]:
```
df_h40 = pd.DataFrame(data_h40, columns=['1','2','3'])
df_h40['1'] = df_h40['1'].astype(int)
df_h40['2'] = df_h40['2'].astype(int)
display(df_h40)
```

Processing math: 100%

	1	2	3
0	33	1	0.01
1	34	1	0.01
2	35	1	0.01
3	37	1	0.01
4	41	2	0.02
5	42	3	0.03
6	43	4	0.04
7	44	5	0.05
8	45	5	0.05
9	46	7	0.07
10	47	3	0.03
11	48	4	0.04
12	49	3	0.03
13	50	6	0.06
14	51	8	0.08
15	52	9	0.09
16	53	5	0.05
17	54	8	0.08
18	55	2	0.02
19	56	4	0.04
20	57	3	0.03
21	58	2	0.02
22	59	3	0.03
23	60	1	0.01
24	61	3	0.03
25	63	1	0.01
26	65	2	0.02
27	66	1	0.01
28	68	2	0.02

Строим гистограммы

In [15]:
```
df_10 = pd.read_csv("data/data_10.csv")

data_10_values = df_10['1'].to_numpy()
data_10_counts = df_10['2'].to_numpy()
data_10_rates = df_10['3'].to_numpy()
```

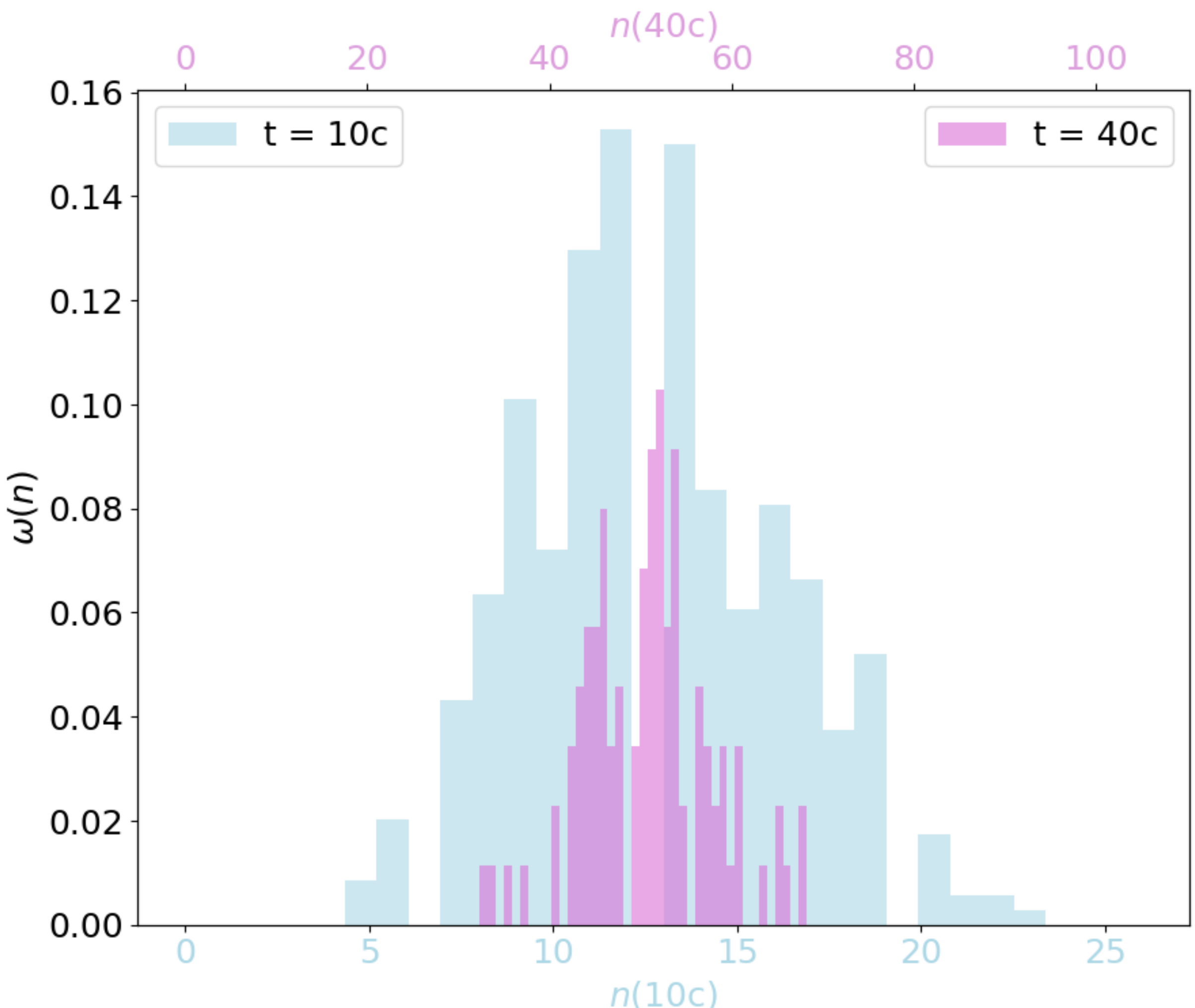
In [17]:
```
fig, ax1 = plt.subplots()
fig.set_figheight(8)
fig.set_figwidth(10)

ax1.set_xlabel('$n$(10c)', fontsize=18, color='lightblue')
ax1.set_ylabel(r'$\omega (n)$', fontsize=18)
ax1.hist(data_10_values, weights=data_10_rates, alpha=0.6, range=(0,26), bins=30, d

ax1.tick_params(axis='x', labelcolor='lightblue',  labelsize=18)
ax1.tick_params(axis='y', labelcolor='black',  labelsize=18)
plt.legend(loc='upper left', fontsize=18)

ax2 = ax1.twiny()
ax2.set_xlabel('$n$(40c)', fontsize=18, color='plum')
ax2.tick_params(axis='x', labelcolor='plum', labelsize=18)
ax2.hist(data_40_values, weights=data_40_rates, alpha=0.6 , range=(0,105), bins=120
plt.legend(loc='upper right', fontsize=18)

plt.show()
```



Расчеты

1. Найдем среднее число срабатываний счетчика за 10с:

```
In [66]: n_1 = sum(data_10_values * data_10_counts) / sum(data_10_counts)
         print(sum(data_10_values * data_10_counts), sum(data_10_counts), n_1, sep='\n')
```

```
5075
400
12.6875
```

$$\bar{n}_1 = \frac{1}{N_1}\sum_{i=1}^{N_1} n_i = \frac{5075}{400} = 12,6875$$

2. Найдем среднеквадратичную ошибку отдельного измерения:

```
In [90]: sigma_1 = np.sqrt(sum(np.power(data_10_values - n_1, 2) * data_10_counts) / sum(dat
         print(sum(np.power(data_10_values - n_1, 2) * data_10_counts), sum(data_10_counts),
```

Processing math: 100%

```
4993.9375
400
3.5333898383846636
```

$$\sigma_1 = \sqrt{\frac{1}{N_1}\sum_{i=1}^{N_1}(n_i - \bar{n}_1)^2} = \sqrt{\frac{4993,9375}{400}} \approx 3,53$$

3. Убедимся в справедливости формулы $\sigma_{\text{отд}} \approx \sqrt{\bar{n}}$:

$$\sigma_1 \approx \sqrt{\bar{n}_1}$$

$$3,53 \approx \sqrt{12,6875} \approx 3,56$$

4. Определим долю случаев, когда отклонение от среднего значения не превышают $\sigma_1, 2\sigma_1$, и сравним с теоретическими оценками:

In [87]:
```python
data_10_counts_1sigma = data_10_counts[abs(data_10_values - n_1) <= sigma_1]
data_10_counts_2sigma = data_10_counts[abs(data_10_values - n_1) <= 2*sigma_1]

p_1 = sum(data_10_counts_1sigma) / sum(data_10_counts)
p_2 = sum(data_10_counts_2sigma) / sum(data_10_counts)

print(sum(data_10_counts_1sigma), p_1)
print(sum(data_10_counts_2sigma), p_2)
```

```
253 0.6325
386 0.965
```

Доля случаев в пределах $1\sigma = \frac{253}{400} = 0.6325 \approx 0.68$

Доля случаев в пределах $2\sigma = \frac{386}{400} = 0.965 \approx 0.95$

5. Найдем среднее число срабатываний счетчика за 40с:

In [88]:
```python
n_2 = sum(data_40_values * data_40_counts) / sum(data_40_counts)
print(sum(data_40_values * data_40_counts), sum(data_40_counts), n_2, sep='\n')
```

```
5075
100
50.75
```

$$\bar{n}_2 = \frac{1}{N_2}\sum_{i=1}^{N_2} n_i = \frac{5075}{100} = 50,75$$

6. Найдем среднеквадратичную ошибку отдельного измерения:

In [91]:
```
sigma_2 = np.sqrt(sum(np.power(data_40_values - n_2, 2) * data_40_counts) / sum(dat
print(sum(np.power(data_40_values - n_2, 2) * data_40_counts), sum(data_40_counts),
```

```
4810.75
100
6.935957035622409
```

$$\sigma_2 = \sqrt{\frac{1}{N_2}\sum_{i=1}^{N_2}(n_i - \bar{n}_2)^2} = \sqrt{\frac{4810,75}{100}} \approx 6,94$$

7. Убедимся в справедливости формулы $\sigma_{\text{о т д}} \approx \sqrt{\bar{n}}$:

$$\sigma_2 \approx \sqrt{\bar{n}_2}$$

$$6,94 \approx \sqrt{50,75} \approx 7,12$$

8. Сравним среднеквадратичные ошибки отдельных измерений для двух распределений: $\bar{n}_1 = 12,6875$; $\sigma_1 = 3,53$ и $\bar{n}_2 = 50,75$; $\sigma_2 = 6,94$:

$$\frac{\sigma_1}{\bar{n}_1} * 100\% = \frac{3,53}{12,6875} * 100\% \approx 28\%$$

$$\frac{\sigma_2}{\bar{n}_2} * 100\% = \frac{6,94}{50,75} * 100\% \approx 14\%$$

9. Определим стандартную ошибку величины $\bar{n}_1$ и относительную ошибку нахождения $\bar{n}_1$ для $N = 400$ измерений по 10 с:

Processing math: 100%

$$\sigma_{\bar{n}_1} = \frac{\sigma_1}{\sqrt{N_1}} = \frac{3,53}{\sqrt{400}} \approx 0,18$$

$$\epsilon_{\bar{n}_1} = \frac{\sigma_{\bar{n}_1}}{\bar{n}_1} * 100\% = \frac{0,18}{12,6875} * 100\% \approx 1,42\%$$

$$\epsilon_{\bar{n}_1} = \frac{100\%}{\sqrt{\bar{n}_1 N1}} = \frac{100\%}{\sqrt{12,6875 * 400}} \approx 1,40\%$$

Относительные ошибки, рассчитанные по двум формулам действительно очень близки.

Тогда, окончательно имеем:

$$n_{t=10c} = \bar{n}_1 \pm \sigma_{\bar{n}_1} = 12,6875 \pm 0,18$$

10. Аналогично находим стандартную ошибку для величины $\bar{n}_2$ и относительную ошибку нахождения $\bar{n}_2$ для $N_2 = 100$ измерений по 40с:

$$\sigma_{\bar{n}_2} = \frac{\sigma_2}{\sqrt{N_2}} = \frac{6,94}{\sqrt{100}} \approx 0,69$$

$$\epsilon_{\bar{n}_2} = \frac{\sigma_{\bar{n}_2}}{\bar{n}_2} * 100\% = \frac{0,69}{50,75} * 100\% \approx 1,36\%$$

$$\epsilon_{\bar{n}_2} = \frac{100\%}{\sqrt{\bar{n}_2 N2}} = \frac{100\%}{\sqrt{50,75 * 100}} \approx 1,40\% = \epsilon_{\bar{n}_1}$$

Окончательно имеем:

$$n_{t=40c} = \bar{n}_2 \pm \sigma_{\bar{n}_2} = 50,75 \pm 0,69$$

Вывод

В ходе выполнения работы познакомился с основными понятиями статистики. Определил среднее число регистрируемых космических лучей за определенные промежутки времени и определил погрешность результата. Выяснил, что средняя интенсивность регистрируемых частиц не зависит от величины интервала τ и числа отсчетов. Исследовал Пуассоновский процесс и распределение Гаусса.