Basic Photon Statistics

Luka Vujeva
luka.vujeva@mail.utoronto.ca

October 2, 2016

Abstract

In this lab, the statistics behind the Photoelectric effect are examined through using the CCD's, or Charge Coupled Devices. The data sets from the various experiments were analysed using both the Poisson and Gaussian distributions, along with comparing the mean and weighted mean with their respective standard deviations. By using statistical analysis over the various data sets, it will be shown that the more number of measurements there are, the closer the Poisson distribution will look like the Gaussian distribution.

1 Introduction

The photoelectric effect is the process in which electrons are emitted when a photons hit a surface. This occurs when the energy absorbed by the electron is greater than the work function of the material, which is essentially the minimum energy required for an electron to escape the material. To measure this effect, we use CCD's, or Charge Coupled Devices, which are especially useful because they have an extremely high quantum efficiency. This means that there is an extremely high ratio of incident photons to converted electrons. In this experiment, we aim to use data, collected using CCD's, of distant stars. From that data we will show that both the data can be modelled very well using both the Gaussian and Poisson distributions.

The programming in this experiment was done individually however there was collaboration on general strategy and theory between the author (Luka Vujeva), and lab partners Pierce Balko, Dong Uk Kim, and Cameron Stockton.

2 Observations & Data

The data for this lab was not acquired by the students, however 4 total sets of data were used to perform the necessary procedure. Two of the data sets were provided in the lab report itself, which were then copy pasted into a text file for python to read. The other two data sets were provided per group specifically. Both data sets were collected using CCD's, or Charged Coupled Devices that use the photoelectric effect to allow us to calculate the amount of incident photons coming in, which then allows us to calculate the distance to the star we are measuring. Error in the case of this lab was provided in the required sections. In the case of the distance measurements, it was measurement error.

3 Data Reduction & Methods

Data was acquired through two main sources: from copying data from the lab handout and pasting into a plain text file, and from plain text files provided on the course website. For the distance data set in section 3 of the lab handout, the distances were in a nice column however the errors were provided in brackets. This would cause major issues with python once imported since python won't ignore the brackets surrounding the values so you would get errors when trying to perform any kinds of calculations. The solution to that problem was individually deleting the brackets surrounding the error values after pasting them into the plain text files. Another form of data cleaning that was done was converting the vertical lists of data into horizontal arrays using the python command np.ndarray.flatten(). This was done to prevent any unexpected errors when performing calculations on vertical arrays, and to ease plotting the data.

Finding the mean, standard deviation, weighted mean and weighted standard deviation of various data sets were also a critical part to this lab. This was done using a series of python functions based on the formulae given in the lecture slides. The first equation was the equation for mean where μ is the mean, is the number of measurements

$$\mu = \frac{1}{N}\left(\sum_{i=1}^{N} x_i\right) \tag{1}$$

Next we had to find the weighted mean which is given by the equation

$$x = \frac{\sum\limits_{i=1}^{n} w_i x_i}{\sum\limits_{i=1}^{n} w_i} \tag{2}$$

Next we have the equation for standard deviation, where σ represents standard deviation

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{3}$$

4 Data Analysis and Modeling

Given that the purpose of this lab was to look at how statistics coincide with the behaviour of photons using Gaussian and Poisson distributions, we must define them. First, the Poisson distribution can be described as

$$P(x) = \frac{e^{-\mu}\mu^x}{x!} \tag{4}$$

The Gaussian distribution can be described as the equation

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2} \tag{5}$$

It is also worth noting that the Poisson distribution is explicit where as the Gaussian distribution is a continuous distribution.

5 Discussion

QUESTION 1: As you can see in figures 1 and 2, the plot and histogram look exactly like the examples given in the lab handout. Not only that, but interestingly enough, we can see that although the plot does not seem to follow any standard distribution upon first glance, figure 2 follows a Gaussian or Poisson distribution extremely accurately. After calculations of mean and standard deviation that can be seen in the Appendices, we can see that the Mean $\pm$ STDDEV = 37.21 $\pm$ 2.65 (pc), just as in the problem in the question. Also, the weighted mean that was calculated using our own equations was (Weighted) Mean $\pm$ STDDEV = 37.50 $\pm$ 0.02 (pc), just as in the problem.

QUESTION 2: After the mean and standard deviation of the data to be Mean $\pm$ STDDEV = 13.2 $\pm$ 3.8 (counts/second), just as in the question, we see that figure 3 loosely follows a Poisson distribution. The explanation of why the fit is so poor could stem from the small amount of measurements done. As we can see in later graphs such as figures 5 and 7, the higher the count of measurements, the better the Poisson and Gaussian Distributions fit to the data.

QUESTION 3: For the two data sets that we were given, the small and large data sets, their individual means and standard deviations were calculated. For the large data set, they ended up being Mean $\pm$

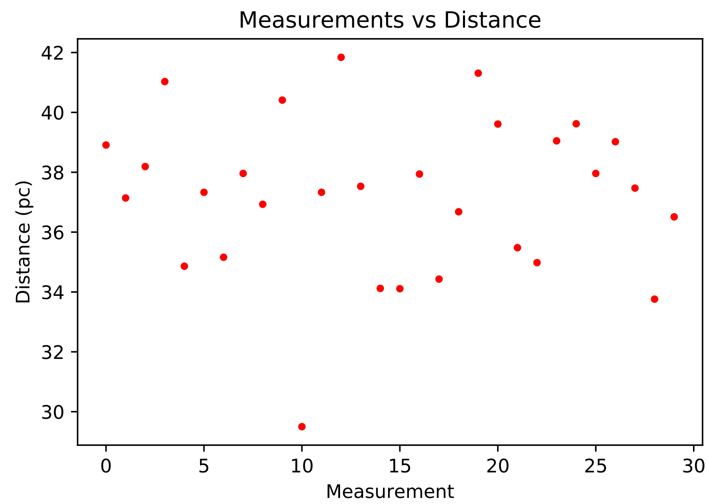

Figure 1: This is the plot from question 1

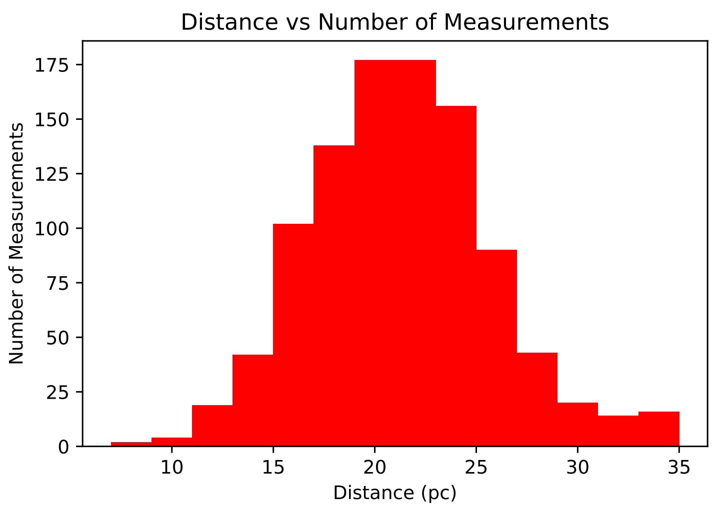

Figure 2: This is the histogram from question 1

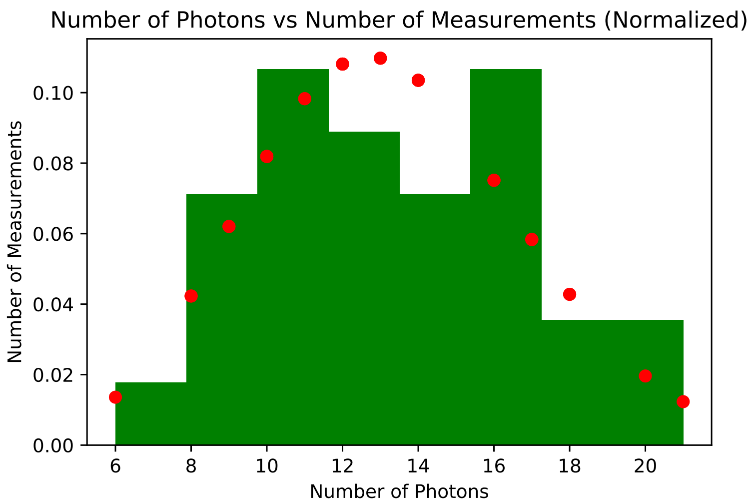

Figure 3: This is the histogram from question 2

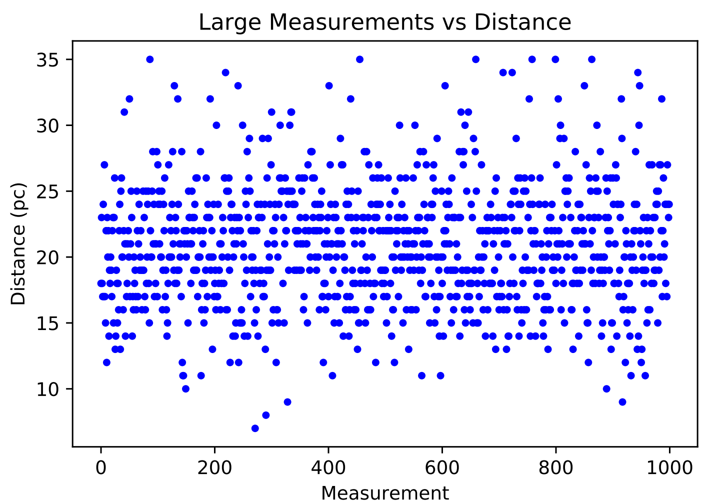

Figure 4: This is the plot of the large data set from question 3

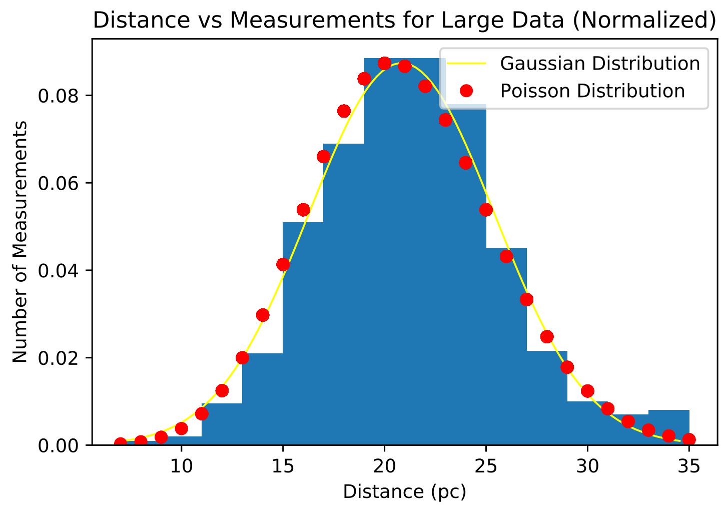

Figure 5: This is the histogram of the large data set from question 3

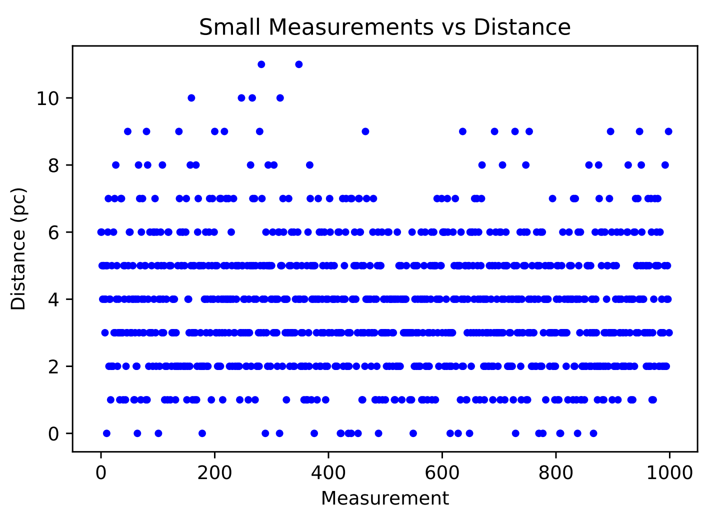

Figure 6: This is the plot of the small data set from question 3

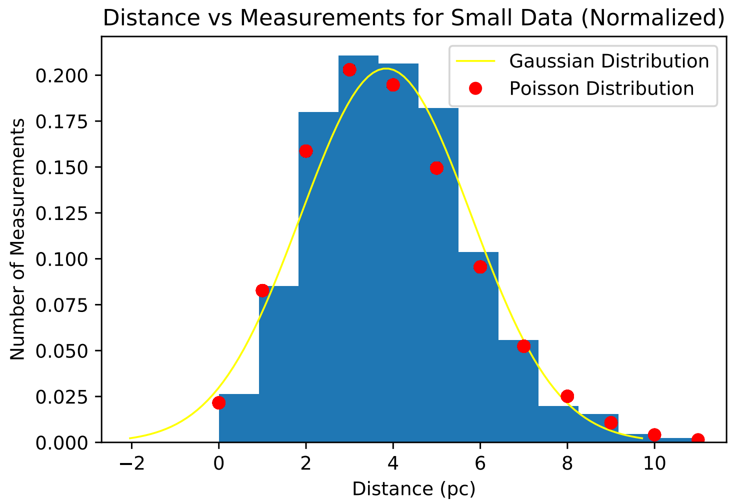

Figure 7: This is the histogram of the small data set from question 3

STDDEV = 20.84 ± 4.51 (pc). For the small data set, it resulted in Mean ± STDDEV = 3.84 ± 1.95 (pc)., Though calculations were done for both of these sets of values, the approximate mean values can be obtained just by looking at the graph, because for both graphs, the maxima of their bars lies on the approximate value of the mean. As we can see in figures 5 and 7, the data looks like it follows both the Poisson and Gaussian distributions very well. However, upon closer inspection the Poisson distribution models the data slightly better than the Gaussian, mostly in estimating where the peak number of measurements is made. This is most apparent in figure 7 where we can see that the peak of the Poisson curve is much closer to the top of the central bar of the distribution than the peak of the Gaussian curve. However, in figure 5, the distributions seem much closer together, thus giving a reasonably similar approximation of the physical distribution. This difference can be seen more drastically when comparing figure 3 to figures 5 and 7. The Poisson distribution in figure 3 very roughly approximates the physical distribution, however it precisely approximates the physical distribution in figures 5 and 7. This is most likely because of the increased number of measurements. As you increase the number of randomized measurements, the Gaussian and Poisson distributions model the physical data better and better.

6 Conclusions

In this experiment, the statistics of photon counting were analyzed through numerous statistical methods such as the Gaussian and Poisson distribution. As seen in the graphs, one can see that the Poisson distribution models the photon counts extremely well. Furthermore, we can see that the Gaussian distribution does a fantastic job of approximating the Poisson distribution. Interestingly enough, we can see that the Poisson and Gaussian distributions look more and more alike as the amount of measurements increases.

7 Appendix A: Mean and Standard Deviation

Below is the code I used for the mean and standard deviation calculations

```
#the number of total measurements was extracted from the data
N = len(distance)

#the mean of each dataset was calculated
mean = (1/N)*(np.sum(distance))
```

```
#the standard deviation was calculated

std = np.sqrt(np.sum((distance - mean)**2)/(N-1))
```

8 Appendix B: Weighted Mean and Standard Deviation

Below is the code I used for the weighted mean and standard deviation calculations

```
#now for the weighted mean

weight  = 1 / (error*2)

weighted_mean = np.sum(weight*distance)/np.sum(weight)

#now for the weighted standard deviation
weighted_std = weighted_mean/np.sum(weight*distance)
```

9 Appendix B: Weighted Mean and Standard Deviation

Below is the code I used for the weighted mean and standard deviation calculations

```
#now for the weighted mean

weight  = 1 / (error*2)

weighted_mean = np.sum(weight*distance)/np.sum(weight)

#now for the weighted standard deviation
weighted_std = weighted_mean/np.sum(weight*distance)
```

10 Appendix C: Plotting the Poisson and Gaussian Distribution

Below is the code I used modelling the Poisson and Gaussian distribution. It should be noted that I had to normalize all bar graphs to get both the Poisson and Gaussian distribution to fit over the data.

```
#the plots of each dataset were shown as a histogram with the gaussian and poisson distributions plotte
plt.hist(data_large, bins = 14, normed = True)

# plot Gaussian distribution
sigma_large = mean_large **(1/2)
x = np.linspace(mean_large-3*sigma_large, mean_large+3*sigma_large, 100)

plt.plot(x, mlab.normpdf(x, mean_large, sigma_large), color=’yellow’, linewidth=’1.0’,
         label=’Gaussian Distribution’)

# plot Poisson distribution
plt.plot(data_large, poisson.pmf(data_large, mean_large), linestyle=’none’, marker=’o’,
         color=’red’, label=’Poisson Distribution’)
sigma = mean_large**(1/2)

plt.title(’Distance vs Measurements for Large Data (Normalized)’)
plt.xlabel(’Distance (pc)’)
```

```
plt.ylabel('Number of Measurements')
plt.legend()
plt.savefig('data_large_hist.pdf')
plt.show()
```