# AST325 Lab 1

Chris Compierchio

October 4 2021

## 1 Introduction

The purpose of this lab was to get a feel for using various statistical analysis tools. Histograms, Poisson distributions and Gaussian plots were used to plot given data on measured photon count rates from a star. Data was collected using both small and large count rates.

## 2 Results

For practice, a histogram using data from the lab document was recreated. 30 data points were given and the mean and standard deviation of the data set were both calculated in order to fit a Poisson distribution onto the histogram representing the data. The plotted results were as shown:

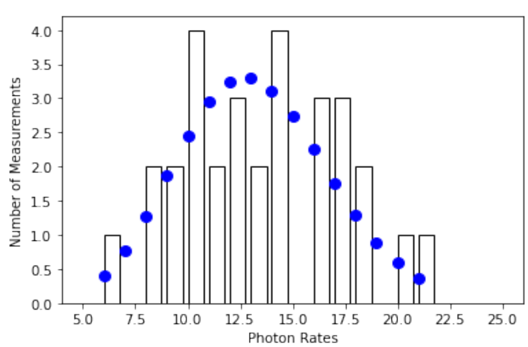
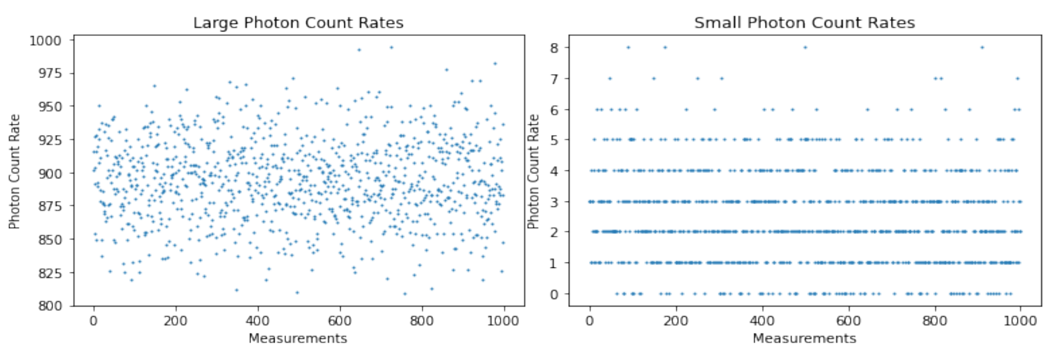



Figure 1: Histogram of given data

Evidently, the Poisson distribution fits well over the histogram. We can interpret the data to see that the expected mean is about 13 counts/second. In other words, it is most probable to see a rate of about 13 counts/second and less probable to see certain values as you move away from 13. This is what the Poisson Distribution suggest, but we can see that the experimental data does not agree. In fact, the experimental does not center around the mean, but instead, it fluctuates as we can see the two highest measurements occurred around 14 and 11 counts/second.

Next, plots were created using the data sets containing small and large integration times. Both data sets were plotted against the number of measurements in scatter plots as shown:

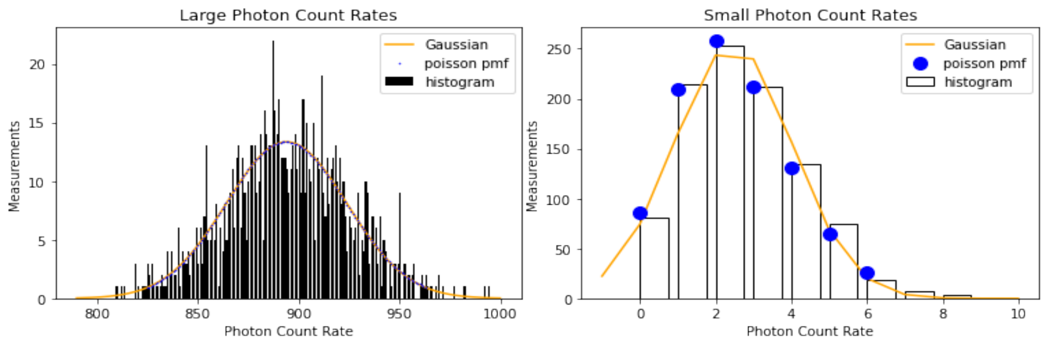



(a) Large Count Rates

(b) Small Count Rates

Figure 2: Scatter Plots for Large and Small Count Rates

By glancing at the plots above, it would be too challenging to try and guess the mean and standard deviation, therefore the data from each scatter plot was used to make histograms. These histograms were overlaid with Poisson and Gaussian distributions in order to compare the expected and collected results as follows:




(a) Large Count Rates

(b) Small Count Rates

Figure 3: Histograms for Large and Small Count Rates

Here, we can estimate a mean of about 875 for the large rates. Using the fact that Poisson distributions follow the rule

$$std = \sqrt{mean} \tag{1}$$

we can calculate an approximate standard deviation (std) of 29.58. This is accurate to the actual values of 894 and 29.9 respectively. Doing the same for the small count rate data, it is evident that the mean is about 2.1 and so the approximate standard deviation would be 1.45. This, again, is similar to the calculated values of 2.46 and 1.57 respectively.

With the large count rates, we have a larger data set and therefore the Poisson and Gaussian distributions are more in line with each other and more accurate to the data set.

# 3 Conclusion

This lab demonstrates the importance of the Poisson distribution. With a large enough data set, one can accurately predict the mean value of the set by plotting a Poisson distribution of the data.

# 4 Appendix A

In [10]:

```
import numpy as np
from scipy.optimize import curve_fit
import matplotlib.pyplot as plt
%matplotlib inline
from scipy.stats import *
import scipy as sp
```

In [11]:

```
vals = np.zeros(30)

vals[0] = 13
vals[1] = 17
vals[2] = 18
vals[3] = 14
vals[4] = 11
vals[5] = 8
vals[6] = 21
vals[7] = 18
vals[8] = 9
vals[9] = 12
vals[10] = 9
vals[11] = 17
vals[12] = 14
vals[13] = 6
vals[14] = 10
vals[15] = 16
vals[16] = 16
vals[17] = 11
vals[18] = 10
vals[19] = 12
vals[20] = 8
vals[21] = 20
vals[22] = 14
vals[23] = 10
vals[24] = 14
vals[25] = 17
vals[26] = 13
vals[27] = 16
vals[28] = 12
vals[29] = 10

mean = np.mean(vals)
std = np.sqrt(mean)
mean, std
```

Out[11]:

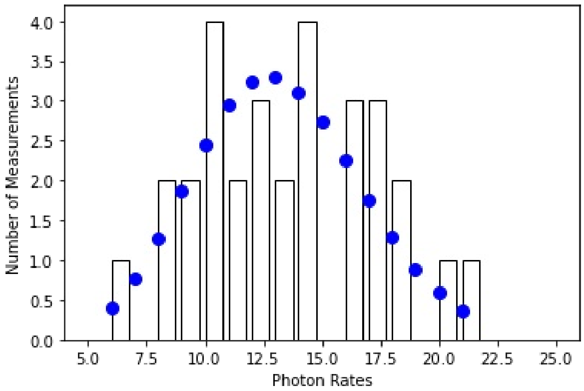
```
(13.2, 3.63318042491699)
```

In [12]:

```
bins = np.round(np.linspace(5, 25, 21))

plt.hist(vals, bins, histtype='bar', align='mid', orientation='vertical', width=.75, density = False, color="whit
e", label = "histogram", edgecolor='black')

z = np.arange(poisson.ppf(0.01, mean),

              poisson.ppf(0.99, mean))

plt.plot(z, 30*poisson.pmf(z, mean), 'bo', ms=8, label='poisson pmf')
plt.xlabel("Photon Rates")
plt.ylabel("Number of Measurements")
```

Out[12]:

```
Text(0, 0.5, 'Number of Measurements')
```




In [13]:

```
large_nums = np.loadtxt('Compierchio-compier1-Large.txt')
small_nums = np.loadtxt('Compierchio-compier1-Small.txt')
```

In [14]:

```
large_mean = np.mean(large_nums)
large_std = np.sqrt(large_mean)

small_mean = np.mean(small_nums)
small_std = np.sqrt(small_mean)

large_mean, large_std, small_mean, small_std
```

Out[14]:

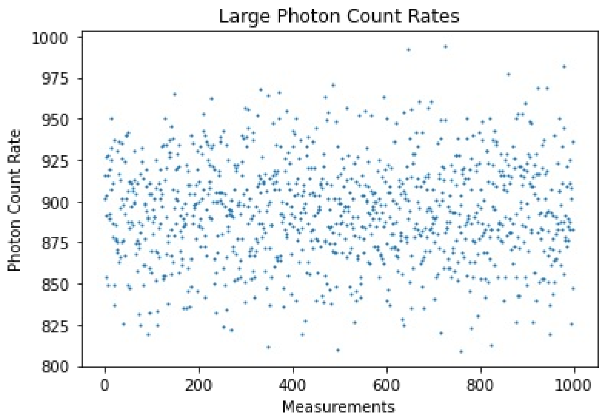
```
(894.082, 29.901203989137294, 2.462, 1.5690761613127644)
```

In [15]:

```
measurements = np.zeros(1000)

for i in range(0, 999):
    measurements[i] = i+1

plt.scatter(measurements, large_nums, s = 1)
plt.xlabel("Measurements")
plt.ylabel("Photon Count Rate")
plt.title("Large Photon Count Rates")
```

Out[15]:

```
Text(0.5, 1.0, 'Large Photon Count Rates')
```

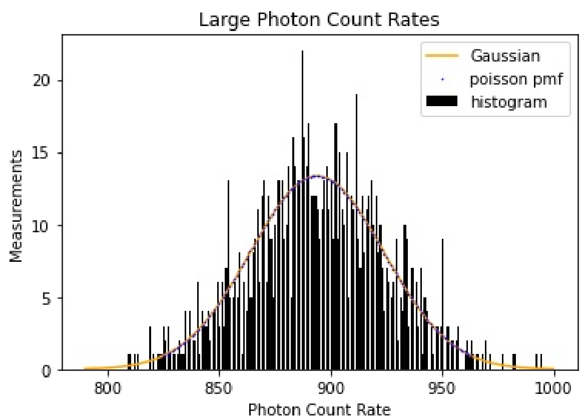



In [16]:

```
bins = np.round(np.linspace(790, 1000, 1000))

histogram_large = plt.hist(large_nums, bins, histtype='bar', align='mid', orientation='vertical', width=.75, colo
r="black", density = False, label = "histogram")

z_large = np.arange(poisson.ppf(0.01, large_mean),

              poisson.ppf(0.99, large_mean))

plt.plot(bins, 1000*sp.stats.norm.pdf(bins, large_mean, large_std), color = "orange", label = 'Gaussian')

plt.plot(z_large, 1000*poisson.pmf(z_large, large_mean), 'bo', ms=0.5, label='poisson pmf')
plt.ylabel("Measurements")
plt.xlabel("Photon Count Rate")
plt.title("Large Photon Count Rates")
plt.legend(loc='best')
```

Out[16]:

```
<matplotlib.legend.Legend at 0x182a19a93d0>
```

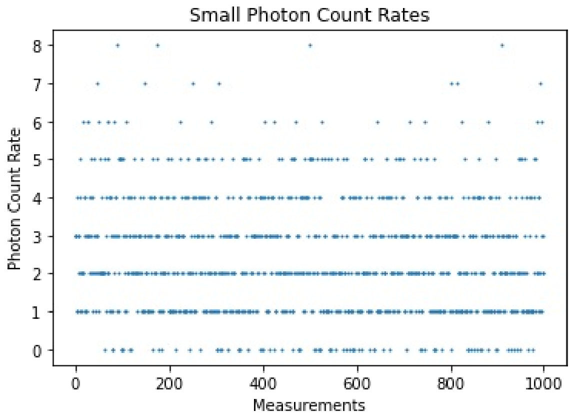


In [17]:

```
plt.scatter(measurements, small_nums, s = 1)
plt.xlabel("Measurements")
plt.ylabel("Photon Count Rate")
plt.title("Small Photon Count Rates")
```

Out[17]:

```
Text(0.5, 1.0, 'Small Photon Count Rates')
```

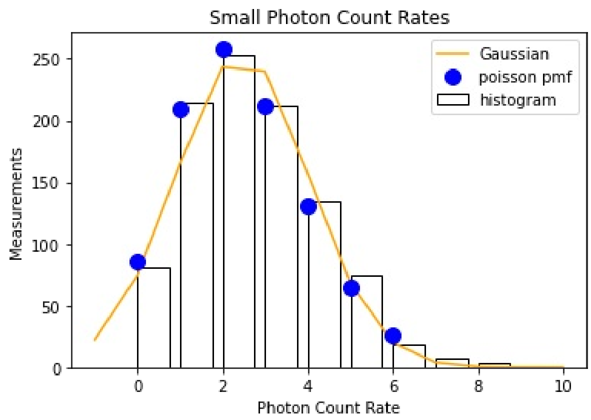



In [18]:

```
bins = np.round(np.linspace(-1, 10, 1000))

histogram_small = plt.hist(small_nums, bins, histtype='bar', align='mid', orientation='vertical', width=.75, color="white", label = "histogram", edgecolor='black')

z_small = np.arange(poisson.ppf(0.01, small_mean),

              poisson.ppf(0.99, small_mean))

plt.plot(bins, 1000*sp.stats.norm.pdf(bins, small_mean, small_std), color = "orange", label = 'Gaussian')

plt.plot(z_small, 1000*poisson.pmf(z_small, small_mean), 'bo', ms=10, label='poisson pmf')
plt.ylabel("Measurements")
plt.xlabel("Photon Count Rate")
plt.title("Small Photon Count Rates")
plt.legend(loc='best')
```

Out[18]:

```
<matplotlib.legend.Legend at 0x182a248a520>
```




In [ ]:

In [ ]: