Лабораторная работа 1.1.4

ИЗУЧЕНИЕ СТАТИСТИЧЕСКИХ ЗАКОНОМЕРНОСТЕЙ НА ПРИМЕРЕ ИЗМЕРЕНИЯ ФОНА КОСМИЧЕСКОГО ИЗЛУЧЕНИЯ

Цель работы: познакомиться с основными понятиями статистики; на примере статистики регистрации фоновых космических частиц изучить статистические закономерности однородного во времени случайного процесса; проверить возможность описания исследуемого процесса статистическими законами Пуассона и Гаусса; измерить среднее число регистрируемых космических лучей в секунду и определить погрешность результата.

В работе используются: счётчик Гейгера—Мюллера, компьютер с интерфейсом для связи со счётчиком.

Введение

В любой физической лаборатории всегда присутствует радиоактивное излучение. Источником излучения являются космические лучи и распад радиоактивных веществ, которые в небольших количествах имеются всюду, в том числе в физических приборах и помещениях. Это излучение является радиоактивным фоном, с которым складывается излучение других источников, если они присутствуют. Основную часть фона обычно составляет космическое излучение.

В данной работе для регистрации космического излучения используется счётчик Гейгера-Мюллера, который представляет собой наполненный газом металлический цилиндр с двумя электродами. Одним из электродов (катодом) служит сам корпус. Другим (анодом) является тонкая нить, натянутая вдоль оси цилиндрического корпуса. Необходимое напряжение (400 В) подаётся на счётчик от смонтированного вместе с ним блока питания через повышающий трансформатор.

Космические частицы — в основном, протоны (92%), альфа-частицы (6%) и электроны/позитроны (1%) — ионизуют газ, которым наполнен счётчик, а также выбивают электроны из его стенок. Двигаясь в сильном электрическом ноле между электродами счётчика, образовавшиеся электроны соударяются с молекулами газа, выбивая из них новые — вторичные электроны. Ускоряясь полем, первичный и вторичные электроны снова ионизуют газ, и т.д. В результате образуется целая лавина электронов, через счётчик протекает кратковременный импульс тока (разряд). Этот импульс и регистрируется установкой, оцифровывается платой аналогово-цифрового преобразователя, и информация о нём через USB-интерфейс подаётся на компьютер.

Число зарегистрированных частиц зависит от времени измерения, размеров счётчика, от давления и состава газа и от материала, из которого сделаны стенки счётчика.

Статистические понятия: базовые сведения

При любом физическом измерении результат, получаемый на опыте, несколько отличается от «истинного» значения измеряемой величины. Погрешности измерений складываются из ошибок, связанных с несовершенством методики измерений и неточностью калибровки приборов (*систематические погрешности*), и из *случайных погрешностей* эксперимента, изменяющих свою величину и знак от опыта к опыту. Частным случаем случайных ошибок являются *статистические ошибки*, вызываемые *флуктуациями* (случайными колебаниями) самой измеряемой величины. К числу таких флуктуирующих величин относится и интенсивность космического излучения.

Далее будем для простоты считать, что все прочие ошибки, помимо статистических, пренебрежимо малы и рассматривать их не будем. В рамках нашего опыта это предположение хорошо выполняется.

Пусть при некотором измерении за время $\tau = 10$ с зарегистрировано n космических частиц. Из этого отнюдь не следует, что в любые следующие 10 с будет регистрироваться именно n частиц. В силу случайных причин при этом можно получить $n-1$, $n+2$ или, вообще говоря, любое другое значение (которое, как правило, всё же не слишком сильно отличается от n). Поэтому физический смысл имеет не столько результат отдельного измерения, сколько совокупность множества результатов и её усреднённые характеристики.

Наиболее важной характеристикой является среднее число регистрируемых частиц в единицу времени. Если n_1, n_2, …n_N — результаты N проведённых в одинаковых условиях измерений, можно вычислить *выборочное среднее* значение числа измерений («выборочное», поскольку определяется из ограниченной «выборки» из N значений):

$$\langle n \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} n_i \,. \qquad (1)$$

Если продолжать проводить измерения, можно ожидать, что выборочное среднее будет стремиться к некоторому конечному пределу, который можно назвать «истинным» средним значением числа регистрируемых частиц (в математике оно называется «*математическим ожиданием*» случайной величины):

$$\bar{n} = \lim_{N \to \infty} \langle n \rangle.$$

Однако, поскольку реальное число измерений всегда конечно, то и значение среднего никогда не известно точно, то есть всегда содержит *погрешность*.

Кроме среднего значения важно знать, насколько сильно флуктуируют значения n_i от опыта к опыту. Количественно меру флуктуаций принято измерять *среднеквадратичным* (или *стандартным*) *отклонением* σ_n. По определению, *средний квадрат отклонения*, называемый также *дисперсией*, (а точнее — *выборочной* дисперсией), равен

$$\sigma_n^2 \equiv \frac{1}{N}\sum_i^N (n_i - \langle n\rangle)^2\,, \qquad (2)$$

что можно короче записать как

$$\sigma_n^2 \equiv \langle (n - \langle n\rangle)^2\rangle.$$

Аналогично, при $N \to \infty$ выборочная дисперсия должна стремиться к некоторому предельному («истинному») значению:

$$\sigma^2 = \lim_{N\to\infty} \sigma_n^2 = \overline{(n - \bar{n})^2}.$$

Из теории погрешностей известно (см. [1], п. 2.4), что упомянутая выше *погрешность среднего значения* $\langle n\rangle$ при независимых измерениях связана со стандартным отклонением (погрешностью *отдельного* измерения) формулой

$$\sigma_{\langle n\rangle} = \frac{\sigma_n}{\sqrt{N}}. \qquad (3)$$

Таким образом, если среднеквадратичное отклонение σ_n стремится к конечному пределу при больших N, погрешность среднего значения убывает с ростом числа измерений как $1/\sqrt{N}$. Иными словами, увеличивая количество измерений, среднее значение можно получать со всё более возрастающей точностью, приближаясь к «истинному» $\bar{n}$. А при конечном N можно записать, что истинное среднее с высокой вероятностью лежит в интервале

$$\bar{n} = \langle n\rangle \pm \frac{\sigma_n}{\sqrt{N}}. \qquad (4)$$

Гистограммы и вероятности

Среднее и дисперсия — это очень важные характеристики, но не дающие *полной* информации о флуктуирующей величине. Более детальную информацию о ней можно получить, если собрать *статистику* того, как часто те или иные значения n встречаются среди многочисленных результатов опыта. Построим график, откладывая по оси абсцисс число частиц, зарегистрированных при измерениях, а по оси ординат — долю случаев (по отношению к общему числу измерений), в которых было зафиксировано данное количество частиц. Например, если некоторое значение n встретилось в серии из N измерений N_n раз, то по вертикали отложим отрезок высотой $w_n = N_n/N$. Построенный график содержит дискретно расположенные точки, которые для наглядности обычно соединяются между собой, изображая их в виде совокупности вертикальных прямоугольников, как это изображено на рис. 1. Высота прямоугольника равна доле наблюдаемых случаев w_n (иногда вместо доли изображается просто их число N_n), а ширина — интервалу значений n, которому эта доля соответствует (в данному случае столбики имеют единичную ширину). Такой график принято называть *гистограммой*.

В пределе $N \to \infty$ столбчатая гистограмма будет стремиться к некоторому предельному состоянию. Предельные значения частот w_n называют *вероятностями* соответствующих событий (того, что в отдельном опыте получится результат n). Для вероятностей можно строить различные теоретические модели, которые можно проверять на опыте, сравнивая практические гистограммы со значениями, предсказываемыми *теорией вероятностей*.

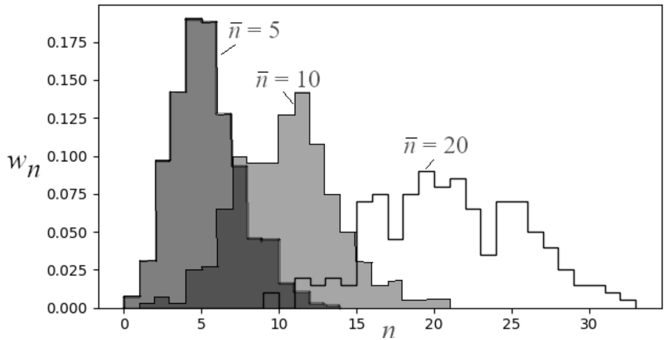

Рис. 1. Примеры результатов опыта для различных средних значений $\bar{n}$

При малых N гистограмма может довольно сильно отличаться от теоретической. По мере роста числа измерений N пик гистограммы будет приближаться к предельному среднему значению $\bar{n}$. Ширина гистограммы по порядку величины совпадает со среднеквадратичным отклонением σ_n. Если величина n близка к $\bar{n}$, её вероятность будет максимальна. А при удалении от $\bar{n}$ на расстояния, превышающие в несколько раз σ_n, вероятность, как правило, быстро падает.

Пуассоновский процесс

Если случайные события (регистрация частиц) *однородны* во времени (не меняют своей средней интенсивности), а каждое последующее событие никак *не зависит* от того, как и когда произошло предыдущее, то последовательность таких событий принято называть *пуассоновским процессом*.

Для пуассоновского процесса может быть получено теоретическое распределение вероятностей — *распределение Пуассона*. Математический вывод распределения и его свойства можно найти в любом руководстве по теории вероятностей (см., например, Приложение к работе 1.1.4 в сборнике «Лабораторный практикум…», Т.1.). Приведём здесь основные известные результаты без вывода.

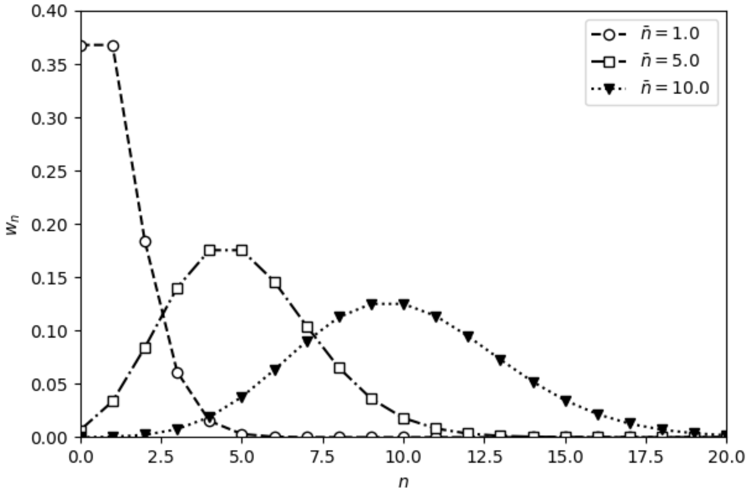

Рис. 2. Примеры теоретически рассчитанных вероятностей пуассоновского процесса для различных средних $\bar{n}$.

Вероятности w_n того, что в эксперименте будет обнаружено n частиц, для распределения Пуассона имеют вид

$$w_n = \frac{\bar{n}^n}{n!} e^{-\bar{n}}, \tag{5}$$

где $n! = 1 \cdot 2 \cdot ... \cdot n$ — факториал числа n, а $\bar{n}$ — основной (и единственный) параметр распределения, совпадающий со средним числом частиц (отметим, что значение $\bar{n}$ может быть дробным, тогда как n всегда целые). Примеры графиков функции $w_n(n)$ для различных $\bar{n}$ приведены на рис.

Из приведённых зависимостей видно, что, во-первых, при малых $\bar{n}$ ($\bar{n} \lesssim 1$) график распределения «прижат» к оси ординат и быстро убывает с ростом n. Кроме того, весьма велика вероятность не обнаружить в отдельном опыте ни одной частицы ($n = 0$). При больших же $\bar{n}$ ($\bar{n} \gtrsim 10$) график распределения стремиться к гладкой симметричной кривой, быстро убывающей к нулю при отдалении от центра. Можно строго показать, что при больших $\bar{n}$ распределение Пуассона переходит в так называемое *нормальное распределение* или *распределение Гаусса* (краткое обсуждение свойств нормального распределения и его значения для физики см. в [1], п. 2.2.).

Одним из наиболее характерным свойств распределения Пуассона является связь между его дисперсией и средним значением. А именно, для пуассоновского процесса (и только для него!) справедливо равенство

$$\sigma = \sqrt{\bar{n}}, \tag{6}$$

то есть среднеквадратичное отклонение равно корню из среднего. На практике можно ожидать приближённое равенство для *выборочных* значений:

$$\sigma_n \approx \sqrt{\langle n \rangle}.$$

Наконец, поскольку отклонения от среднего для такого процесса не велики, можно ожидать равенство *по порядку величины*

$$\sigma_n \sim \sqrt{n_i},$$

где n_i — результат любого отдельного измерения.

В качестве второго характерного свойства можно воспользоваться тем, что при достаточно больших $\bar{n}$ (на практике, $\bar{n} > 10$) распределение Пуассона приближается к *нормальному распределению* (*распределению Гаусса*). Для последнего же известно следующее характерное свойство: примерно в 2/3 случаев (с вероятностью 68%) отдельное измерение отличается от среднего значения не более, чем на одну среднеквадратичную ошибку ($\pm\sigma$), с вероятностью 95% — не более чем на две среднеквадратичные ошибки ($\pm 2\sigma$), и, наконец, теоретическая вероятность отклонения в пределах $\pm 3\sigma$ равна 99,7%.

Эти два простых свойства можно использовать для проверки того, насколько хорошо реальный случайный процесс регистрации космических частиц соответствует идеализированной теоретической модели пуассоновского процесса.

Погрешность эксперимента

Рассмотрим опыт, в котором интервал измерения t разбит на $N = t/\tau$ промежутков, длительностью τ. В качестве основного результата опыта нас прежде всего интересует среднее число частиц $\bar{n} \approx \langle n \rangle$, регистрируемое за время τ. Подставим основное свойство распределения Пуассона (6) в формулу (3). Получим среднеквадратичную погрешность определения среднего:

$$\sigma_{\langle n \rangle} = \frac{\sigma_n}{\sqrt{N}} = \sqrt{\frac{\langle n \rangle}{N}}.$$

Обычно больший интерес представляет не абсолютное, а *относительное* значение погрешности. Для него находим:

$$\varepsilon_{\langle n \rangle} = \frac{\sigma_{\langle n \rangle}}{\langle n \rangle} = \frac{1}{\sqrt{\langle n \rangle N}}.$$

В знаменателе полученного выражения, как нетрудно видеть из (1), стоит *полное число частиц* $N_0 = \langle n \rangle N = \sum n_i$, зарегистрированных за всё время измерений t. То есть относительная погрешность опыта не зависит от интервалов τ разбиения серий, и убывает обратно пропорционально корню из общего числа частиц N_0. Этого, конечно, и следовало ожидать, так как все измерения вместе составляют одно более продолжительное измерение, в котором зарегистрировано $\sum n_i = N_0$ отсчётов.

Таким образом, единственный способ увеличить точность опыта — увеличивать общее число регистрируемых частиц за счёт увеличения совокупного времени измерений t. Например, для достижения точности измерения интенсивности фона в 1% необходимо зарегистрировать в общей сложности не менее 10^4 частиц.

ЗАДАНИЕ

Основной эксперимент

1. Ознакомьтесь с устройством измерительного прибора (счётчик Гейгера—Мюллера). Убедитесь, что он подключён к сети и к USB-порту компьютера. Включите питание прибора. Периодически моргающий зелёный светодиод означает, что счётчик регистрирует пролетающие через него космические частицы. Включите компьютер и запустите программу (ярлыком на рабочем столе).
2. Перейдите в раздел «Эксперимент», выберите «Запуск эксперимента» и введите свои данные для создания папки, в которой будут сохраняться данные эксперимента и симуляций. Если графики в окне программы начнут обновляться, это означает, что эксперимент успешно запущен. При возникновении проблем с запуском обратитесь к лаборанту или преподавателю. Длительность эксперимента по умолчанию $t = 4000$ с (можно изменить в настройках программы до запуска эксперимента), число зарегистрированных частиц записывается с интервалом $\tau_0 = 1$ с (всего $N = t/\tau_0 = 4000$ точек).
3. В процессе сбора данных можно изучать графики и гистограмму процесса, а также выполнить «демонстрационное задание» (см. ниже). При возврате в главное меню эксперимент не прервётся (прерывание состоится только при выходе из программы или нажатии кнопки «Остановка эксперимента»). Все данные сохраняются автоматически в отдельный файл.
4. После завершения эксперимента найдите файл с записанными данными и скопируйте его себе. Положение файлов по умолчанию: Рабочий стол -> папка с именем «Фамилия_Группа».
5. Обработайте экспериментальные данные согласно заданию «Обработка данных» (см. ниже).

Демонстрационное задание

6. Нажав кнопку «Назад», вернитесь в главное меню и перейдите в раздел «Симуляция». Симуляция позволяет в ускоренном режиме наблюдать процесс сбора статистических данных, порождаемых генератором псевдослучайных чисел.
7. Подготовьте запуск симуляции с настройками по умолчанию (распределение Пуассона, интенсивность — 1,0, ускорение — 10). «Ускорение» задаёт количество данных, порождаемых генератором в одну секунду (т.е. при ускорении 10

«компьютерные» часы идут в 10 раз быстрее «реальных»). При желании «ускорение» можно увеличить (например, до 100).

Дополнительно при изучении результатов можно изменять параметр «Группировка» — он задаёт длительность τ, в течение набираются данные для отдельной экспериментальной точки (соседние данные для числа части за 1 секунду группируются в блоки и суммируются за τ секунд).

8. Запустите симуляцию (кнопка «Запуск симуляции»). В процессе сбора демонстрационных данных наблюдайте (качественно), как изменяются в зависимости от количества N собранных экспериментальных точек:
   а. среднее число зарегистрированных частиц $\langle n \rangle$;
   б. среднеквадратичное отклонение $\sigma_n = \sqrt{\langle (n - \langle n \rangle)^2 \rangle}$;
   в. насколько хорошо выполняется свойство распределения Пуассона $\sigma_n = \sqrt{\langle n \rangle}$;
   г. как изменяется общий вид гистограммы и величина флуктуаций отдельных столбцов и какова степень совпадения экспериментальной гистограммы с теоретическим кривыми (распределениями Пуассона и Гаусса).

   Рассмотрите также, как изменяются вышеперечисленные параметры при изменении параметра группировки результатов τ (рекомендуется использовать значения $\tau = 1; 10; 20; 80$).
9. Сделайте выводы о наблюдаемых зависимостях и запишите их в лабораторном журнале. Сохраните или сфотографируйте несколько (4-5) характерных изображений гистограмм и других графиков, иллюстрирующих ваши выводы, и приложите их к отчёту о работе.

   Вы можете сохранить данные симуляции либо в текстовом формате (кнопка «Сохранить данные», данные сохранятся в ту же папку, что и для основного эксперимента), либо сохранить любой график в формате изображения (правая кнопка мыши на графике -> Export, далее выбрать формат и место сохранения файла).
10. Запустите несколько демонстраций в максимально ускоренном режиме (ускорение > 1000), изменяя среднюю интенсивность числа частиц (параметр μ). Как при этом изменяются наблюдаемые гистограммы? Сохраните или сфотографируйте несколько (3-4) характерных изображений гистограмм и опишите результат в лабораторном журнале.
11. *Изучите влияние характера распределения случайной величины на итоговую статистику результатов. Как изменяется характер результатов (вид гистограммы) при использовании вместо распределения Пуассона а) экспоненциального распределения, б) распределения Парето при различных показателях α (попробуйте $\alpha = 1{,}0$ и $\alpha = 2{,}0$). Стремятся ли

гистограммы результатов к распределению Гаусса при большом числе измерений $N \gg 1$? Для любых ли распределений можно говорить, что выборочные среднее или среднеквадратичное отклонение стремятся при больших N к некоторым определённым значениям?

Обработка результатов

12. Откройте файл с экспериментальными данными. Файл имеет текстовый формат, где все значения отделены новой строкой (кроме строк с комментариями, начинающимися с символа #). Для обработки данных рекомендуется использовать электронные таблицы (например, *LibreOffice Calc*) или любые другие доступные вам инструменты (выбранные самостоятельно или по рекомендации преподавателя).
13. Сгруппируйте и просуммируйте соседние данные с различными интервалами группировки, например: $\tau = 10$ с; 20 с; 40 с; 80 с (используйте не менее 3-х значений τ). Для каждого разбиения вычислите частоты w_n, с которыми каждое число отсчётов n встречается среди результатов.
14. Для каждого τ постройте столбчатые (ступенчатые) гистограммы результатов, откладывая по оси абсцисс число отсчётов n, а по оси ординат столбики, высота которых соответствует частоте w_n. Подберите (по своему усмотрению или по заданию преподавателя) масштабы и взаимное расположение гистограмм так, чтобы их было удобно сравнить друг с другом (в том числе, можно изобразить несколько гистограмм на одном графике).
15. Для каждого τ вычислите а) среднее число регистрируемых частиц $\langle n \rangle$, б) среднеквадратичное (стандартное) отклонение σ_n, в) погрешность среднего значения $\sigma_{\langle n \rangle}$, г) среднюю интенсивность регистрируемых частиц в секунду $j = \langle n \rangle / \tau$ и её погрешность σ_j. Сделайте выводы о зависимости этих результатов от величины интервала разбиения τ и числа точек $N = t/\tau$.
16. Используя найденные значения $\langle n \rangle$ и σ_n, наложите поверх экспериментальных гистограмм (как минимум, на одну из гистограмм) теоретические распределения Пуассона или Гаусса (по собственном выбору или по указанию преподавателя). Оцените (качественно), насколько хорошо экспериментальные гистограммы согласуются с теоретическими.
17. Проверьте справедливость основного свойства распределения Пуассона $\langle n \rangle = \sigma_n$. Можно ли утверждать, что данное равенство выполняется с удовлетворительной точностью? Какой вывод можно сделать о характере распределения регистрируемых частиц?

18. Определите доли случаев, когда отклонение числа отсчётов n от среднего значения не превышает (по модулю) одного, двух и трёх стандартных отклонений: $|n - \langle n \rangle| \leq \sigma_n$, $|n - \langle n \rangle| \leq 2\sigma_n$ и $|n - \langle n \rangle| \leq 3\sigma_n$. Сравните результаты с теоретическими для распределения Гаусса. Какой вывод можно сделать о характере распределения регистрируемых частиц?

Литература

1. Попов П.В., Нозик А.А. «Обработка результатов учебного эксперимента», 2019

Составитель:
Попов П.В., 03.09.2023
(на основе предыдущей версии,
Кузьмичев С.Д., Долженко В.Ю., Соркин В.Ю., 1997)

Авторы программы:
Лапушкин А.Г., Попов П.В.