Lab #1 – Basic Photon Statistics

AST326 – October 10, 2018

Ayush Pandhi (ayush.pandhi@mail.utoronto.ca)

Group E: Ayush Pandhi, Hansen Jiang, Keslen Murdock

1 Abstract

CCD data sets of distances and photon count rate are analyzed using Poisson and Gaussian statistics to show that photon noise is well represented as a Poisson distribution of a given average expected rate and that, for a large enough data set, it is similar to a Gaussian distribution. The analysis is performed by over-plotting Poisson distributions for the computed μ parameter and Gaussian distributions for the computed μ and σ of the given data sets. It is also shown that the data sets follow the condition that $\sqrt{Mean} = STDDEV$, which is a predicted outcome of Poisson statistics. In the limit that the independent variable $x \rightarrow \infty$, and as a result μ is large, it is clear that the large data set which is well described by Poisson statistics is also very similar to a Gaussian distribution. However, for the smaller data set a Gaussian distribution proves to be a poor fit in comparison to a Poisson distribution.

2 Introduction

Photon based observations are the most direct method by which astronomers can accurately observe celestial bodies and discern vital information such as position and proper motion of objects. Most modern photon based astronomical observation devices rely heavily on the Photo-electric Effect. This effect, as theorized by Einstein, describes the emission of electrons from the surface of an object in response to an incident light source. Photons themselves carry a characteristic energy which depends on the wavelength (or frequency) of the incident light source. When the photon interacts with the outermost electrons of a material, the energy from the photon is absorbed into the electron. Given that this increase in energy is sufficient to crest the threshold energy of the material's work function, the electron is emitted from the object. A key distinction to make is that increasing the intensity of the incident light does not affect whether or not an electron will be emitted or the energy that it carries because this effect solely depends on the interaction between a singular photon transferring energy to an electron. Charge Coupled Device (CCD) sensors use this effect to generate charge along an array of pixels which is proportional to the number of incoming photons at each pixel. A capacitor then reads the charge along every row of the sensor and the resulting voltage is amplified by both an internal and external amplifier and passed through an analog-to-digital converter to convey the data on a computer in analog to digital units (ADU).

The following experiment uses predetermined data sets from CCD observation of distances and photon count rate that is provided directly by the administration. Theoretically photon noise can be described reasonably well by Poisson statistics at short wavelengths. Thus, the data is analyzed through the use of Poisson statistics to measure if it can be described well by a Poisson distribution for a given average expected rate. Furthermore, the data is also analyzed as a Gaussian distribution for the limit in which the independent variable of the data set becomes large.

3 Observation and Data

All data sets for this experiment were provided digitally through the Astronomy and Astrophysics Department at the University of Toronto[1]. The first set of data (available on September 10th, 2018) describes a series of thirty astronomical observations of a star's distance in parsecs using a CCD with corresponding measurement error for each of the thirty trials. The uncertainty of this data is taken to be the standard deviation or weighted standard deviation as is described in more detail in the next section (Section 4).

The second data set (available on September 10th, 2018) is another series of thirty observation of the photon count rate from an astronomical source for which the expected average count per second is 12 counts/sec. Again, the uncertainty of the set is taken to be its standard deviation (Section 4).

The final two data sets (available on September 24th, 2018) are sets of a thousand observations of the photon count rate of an astronomical source organized into large and small data sets based on the length of the integration time. The standard deviation of each set of observations is computed to be the uncertainty in that data set (Section 4).

4 Data Reduction and Methods

The primary focus of this experiment was to conduct statistical analysis of the provided data sets. The mean, μ_x, and standard deviation, σ_x, of a given parameter x were computed as:

$$\mu_x = \frac{\sum_i x_i}{N} \quad and \quad \sigma_x = \sqrt{\frac{\sum_i (x_i - \mu_x)^2}{N}} \tag{1}$$

Where x_i is the i^{th} value of the parameter and N is the total number of data points in the set. The standard deviation represents the systematic error of the system; consistent error that is innate to the experiment as opposed to random error. Furthermore, given the individual measurement errors for distance within the first data set, the measurements with smaller error should theoretically be more reliable than those with larger error. Thus, it is more accurate to use weighted mean and standard deviation when dealing with this particular set of data. the weight of the i^{th} data point, w_i, is defined as:

$$w_i = \frac{1}{e_i} \tag{2}$$

Where e_i is the measurement error corresponding to the i^{th} data point. Therefore, the weighted mean and standard deviation can be written as:

$$\mu_{weighted,x} = \frac{\sum_i w_i x_i}{\sum_i w_i} \quad and \quad \sigma_{weighted,x} = \sqrt{\frac{1}{\sum_i w_i}} \tag{3}$$

Importantly, $\mu_{weighted,x}$ is simplified to μ_x when the weights on each i^{th} component are equal and sum to 1. The histograms produced from the data sets were analyzed with Poisson and Gaussian statistics. Firstly, Poisson statistics is particularly useful for describing the results of an experiment in which one counts events that occur at random but at an expected average rate. This applies to many astronomical observations such as counting the disintegrations of radioactive nuclei, arrival of cosmic ray particles, and photons hitting a CCD. The probability distribution given by Poisson statistics is described by:

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

Which depicts the probability of having x events when the mean of that parameter is μ. It should be noted that $P(x, \mu)$ is a discrete function that exists for $x > 0$ and $x \epsilon \mathbb{Z}$. The standard deviation of a Poisson distribution is $\sqrt{\mu}$; bright objects with a large mean therefore also have a large standard deviation. Hence, the signal-to-noise for an observation that follows this distribution depends on $\sqrt{\mu}$. For large numbers in the independent variable (i.e. in the limit that $x \rightarrow \infty$ because then μ is large), the Poisson distribution becomes similar to a Gaussian distribution.

Gaussian statistics is often used to describe the distribution of random measurements. The probability distribution of a Gaussian is normalized to 1 and centered at the mean value of the parameter, μ, and follows:

$$G(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \tag{5}$$

It is common practice for a Gaussian to describe the width of the distribution in terms of σ such that: 1σ corresponds to the width in which the probability of finding any given x is approximately 68%, 2σ corresponds to roughly 95%, 3σ to about 99% and so on. Therefore, a small σ leads to precise measurements as there is a high probability of measured values being near the mean of the parameter.

5 Data Analysis and Modeling

The detailed python code for generating all figures and data analysis in this section can be found in the Appendix. Using equation (1) the mean and standard deviation of the first data set was computed to be:

$$Mean \pm STDDEV = 37.21 \pm 2.61\ (pc) \tag{6}$$

However as mentioned previously, this data set contains individual measurement errors corresponding to each distance measurement and thus it is more reasonable to apply the weighted mean and standard deviation formula from equation (3). This leads to a more accurate result with a drastically lower uncertainty:

$$(Weighted)\ Mean \pm STDDEV = 37.50 \pm 0.02\ (pc) \tag{7}$$

Additionally, visualizing this data set as a histogram is more effective than a scatter plot which tends to be convoluted and difficult to decipher (Figure 1). The histogram is not normalized to 1 to maintain physical accuracy of the data as the dependent axis is the number of measurements for any given distance bin.

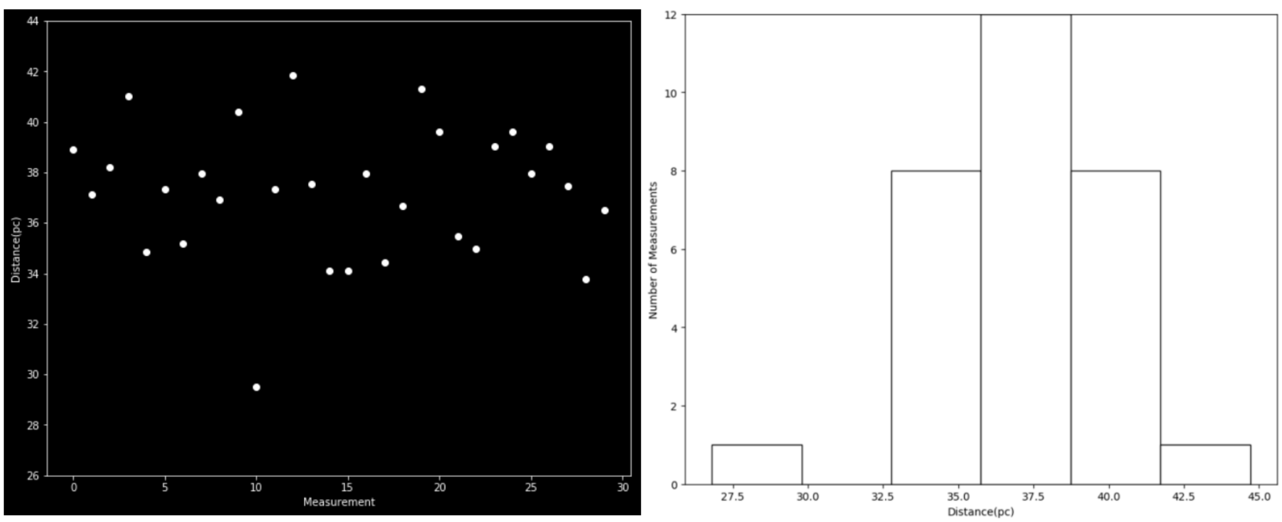

Figure 1: *a) (Left) Scatter plot of 30 obtained distances; b) (Right) Histogram of 30 obtained distances (bin-width = 3).*

Performing an analogous statistical analysis of the mean and standard deviation of the second data set, as per equation (1), leads to the result:

$$Mean \pm STDDEV = 13.2 \pm 3.8\ (counts/second) \tag{8}$$

However, this data set is dealing with the photon count rate per second for 30 consecutive seconds and does not have corresponding measurement errors for each measurement; therefore a weighted analysis of mean and standard deviation cannot be performed. To visualize how well this data set is represented by Poisson statistics a Poisson distribution for $\mu = 12$, which is the expected mean count rate per second from the source for the second data set, is over-plotted with its histogram (Figure 2).

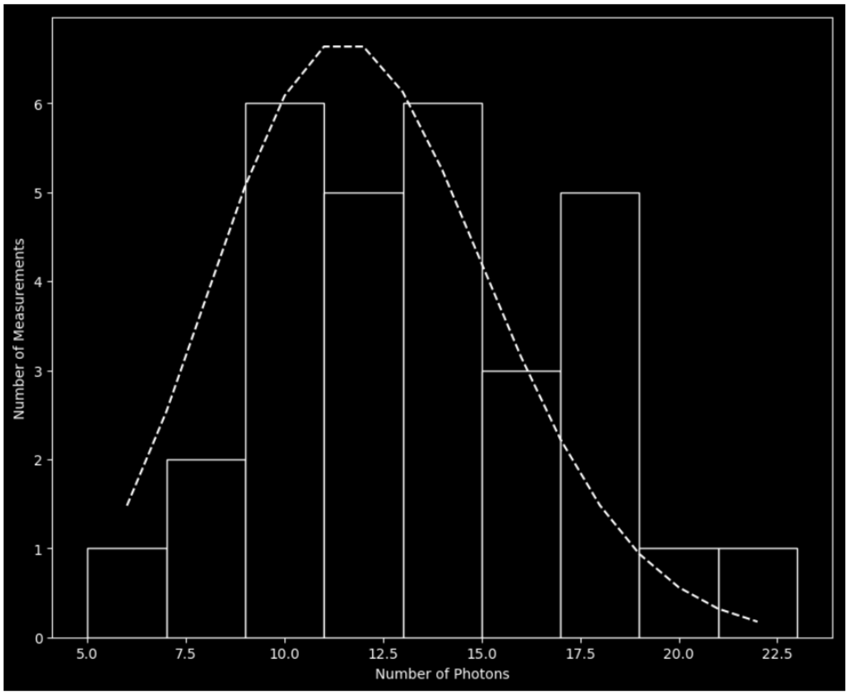

Figure 2: *Histogram of 30 photon count rate measurements (bin-width = 2); (Dashed Line) Expected Poisson distribution of* $\mu = 12$.

Each point in the Poisson distribution here is normalized to match the histogram using:

$$P(x_i, \mu)_{normalized} = C \frac{P(x_i, \mu)}{\sum_i P(x_i, \mu)} (x_{i+1} - x_i) \tag{9}$$

$$C = Total\ number\ of\ measurements(number\ of\ intervals\ in\ x/number\ of\ bins) \tag{10}$$

Specifically for Figure 3, $C = 30(17/9)$. This normalizes the given distribution to the total number of measurements and the same technique is used for following plots in this section. For more specifics on the normalization of the distribution in this section, refer to the python code found in the Appendix.

The next set of data files each contained 1000 measurements of photon count rate and a similar analysis from previous results was performed on it. It was found that the mean and standard deviation of each data set is:

$$Large\ Data\ Set: Mean\ \pm STDDEV = 22.1\ \pm 4.6\ (counts/second) \tag{11}$$

$$Small\ Data\ Set: Mean\ \pm STDDEV = 3.8\ \pm 1.9\ (counts/second) \tag{12}$$

When plotting each data set as a scatter plot similar to Figure 1a, the data becomes extremely disorganized and difficult to read (Figure 3).

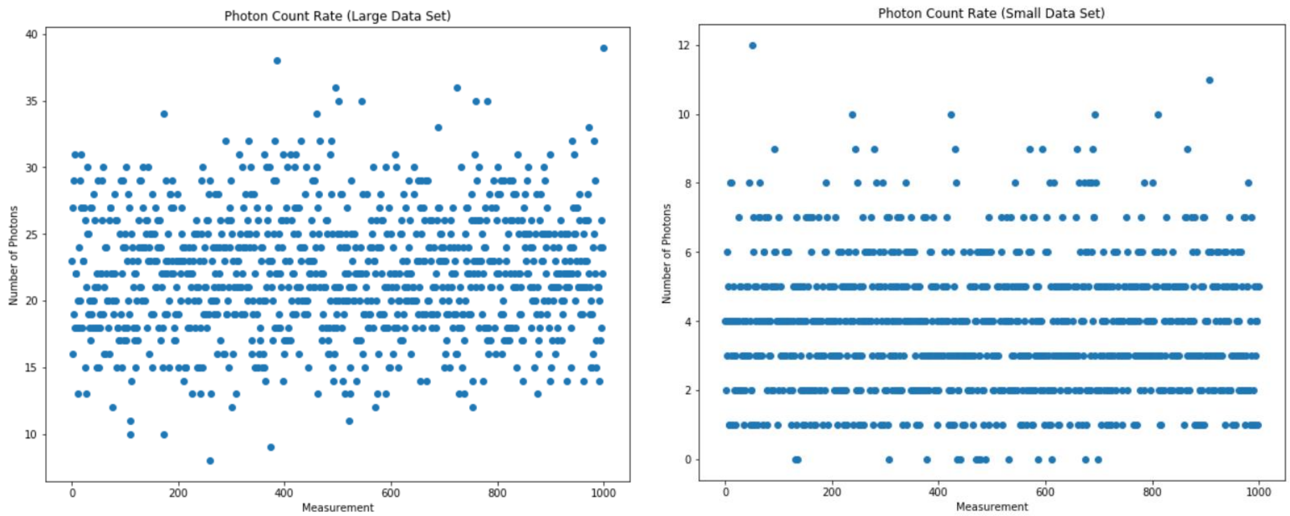

Figure 3: *a) (Left) Scatter plot of 1000 photon count rate measurements of the large data set; b) (Right) Scatter plot of 1000 photon count rate measurements of the small data set.*

Organizing the data into a histogram instead provides a much clearer visualization of the distribution (Figure 4) and the mean and standard deviation can be roughly estimated by visual inspection to be: 22.0 ± 4.0 (counts/second) and 4.0 ± 2.0 (count/second) for the large and small data sets respectively. These values are crude estimations from visual inspection alone but they roughly agree with the previous results and thus inspire confidence in the computation being correct.

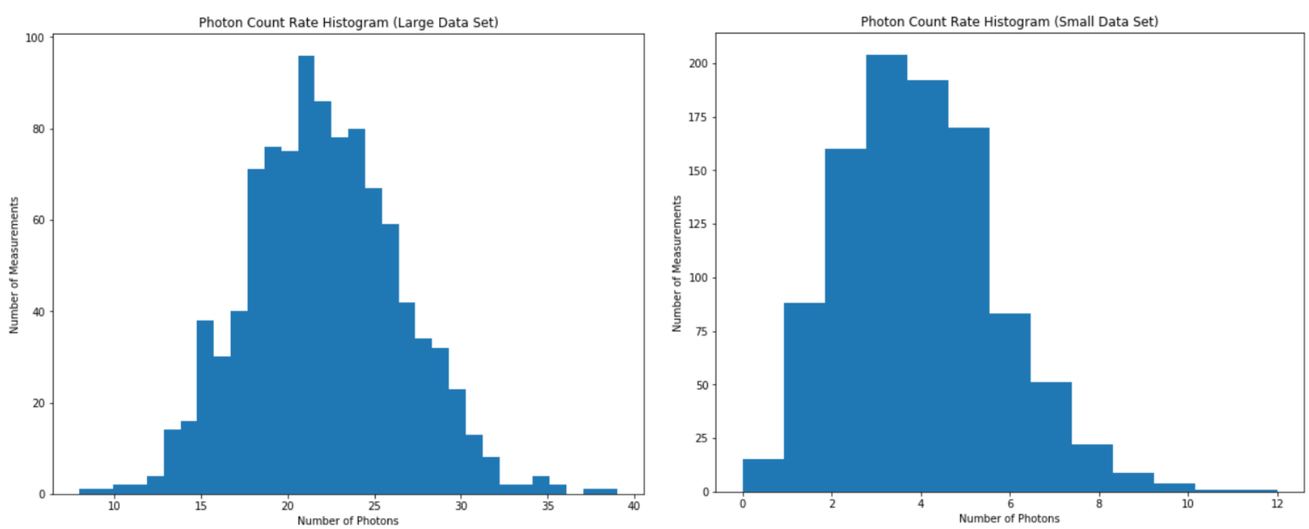

Figure 4: *a) (Left) Histogram of 1000 photon count rate measurements of the large data set (bin-width = 1); b) (Right) Histogram of 1000 photon count rate measurements of the small data set (bin-width = 1).*

Again, the Poisson distribution for the expected photon count rate per second of $\mu_{large} = 22.1$ and $\mu_{small} = 3.8$ respectively is over-plotted against the histogram data for this set (Figure 5).

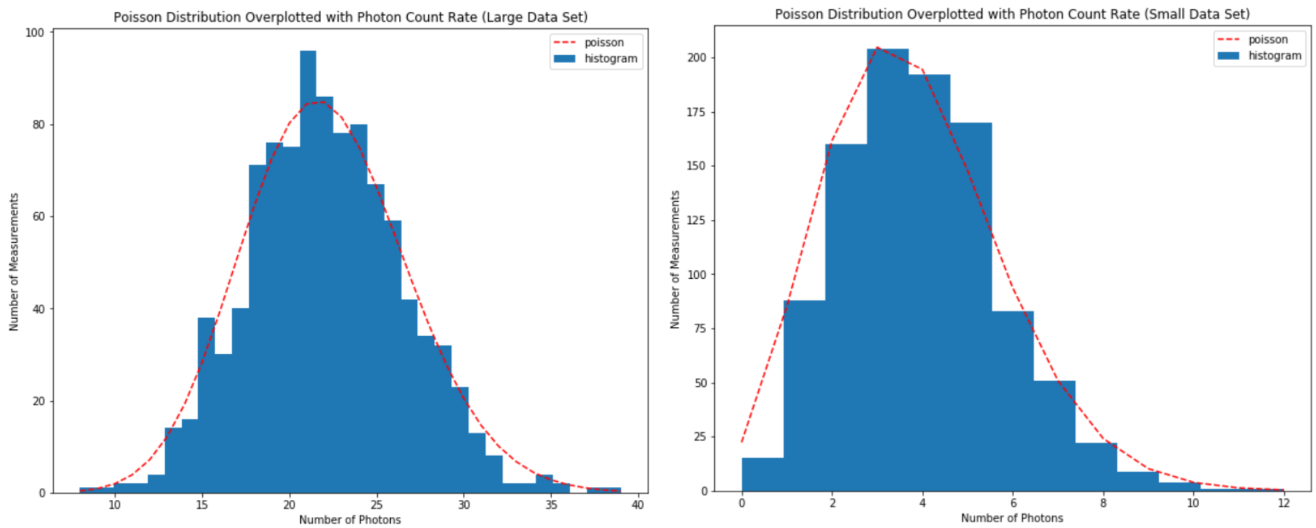

Figure 5: *a) (Left) Identical histogram from Figure 4a with over-plotted Poisson distribution of* $\mu = 22.1$*; b) (Right) Identical histogram from Figure 4b with over-plotted Poisson distribution of* $\mu = 3.8$*.*

To further study this distribution, Gaussian statistics is directly compared to Figure 5 by over-plotting a Gaussian distribution (Figure 6) with the mean and standard deviation parameters set to those that were computed earlier in equations (11) and (12).

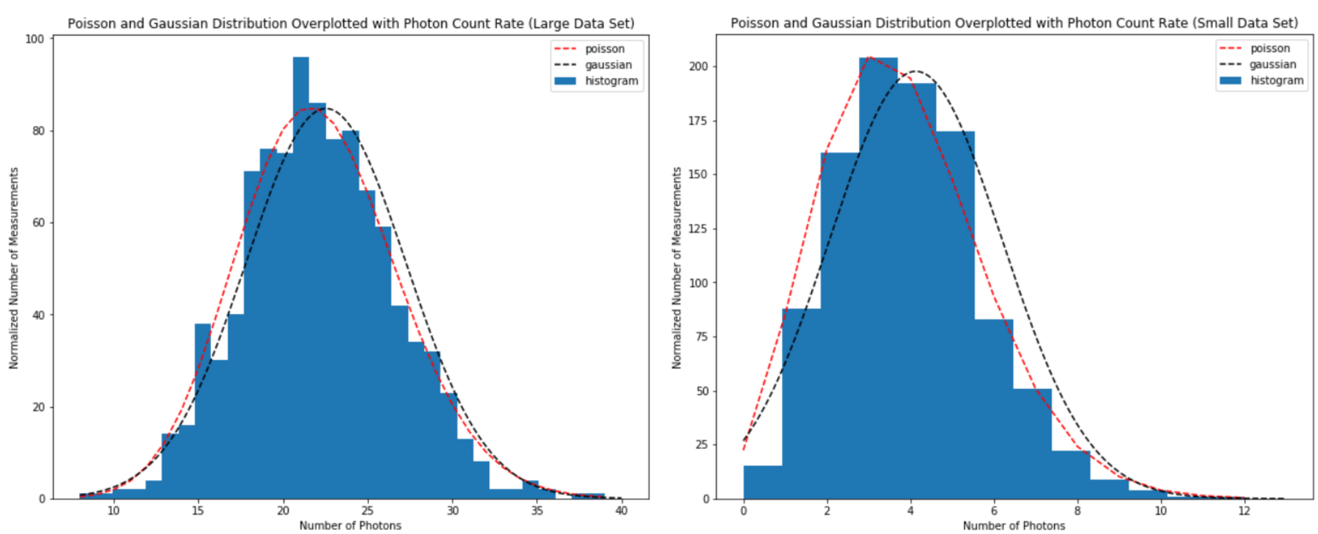

Figure 6: *a) (Left) Identical plot from Figure 5a with additional over-plotted Gaussian distribution of* $\mu = 22.1$ *and* $\sigma = 4.6$*; b) (Right) Identical plot from Figure 5b with additional over-plotted Gaussian distribution of* $\mu = 3.8$ *and* $\sigma = 1.9$*.*

6 Discussion

In Figure 2 it was found that the Poisson distribution for $\mu = 12$ is a somewhat reasonable fit for the second data set, however due to the limited number of measurements (30 total measurements) the histogram of this data is not closely distributed around the mean. The histogram is expected to become closer to the Poisson distribution for a much larger total number of measurements as this would smooth out outliers and create a more uniform distribution based on the Central Limit Theorem. As mentioned earlier in section 4, Poisson statistics can describe photon noise observations reasonably well for lower wavelengths. Additionally, for data that follows this distribution it was found that the mean is related to the square of standard deviation. Attempting to verify this condition in the second data set results in:

$$\sqrt{Mean} = \sqrt{13.2} \cong 3.6 \pm 0.5 \tag{13}$$

Thus, bolstering the notion that the second data set is well represented by Poisson statistics for $\mu = 12$ since the computed standard deviation from equation (8) is within the uncertainty margin of this value.

A similar analysis is performed on the next two sets of data with a comparison to Poisson statistics for $\mu_{large} = 22.1$ and $\mu_{small} = 3.8$ for each respective data set. From visual inspection it seems that both distributions reasonably follow a Poisson distribution (Figure 5). Additionally, to verify the relationship between mean and standard deviation that is true for Poisson statistics once again for this data set:

$$Large\ Data\ Set: \sqrt{Mean} = \sqrt{22.1} \cong 4.7 \pm 0.5 \tag{14}$$

$$Small\ Data\ Set: \sqrt{Mean} = \sqrt{3.8} \cong 1.9 \pm 0.5 \tag{15}$$

These values are resoundingly similar to the computed results from equation (11) and (12); thus it can be concluded that indeed both large and small data sets follow a Poisson distribution for $\mu_{large} = 22.1$ and $\mu_{small} = 3.8$ respectively. In addition, it was found that a Poisson distribution, in theory, is similar to a Gaussian distribution when the independent variable, x, is large (i.e. in the limit that $x \rightarrow \infty$ because then μ is large). From Figure 6, the large data set is clearly well represented by a Gaussian distribution, which is very similar to the Poisson distribution as was predicted. For the small data set, Gaussian statistics is not a good approximation of the data which more resembles Poisson statistics instead. Again as expected, the Gaussian and Poisson distributions do not agree very well in the case that the independent variable, x, is small.

The uncertainty in the experiment primarily arises from systemic error in each data set which is computed to be the standard deviation of the set. In the first data set it is possible to compute the weighted standard deviation because corresponding errors were provided for each measurement. The weighted method provides a low standard deviation and thus a very precise distribution. However for the remaining data sets, weights are not considered and regular standard deviation provides the uncertainty in the data set. Furthermore, when verifying if $\sqrt{Mean} = STDDEV$ for any given set of data, the error is propagated using the rule for polynomial functions:

$$E_{\sqrt{mean}} = \left|\sqrt{Mean}\right| \left|\frac{1}{2}\right| \frac{E_{mean}}{Mean} \tag{16}$$

Where $E_{\sqrt{mean}}$ is the uncertainty of the resulting $\sqrt{Mean}$ value and E_{mean} is the uncertainty (standard deviation) computed for that particular data set.

7 Conclusions

Poisson statistics serves as a good approximation for describing photon based observations such as CCD observations since this type of data deals with counts of events that occur at random but at an expected average rate. Computed means and standard deviations of respective data sets further bolster this conclusion as they follow the relationship that the mean depends on the square of standard deviation, which is an attribute of Poisson distributions. Furthermore, a Poisson distribution for which the independent variable, x, is large (and thus μ is large) becomes very similar to a Gaussian distribution. This discrepancy is seen when comparing the large and small data sets, which are both well represented by Poisson statistics, to a Gaussian distribution based on their computed mean and standard deviation. Here the large data set closely resembles the Gaussian while the small data set does not.

8 References

[1] Department of Astronomy and Astrophysics, University of Toronto, Toronto, ON. "Astronomy 325/326 Lab 1: Basic Photon Statistics", October 2018. Retrieved from http://www.astro.utoronto.ca/~astrolab/.

9 Appendix

9.1 Python Code for Figures 1 - 2

```
#AST326 Lab 1 Part 1 & 2
#Ayush Pandhi (1003227457)
#October 10, 2018

#Importing required modules
import numpy as np
```

```
import matplotlib.pyplot as plt
import math as m

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

#Defining required functions
def mean(data):
  n = len(data)
  mu = (sum(data)/n)
  return mu

def stddev(data):
  n = len(data)
  mu = mean(data)
  numerator = []
  for x in data:
    numerator.append((x - mu)**2)
  sigma = (sum(numerator)/n)**(1/2)
  return sigma

def wgtmean(data, errors):
  w = []
  for i in range(len(data)):
    w.append(1/(errors[i])**2)
  num = []
  for i in range(len(data)):
    num.append(data[i]*w[i])
  wgtmean = sum(num)/sum(w)
  return wgtmean

def wgtstddev(data, errors):
  w = []
  for i in range(len(data)):
```

```
        w.append(1/(errors[i])**2)
    wgtstddev = (1/sum(w))**0.5
    return wgtstddev

def poisson(min, max, mu):
    prob = [(((mu**x)/m.factorial(x))*(2.71828**(-mu))) for x in range(min, max)]
    return prob

def gaussian(data, mu, sigma):
    prob = [(1/sigma*(2*m.pi)**0.5)**((-0.5)*((x - mu)/sigma)**2) for x in data]
    return prob

#Mean and standard deviation for the data
mean_distances = mean(distances)
stddev_distances = stddev(distances)
print ('Mean and standard deviation of data set 1: ', mean_distances, ' +- ', stddev_distances)

#Weighted mean and standard deviation for the data
mean_wgtdistances = wgtmean(distances, stddevs)
stddev_wgtdistances = wgtstddev(distances, stddevs)
print ('Weighted mean and standard deviation of data set 1: ', mean_wgtdistances, ' +- ', stddev_wgtdistances)

#Recreating figure 3 from the handout
plt.style.use('dark_background')
plt.figure(figsize = (10,8))
plt.plot(distances, 'wo')
plt.xlabel('Measurement')
plt.ylabel('Distance(pc)')
plt.ylim(26, 44)
plt.show()

#Recreating figure 4 from the handout
plt.style.use('default')
plt.figure(figsize=(10,8))
plt.hist(distances, bins = 6, range = (26.8, 44.7), edgecolor='k', color='w')
plt.xlabel('Distance(pc)')
plt.ylabel('Number of Measurements')
```

```
plt.ylim(0,12)
plt.show()

#Photon count rate data
pcr = [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]

#Mean and standard deviation of pcr data
mean_pcr = mean(pcr)
stddev_pcr = stddev(pcr)
print ('Mean and standard deviation of data set 2: ', mean_pcr, ' +- ', stddev_pcr)

#Recreating figure 5 from the handout
xlin = np.arange(6, 23)
output = poisson(6, 23, 12)
plt.style.use('dark_background')
plt.figure(figsize=(10,8))
plt.hist(pcr, bins = 9, range = (5, 23), normed=0, edgecolor='w', color='k')
plt.xlabel('Number of Photons')
plt.ylabel('Number of Measurements')
plt.plot(xlin, 30*((23-6)/9)*(output/np.sum(output)/(xlin[1] - xlin[0])), 'w--')
plt.show()
```

9.2 Python Code for Figures 3 - 6

```
#AST326 Lab 1 - Part 3
#Ayush Pandhi (1003227457)
#October 10, 2018

#Importing required modules
import numpy as np
import matplotlib.pyplot as plt
import math as m

#Defining required functions
def mean(data):
    n = len(data)
    mu = (sum(data)/n)
```

```
    return mu

def stddev(data):
    n = len(data)
    mu = mean(data)
    numerator = []
    for x in data:
        numerator.append((x - mu)**2)
    sigma = (sum(numerator)/n)**(1/2)
    return sigma

def wgtmean(data, errors):
    w = []
    for i in range(len(data)):
        w.append(1/(errors[i])**2)
    num = []
    for i in range(len(data)):
        num.append(data[i]*w[i])
    wgtmean = sum(num)/sum(w)
    return wgtmean

def wgtstddev(data, errors):
    w = []
    for i in range(len(data)):
        w.append(1/(errors[i])**2)
    wgtstddev = (1/sum(w))**0.5
    return wgtstddev

def poisson(min, max, mu):
    prob = [(((mu**x)/m.factorial(x))*(2.71828**(-mu))) for x in range(min, max)]
    return prob

def gaussian(data, mu, sigma):
    xdata = np.linspace(min(data), max(data), 1000)
    prob2 = [(1/(2*m.pi)*sigma**2)*(2.71828**(-(x - mu)**2/(2*sigma**2))) for x in xdata]
    return prob2

#Loading Large and Small data files
```

```
largedata = np.loadtxt('largedata.txt')
smalldata = np.loadtxt('smalldata.txt')

#Calculating and printing mean and standard deviation of each data set
#Getting weighted mean and stddev is not possible as there are no errors in this data set
mean_large = mean(largedata)
stddev_large = stddev(largedata)
mean_small = mean(smalldata)
stddev_small = stddev(smalldata)
print ('Mean and Standard Deviation of Large Data: ', mean_large, ' +- ', stddev_large)
print ('Mean and Standard Deviation of Small Data: ', mean_small, ' +- ', stddev_small)

#If we account for significant figures we get the the following instead
print ('Mean and Standard Deviation of Large Data (with significant figures): 22.1 +- 4.6')
print ('Mean and Standard Deviation of Small Data (with significant figures): 3.8 +- 1.9')

#Scatter plot for large data set
plt.figure(figsize = (10, 8))
plt.plot(largedata, 'o')
plt.title('Photon Count Rate (Large Data Set)')
plt.xlabel('Measurement')
plt.ylabel('Number of Photons')
plt.show()

#Scatter plot for small data set
plt.figure(figsize = (10, 8))
plt.plot(smalldata, 'o')
plt.title('Photon Count Rate (Small Data Set)')
plt.xlabel('Measurement')
plt.ylabel('Number of Photons')
plt.show()

#Histogram for large data set
plt.figure(figsize = (10, 8))
plt.hist(largedata, bins = 32)
plt.title('Photon Count Rate Histogram (Large Data Set)')
plt.xlabel('Number of Photons')
plt.ylabel('Number of Measurements')
```

```
plt.show()

#Hitogram for small data set
plt.figure(figsize = (10, 8))
plt.hist(smalldata, bins = 13)
plt.title('Photon Count Rate Histogram (Small Data Set)')
plt.xlabel('Number of Photons')
plt.ylabel('Number of Measurements')
plt.show()

#Overplotting the poisson_output distriution onto the large data histogram (mu = 22)
xlin = np.arange(8, 40)
poisson_output = poisson(8, 40, 22)
gaussian_output = gaussian(xlin, mean_large, stddev_large)
plt.figure(figsize = (10, 8))
plt.hist(largedata, bins = xlin.shape[0], normed=0, label='histogram')
plt.plot(xlin, 1000*(poisson_output/np.sum(poisson_output)/(xlin[1] - xlin[0])), 'r--', label='poisson')
plt.title('Poisson Distribution Overplotted with Photon Count Rate (Large Data Set)')
plt.xlabel('Number of Photons')
plt.ylabel('Number of Measurements')
plt.legend()
plt.show()

#Overplotting the poisson_output distriution onto the small data histogram (mu = 4)
xlin2 = np.arange(0, 13)
poisson_output2 = poisson(0, 13, 4)
gaussian_output2 = gaussian(xlin2, mean_small, stddev_small)
plt.figure(figsize = (10, 8))
plt.hist(smalldata, bins = xlin2.shape[0], normed=0, label='histogram')
plt.plot(xlin2, 1000*(poisson_output2/np.sum(poisson_output2)/(xlin2[1] - xlin2[0])), 'r--',
label='poisson')
plt.title('Poisson Distribution Overplotted with Photon Count Rate (Small Data Set)')
plt.xlabel('Number of Photons')
plt.ylabel('Number of Measurements')
plt.legend()
plt.show()

#Overplotting the poisson_output distriution onto the large data histogram (mu = 22)
```

```
xlin = np.arange(8, 40)
poisson_output = poisson(8, 40, 22)
gaussian_output = gaussian(xlin, mean_large, stddev_large)
plt.figure(figsize = (10, 8))
plt.hist(largedata, bins = xlin.shape[0], normed=0, label='histogram')
plt.plot(xlin, 1000*(poisson_output/np.sum(poisson_output)/(xlin[1] - xlin[0])), 'r--', label='poisson')
plt.plot(np.linspace(8, 40, 1000), 1000*(gaussian_output/np.sum(gaussian_output)/(np.linspace(8, 40, 1000)[1] - np.linspace(8, 40, 1000)[0])), 'k--', label='gaussian')
plt.title('Poisson and Gaussian Distribution Overplotted with Photon Count Rate (Large Data Set)')
plt.xlabel('Number of Photons')
plt.ylabel('Normalized Number of Measurements')
plt.legend()
plt.show()

#Overplotting the poisson_output distriution onto the small data histogram (mu = 4)
xlin2 = np.arange(0, 13)
poisson_output2 = poisson(0, 13, 4)
gaussian_output2 = gaussian(xlin2, mean_small, stddev_small)
plt.figure(figsize = (10, 8))
plt.hist(smalldata, bins = xlin2.shape[0], normed=0, label='histogram')
plt.plot(xlin2, 1000*(poisson_output2/np.sum(poisson_output2)/(xlin2[1] - xlin2[0])), 'r--', label='poisson')
plt.plot(np.linspace(0, 13, 1000), 1000*(gaussian_output2/np.sum(gaussian_output2)/(np.linspace(0, 13, 1000)[1] - np.linspace(0, 13, 1000)[0])), 'k--', label='gaussian')
plt.title('Poisson and Gaussian Distribution Overplotted with Photon Count Rate (Small Data Set)')
plt.xlabel('Number of Photons')
plt.ylabel('Normalized Number of Measurements')
plt.legend()
plt.show()
```