Lab #1: Basic Photon Statistics

Lab report is due Wednesday, October 10, 2018, before 11:59 pm EDT

1. Overview

This handout provides a description of the activities that we will explore in the first lab. Because we use this first lab to introduce many new skills, especially Python programming, this handout is a straightforward guide that your lab group should follow sequentially.

1.1 Schedule

This is a four-week lab with activities and lectures on September 10, 17, 24, and October 1. The report due is Oct 10. (Note that a late submission is unacceptable.) See §5 for detailed steps and milestones.

1.2 Goals

Learn how conduct simple statistical analysis of sample data using Python programming language. This will be a stepping-stone to more advanced analysis.

1.3 Reading assignments (available online on course website)

- Linux
    - To master the first lab, you will need basic Linux system skills so that you can collect data and manipulate files.
- Python
    - Work through the Python/IPython tutorials to learn how to use this powerful programming language to compute statistical quantities and make plots. Check out the lab we page and "Getting started" at:
        - https://www.scipy.org/getting-started.html
    - Fundamental Python references are:
        - Numerical computing: http://www.numpy.org
        - Plotting: http://matplotlib.org
- Document preparation
    - You can use the TeX/LaTeX document preparation system to create publication quality reports. Download the examples and make sure that you can generate PDF output. You can then use this example as a template for your first report.
- Skim the statistics handout on the class web page.

1.4 Key steps

You will execute eight key steps in this lab:

1. Understand how photoelectric effect is used for the detection of photons of visible wavelengths with a charge-coupled device.
2. Familiarize yourself with basic statistical principles and noise performance needed for professional data reduction.
3. Plot histograms to visualize the data and compute the mean and standard deviation to understand the data properties.
4. Understand the Poisson distribution and investigate if the data is represented by the distribution.

5. Write your lab report.

2. Photo-electric Effect and Charge Coupled Devices (CCDs)

Photo-Electric Effect: At the wavelengths of visible, near-infrared, ultra-violet and X-ray, most astronomical instruments employ the photoelectric effect to convert photons into electrons, which can then be counted and recorded. Since the 1930's astronomers have used photomultiplier tubes, which employ the photoelectric effect in vacuum. Modern detectors, such as CCDs, are based on the solid-state devices constructed from semiconductor materials, but still employing the photo-electric effect as the basic method of detecting photons from astronomical sources. Note that the process of photon detection and counting is not perfect–detectors have flaws that introduce noise and various systematic errors into the measurement process.

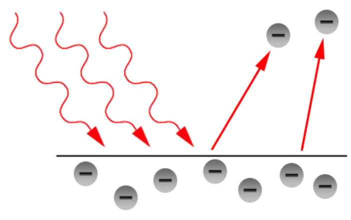

Figure 1: Basic image illustrating the photoelectric effect.

CCD Image Sensor: A CCD sensor is usually a two-dimensional array of pixels where charges are shifted down each column one row at a time to a horizontal shift register. (Note that there are also one-dimensional linear CCD arrays.) Charge is generated in each pixel by the photoelectric effect in proportion to the number of incoming photons incident at that location. The charge within pixels in each row is then shifted onto a capacitor and read one pixel at a time. This voltage is then amplified by a preamplifier within the sensor and by another adjustable gain amplifier within the camera. Finally, this analog signal is digitized by an analog-to-digital convertor, and the data is conveyed to a computer. The digital signal is recorded in units known as analog to digital units (ADU) after being amplified by both the sensor and the camera electronics as shown in Figure 1. Assuming that the measuring circuits are linear there is a linear relationship between ADU and the number of photoelectrons collected; this constant of proportionality is known as the **gain**. The entire CCD array is reset before the start of an exposure by clearing any remaining charge built up on the sensor. This establishes a well-defined charge state at the start of the subsequent exposure. The photo-charge (Q) is determined by measuring the voltage (V) induced on a capacitor of capacitance (C)

$$V = Q/C \qquad (1)$$

This voltage is therefore directly proportional to the number of incident photons.

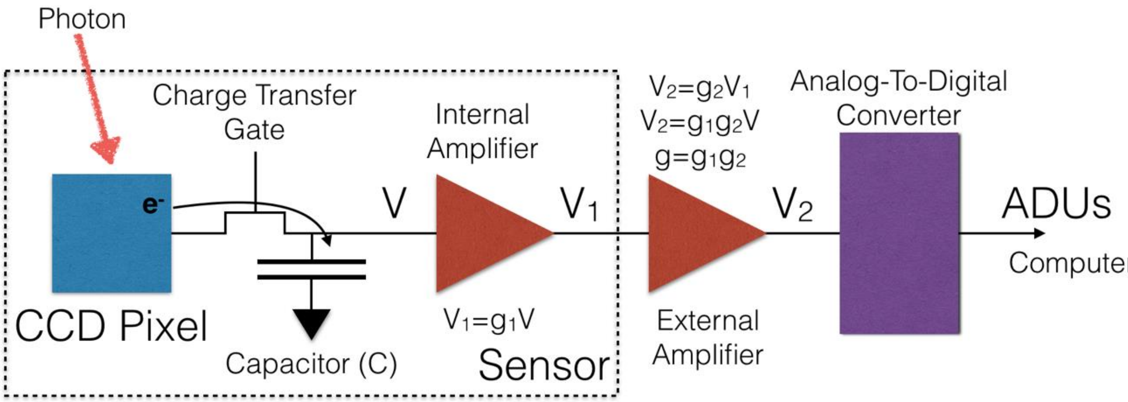

Figure 2: Analog signal chain for a CCD sensor.

In order to understand the photon-detection process, we first need to know how to analyze the obtained data. **The purpose of this lab is to learn basic computational skills needed for photon statistics, which will serve as a stepping-stone for the forthcoming labs.** We will learn more about the CCDs and real data acquisition using a CCD in Lab #3.

3. Data Visualization and Calculation of Basic Statistical Parameters

(Mean, Median, Standard Deviation and Weighted Median)

Let's assume that you have conducted a series of astronomical observations using a CCD in order to determine the distance to a nearby star. As below, you repeated the same experiments 30 times and obtained the following results – the number is the distance (in parsec) to the star obtained in each trial, whereas the number in the parenthesis is the measurement error.

Distance Measurement:

38.91 (1.41)	37.14 (0.36)	38.19 (0.69)	41.03 (3.53)	34.86 (2.64)
37.33 (0.17)	35.16 (2.34)	37.96 (0.46)	36.93 (0.57)	40.41 (2.91)
29.50 (8.00)	37.33 (0.17)	41.84 (4.34)	37.53 (0.03)	34.12 (3.38)
34.11 (3.39)	37.94 (0.44)	34.43 (3.07)	36.68 (0.82)	41.31 (3.81)
39.61 (2.11)	35.48 (2.02)	34.98 (2.52)	39.05 (1.55)	39.62 (2.12)
37.96 (0.46)	39.02 (1.52)	37.47 (0.03)	33.76 (3.74)	36.51 (0.99)

The first thing you want to do with the data is to see the distribution of the measurements. Figure 3 shows the plot comparing all the 30 measurements, while Figure 4 shows a histogram of it. Do you see the difference? Which is better? The next step is to calculate some basic parameters such as the mean and standard deviation (STDDEV), and I have the following numbers:

$$\text{Mean} \pm \text{STDDEV} = 37.21 \pm 2.65 \text{ (pc)}$$

However, each measurement has a different error, and the measurements with a smaller error should be more reliable than those with a large error. Using weighted mean statistics, now I have the following numbers:

$$\text{(Weighted) Mean} \pm \text{STDDEV} = 37.50 \pm 0.02 \text{ (pc)}$$

Now you can see that the error (= standard deviation) in the weighted mean is much smaller and more reliable.

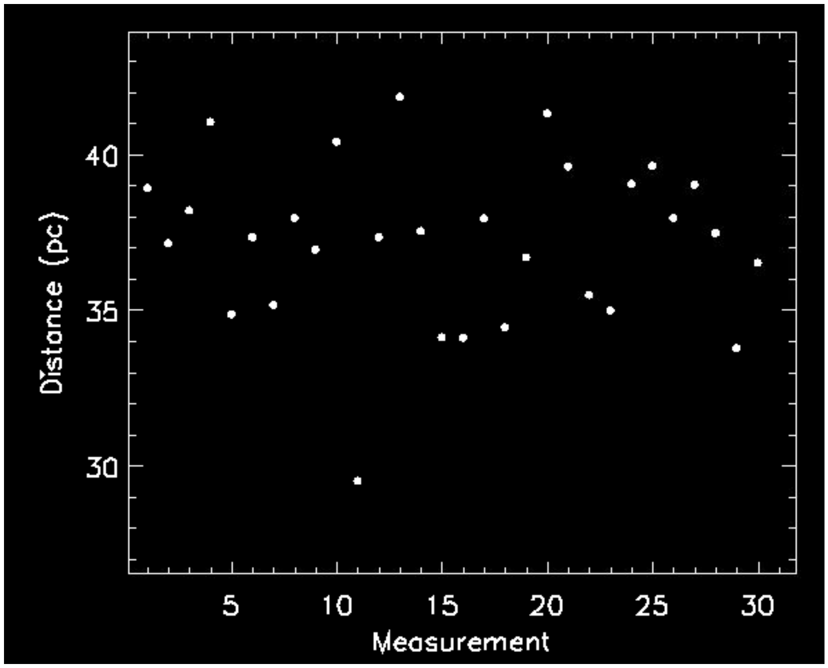

Figure 3. Distribution of the 30 obtained distances.

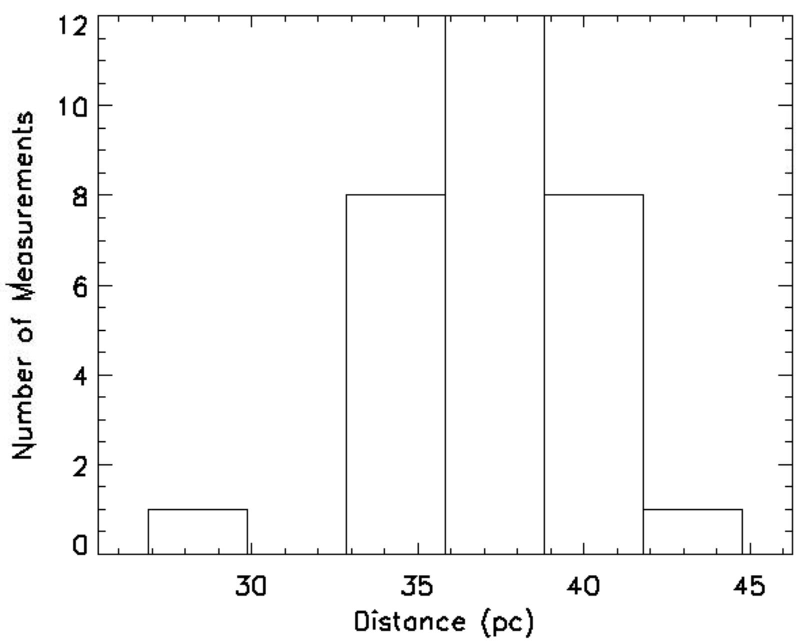

Figure 4. Histogram of the 30 distance measurements in Figure 3.

4. Poisson Distribution and Time Series Data

Poisson distribution: Often astronomical signal follows the Poisson statistics as below

$$P(x;\mu) = \frac{\mu^x}{x!}e^{-\mu} \quad (2)$$

where *P(x; μ)* is the probability *x* events will be measured within an interval given a mean of μ events. (Refer to lecture notes for the details of the Poisson distribution). In comparison, the Gaussian probability distribution follows the following form

$$P(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \quad (3).$$

Under certain conditions, the Gaussian probability distribution can be a reasonable approximation of the Poisson distribution.

Now let's assume that you measured the number of incoming photons from an astronomical source per second 30 consecutive times as below. The expected average count per second from this source is 12 counts/sec.

Photon Count Rate (= Number of incoming photons per second)

13	17	18	14	11	8	21	18	9	12
9	17	14	6	10	16	16	11	10	12
8	20	14	10	14	17	13	16	12	10

First, let's calculate the mean and standard deviation as follows:

Mean ± STDDEV = 13.2 ± 3.8 (counts/second)

The Poisson distribution is known to have the following relation between the mean and standard deviation: Mean = $(\text{STDDEV})^2$. Figure 5 below compares a histogram of the measured photon count rate (solid line) and the expected Poission distribution with μ = 12. Based on the calculated mean and STDDEV and the comparison in Figure 2, do you think the obtained data follow the Poission distribution?

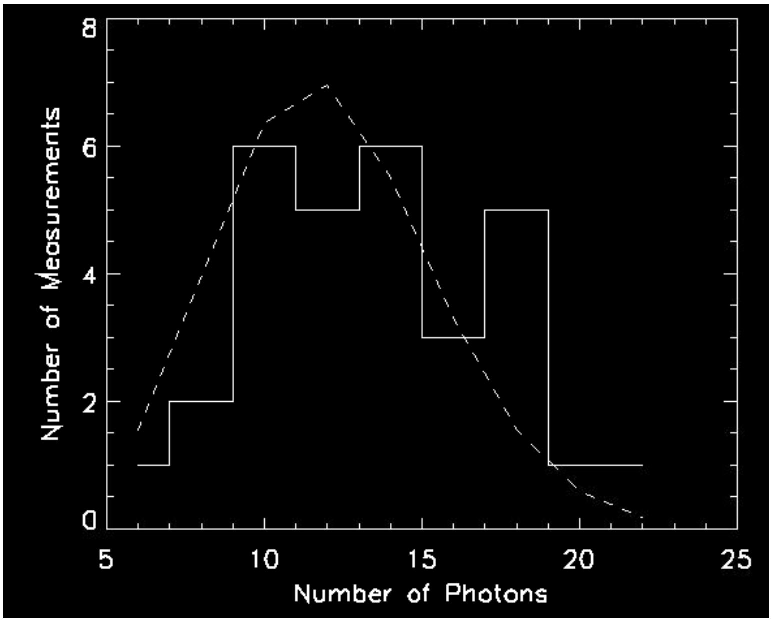

Figure 5. (Solid line) Histogram of the photon count rate; (Dashed line) Expected Poisson distribution with $\mu = 12$.

5. Lab #1 Assignment Steps and Milestones

Below is a list of the steps that you need to accomplish in Lab #1.

(1) Using your own Python programs, repeat §3 by calculating the mean and stddev, weighted mean and stddev. Also make your own plots of Figure 3 and 4.
(2) As above, repeat §4 using your own Python programs by calculating the mean and stddev and making your own plot of Figure 5.
(3) Retrieve the data for your group from the class web page and perform the analysis as described.

Steps (1) and (2) above by September 24. Also you should be able to use Latex for documentation by this day. Step (3) above by October 1 and you should begin writing a report by this day. Remember that writing always takes longer than you expect, so **start early!**