Отношение сигнал/шум в астрофизических наблюдениях

Лекция 4

Фотонный и пуассоновский шум. Статистика счёта фотонов и накопление сигнала

КазНУ имени аль-Фараби

Цель и план лекции

Цель

Изучение статистики Пуассона как основы счёта фотонов, анализ закона накопления сигнала и приобретение навыков вычисления доверительных интервалов при малом числе зарегистрированных событий.

1. Введение
2. Распределение Пуассона
3. Физические причины и отклонения от пуассоновского закона
4. Квантовая эффективность и накопление
5. Малые числа событий
6. Применение счёта фотонов
7. Статистика фотонов при обработке данных

1 Введение

- Свет регистрируется детектором в виде отдельных квантов. Даже при неизменном потоке излучения число фотонов, поступивших за заданный интервал времени, случайно.
- Особенно важна пуассоновская статистика в рентгеновской и гамма-астрономии, где число зарегистрированных квантов может составлять единицы, а также при наблюдениях слабых объектов с помощью счётчиков фотонов.

2 Распределение Пуассона

- Если события происходят независимо со средней интенсивностью r, то число событий k за время t подчиняется распределению Пуассона:

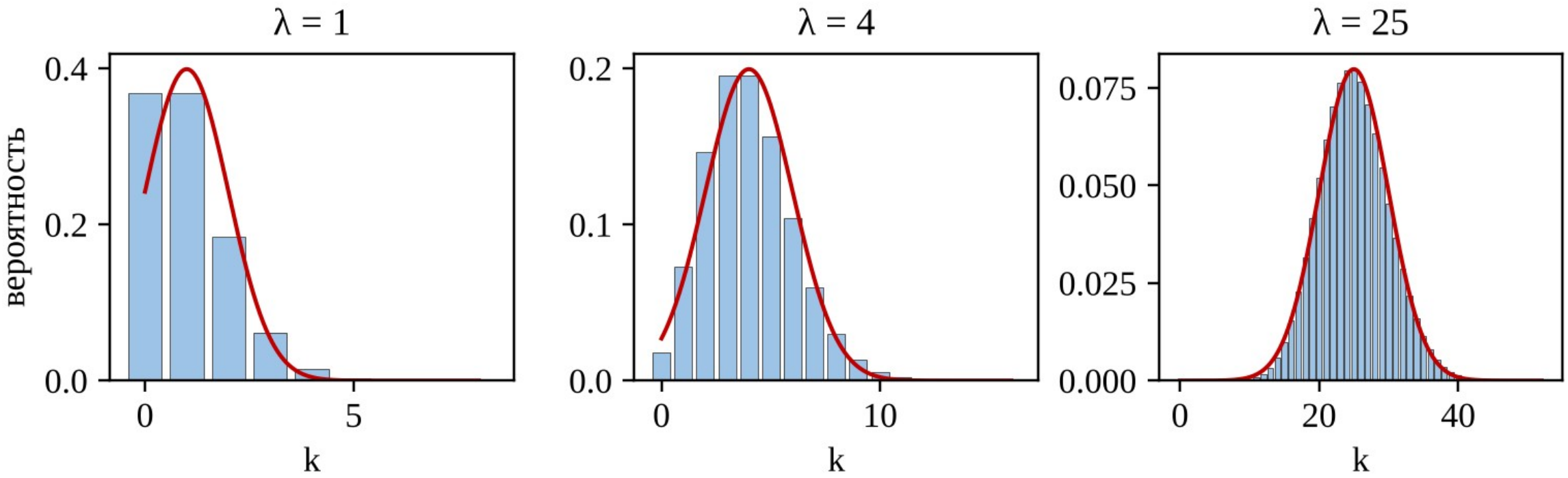
$$P(k;\lambda) = \lambda^k e^{-\lambda} / k!, \quad \lambda = r\,t.$$

- Основное свойство состоит в равенстве математического ожидания и дисперсии: E[k] = D[k] = λ.

$$\sigma = \sqrt{\lambda}, \quad S/N = \sqrt{\lambda}.$$

- Таким образом, для получения S/N = 100 необходимо зарегистрировать 10000 электронов.

2 Распределение Пуассона (продолжение)



Распределение Пуассона при различных значениях λ и его нормальное приближение

2 Распределение Пуассона (продолжение)

- Сумма двух независимых пуассоновских величин также распределена по закону Пуассона со средним, равным сумме средних.

3 Физические причины и отклонения от пуассоновского закона

- Фотоны теплового источника в действительности группируются, что описывается статистикой Бозе–Эйнштейна.

$$\sigma^2 = \bar{n} + \bar{n}^2.$$

- Первое слагаемое описывает дробовой (корпускулярный) шум, второе – волновой шум.
- Распределение может отличаться от пуассоновского из-за особенностей детектора.

3 Физические причины и отклонения от пуассоновского закона (продолжение)

Фактор Фано

Для рентгеновских фотонов один квант создаёт много электронов; дисперсия их числа меньше пуассоновской и определяется фактором Фано (F ≈ 0,12 для кремния).

Умножение заряда

Для ЭПЗС (EMCCD) процесс умножения вносит дополнительный шум, что эквивалентно увеличению дисперсии вдвое.

Мёртвое время

При больших потоках счётчик не успевает регистрировать все фотоны, и наблюдаемая статистика становится субпуассоновской.

Нестационарность

Изменение прозрачности или переменность приводит к сверхпуассоновскому разбросу.

4 Квантовая эффективность и накопление

- Детектор регистрирует только часть падающих фотонов, определяемую квантовой эффективностью QE. Прореживание потока Пуассона также даёт пуассоновский поток со средним QE·λ.

$$S/N = \sqrt{(QE \cdot \lambda)}.$$

- Обычно эффективность всей системы включает пропускание атмосферы, зеркал, фильтра и QE детектора.
- Если объект наблюдается на фоне с интенсивностью b, необходимо раздельно измерить сигнал с фоном (on) и фон (off).

5 Малые числа событий

- При малом числе зарегистрированных событий нельзя указывать погрешность как √N. Герельс (1986) привёл приближения и таблицы доверительных интервалов.

$$\lambda_{low} = \tfrac{1}{2}\,\chi^2_{\alpha/2}(2k), \quad \lambda_{up} = \tfrac{1}{2}\,\chi^2_{1-\alpha/2}(2k+2).$$

- Если k = 0, то верхний предел 95 % односторонний равен примерно 3. Функция правдоподобия также применяется для подбора моделей: статистика Кэша (1979) $C = 2\Sigma(m_i - k_i \ln m_i)$ заменяет χ^2 при малом числе отсчётов.

5 Точные интервалы и приближение корня

Листинг 4.1.

```
import numpy as np
from scipy import stats

def poisson_ci(k, cl=0.6827):
    a = 1 - cl
    lo = 0.0 if k == 0 else 0.5 * stats.chi2.ppf(a / 2, 2 * k)
    up = 0.5 * stats.chi2.ppf(1 - a / 2, 2 * k + 2)
    return lo, up

print(" k    точный 68% интервал    k +- sqrt(k)")
for k in (0, 1, 3, 10, 100, 1000):
    lo, up = poisson_ci(k)
    print(f"{k:4} [{lo:9.2f}, {up:9.2f}]   "
          f"[{k - np.sqrt(k):8.2f}, {k + np.sqrt(k):8.2f}]")
print(f"Верхний предел 95% при k=0: {-np.log(0.05):.2f}")
```

5 Результат: Точные интервалы и приближение корня

```
 k    точный 68% интервал    k +- sqrt(k)
   0 [     0.00,      1.84]   [    0.00,     0.00]
   1 [     0.17,      3.30]   [    0.00,     2.00]
   3 [     1.37,      5.92]   [    1.27,     4.73]
  10 [     6.89,     14.27]   [    6.84,    13.16]
 100 [    90.02,    111.03]   [   90.00,   110.00]
1000 [   968.38,   1032.63]   [  968.38,  1031.62]
Верхний предел 95% при k=0: 3.00
```

5 Накопление сигнала и разбиение экспозиции

Листинг 4.2.

```
rng = np.random.default_rng(4)
rate, T = 200.0, 100.0

one = rng.poisson(rate * T, 200000)
split = rng.poisson(rate * T / 10, (200000, 10)).sum(1)

print(f"Одна экспозиция: среднее {one.mean():.0f}, "
      f"СКО {one.std():.1f}, S/N {one.mean() / one.std():.1f}")
print(f"10 частей: среднее {split.mean():.0f}, "
      f"СКО {split.std():.1f}, S/N {split.mean() / split.std():.1f}")

RN = 10.0
noisy = rng.normal(split, np.sqrt(10) * RN)
print(f"10 частей с RN=10: S/N {noisy.mean() / noisy.std():.1f}")
print(f"Теория: {rate * T / np.sqrt(rate * T + 10 * RN**2):.1f}")
```

5 Результат: Накопление сигнала и разбиение экспозиции

```
Одна экспозиция: среднее 20000, СКО 141.5, S/N 141.4
10 частей: среднее 20000, СКО 141.2, S/N 141.6
10 частей с RN=10: S/N 138.2
Теория: 138.0
```

Результат показывает, что при отсутствии шума считывания разбиение не изменяет S/N, но уже при значении RN = 10 e^- с 10 считываниями значение немного снижается.

6 Применение счёта фотонов

Счёт фотонов применяется в разных областях.

Рентгеновская астрономия

Для каждого кванта регистрируются время, координаты и энергия. Массивы фотонных событий обрабатываются методами максимального правдоподобия.

Высокоскоростная фотометрия

Устройства с малым шумом считывания позволяют исследовать быстрые изменения блеска.

Поляриметрия

Разность интенсивностей двух лучей мала, поэтому требуется большое число фотонов: для измерения поляризации с точностью 0,1 % необходимо порядка 10^6 зарегистрированных фотонов.

Фотонный шум задаёт теоретически достижимую точность, называемую нижней границей Крамера–Рао.

6 Применение счёта фотонов (продолжение)

Отсчёт с нулевым фоном

Для космических телескопов малые числа событий позволяют достичь высокой значимости: S/N определяется вероятностью случайного появления такого числа фотонов.

Фотонный шум задаёт теоретически достижимую точность, называемую нижней границей Крамера–Рао.

7 Статистика фотонов при обработке данных

- Пуассоновская природа данных накладывает ограничения на корректность обработки.
- Для пикселей с малым числом электронов шум считывания и шум фотонов имеют разную статистику: распределение значений представляет собой свёртку пуассоновского и нормального распределений.
- Вариационно-стабилизирующее преобразование Анскомба преобразует пуассоновскую величину x в величину $2\sqrt{(x + 3/8)}$, дисперсия которой почти равна единице.
- Для рентгеновских наблюдений обычно проводится группировка энергетических каналов так, чтобы в каждом бине было не менее заданного количества отсчётов (например, 20), что позволяет применять χ^2.

7 Статистика фотонов при обработке данных (продолжение)

- В качестве упражнения вычислите, сколько фотонов необходимо для определения блеска с точностью 1 %, 0,1 % и 0,01 %.

8 Разбор примера

1. Пусть зарегистрировано $k = 9$ фотонов при ожидании фона 1 фотон.
2. Оценка источника равна 8.
3. Приближённая погрешность равна $\sqrt{9} = 3$, поэтому $S/N \approx 8/3 \approx 2{,}7$.
4. Однако более информативно вычислить вероятность случайного появления 9 или более отсчётов при $\lambda = 1$: она имеет порядок 10^{-6}, поэтому событие значимо.
5. Этот пример показывает, что при малых числах значимость нельзя определять просто как отношение $S/\sqrt{N}$.
6. Если $S/N = 25$ достигается за 100 с, то при фотонном шуме значение $S/N = 50$ будет достигнуто за 400 с.

9 Вопросы для обсуждения на занятии

Предел

Может ли измерение быть точнее фотонного предела? Обсудите сжатый свет и дополнительную информацию.

Малые числа

Почему при малых k нельзя использовать простое выражение √k для погрешности?

Детектор

В каких режимах наблюдения выгодно использовать EMCCD, несмотря на избыточный шум?

Контрольные вопросы

1. Запишите распределение Пуассона и его основные свойства.
2. Почему S/N растёт как √N при фотонном шуме?
3. В каких случаях справедливо приближение нормальным распределением?
4. Чем определяются отклонения от пуассоновской статистики?
5. Как квантовая эффективность влияет на S/N?
6. Как вычитание фона влияет на дисперсию оценки?
7. Как определить доверительный интервал при малом числе событий?
8. Что такое граница Крамера–Рао для определения положения источника?

Список литературы

1. Gehrels N. Confidence limits for small numbers of events in astrophysical data // The Astrophysical Journal. – 1986. – Vol. 303. – P. 336–346.
2. Cash W. Parameter estimation in astronomy through application of the likelihood ratio // The Astrophysical Journal. – 1979. – Vol. 228. – P. 939–947.
3. Wall J.V., Jenkins C.R. Practical Statistics for Astronomers. – 2nd ed. – Cambridge: Cambridge University Press, 2012.
4. Bevington P.R., Robinson D.K. Data Reduction and Error Analysis for the Physical Sciences. – 3rd ed. – New York: McGraw-Hill, 2003.
5. Ivezić Ž., Connolly A.J., VanderPlas J.T., Gray A. Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data. – Updated ed. – Princeton: Princeton University Press, 2020.