Lab 1: Photon Counting and the Statistics of Light

Jeffrey Vargas

Armando Montejano, Alex Nguyen, Brian Timar

vargas.jeffrey52@berkeley.edu

I. Abstract

Instruments used in modern astronomy operate under the principals of photon counting. Information collected through photons allows astronomers analyze the statistical properties of light. The detectors are able to count photons per some interval time. In this experiment we analyzed the physical limitation on the detection of light through the collection of 100,000 photons from an LED light source. We found that the mean photons arrival time is 7.1×10^5 clock ticks, where clock ticks are in the scale of Nano seconds. After removing after pulses, we observe that the Standard deviation of the mean interval clock ticks exhibits exponential behavior. Finally, we analyzed how the statistical properties of light vary with brightness.

II. Introduction

During the 19th century it was observed that light incident on certain metals caused an electron to be emitted from its surface. This phenomena is known as the Photoelectric effect and the electrons emitted are called photoelectron. This occurs because photons carry energy in discrete packages, that is E = hf [1]. Photons with the required energy cause an electrons to get ejected from the metal.

In this experiment we analyzed the physical limitation with the detection of light. We used a photo-multiplier tube (PMT), which utilizes photoelectric effect and amplifies the number of emitted photons, to collect photon samples and observed the collective statistical properties of our sample. Because light travels at 3×10^8 meter per second, our detector had to operate at a comparable scale. As result, we used a coincidence processor (CoinPro) to measures the time interval between pulses from the PMT with sub-nanosecond resolution.

We observed that the data collected had statistical characteristics which could be analyzed. Our plots illustrate the statistics of photon counting;such as, the Poisson distribution, the Gaussian limit, histograms, descriptive statistics, mean, standard deviation, error propagation, and the standard error related to our measurements.

III. Observations and Data

I. Gathering Data

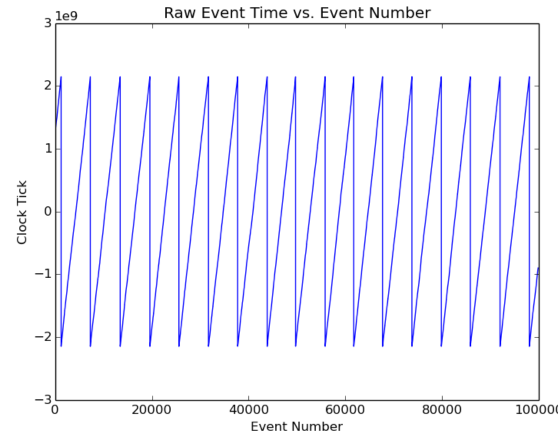
The PMT along with the CoinPro allowed us to measure the detection of photons. The photons we collected within the elapse time we reference it as an event. events occur at the scale of nanosecond.

[1] h represents Planks Constant 6.6260700 $\times 10^{-34} m^2 \times kg \times s^{-1}$ and f represents the frequency.

Fortunately, the CoinPro measures time elapse by measuring time in pulses,which we reference as 'Clock Ticks'. Events are measured every 0.833 ns which represent clock ticks or time elapsed since the previous event. Additionally, 'Clock Ticks' are recorded in 32 bit, hence data runs from -2^{31} to $2^{31}-1$, that is, -2147483648 to 2147483647, before rolling over. Therefore, the clock rolls over approximately $2^{32} \times .833 \times 10^{-9}s \approx 3.58s$. Figure1a gives a sample of the initial data collected while figure1b shows the number of events collected versus 'Clock Ticks'. For our experiment we used a sample size which consisted of 100,000 events, i.e. photons.

2	1230515798
2	1232194204
2	1232419276
2	1233121064
2	1233312955
2	1235227212
2	1237237673
2	1238185988
2	1239034118
2	1239034385

(a) Sample Data



(b) Initial data plotted

Figure 1: The first column in figure1a describes the channel on the CoinPro where the PMT was connected to, the second column describes the elapses time, i.e. 'Clock Ticks', of subsequent event. Figure1b illustrates the sample size and time elapse in a plot. Note that clock thick oscillate according to 32 bit format.

It is important we stress the importance of the type of data collected, 32 bit, because the program we used (numpy[2]) to read the data has a default setting of 16bit. Figure 2 displays what would occur if we did not specify the data type.

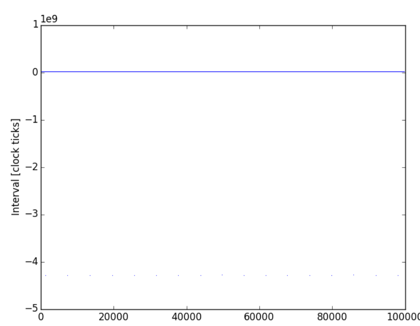
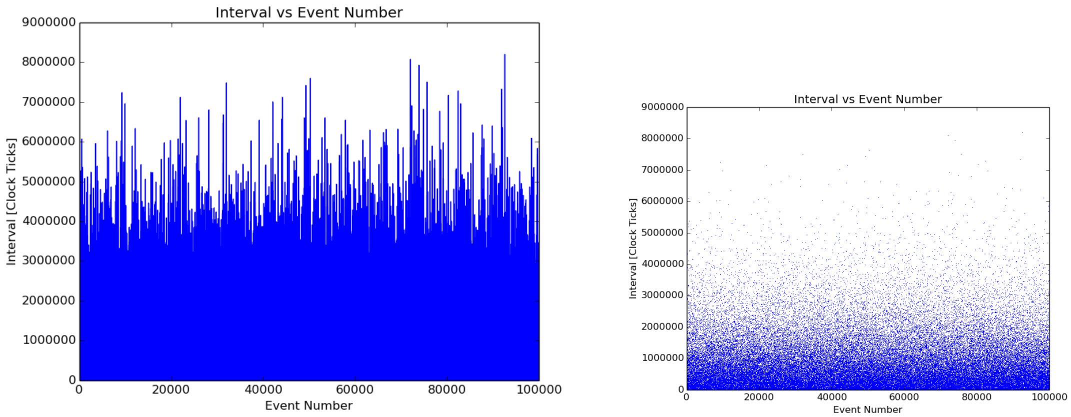


Figure 2: The plot describes the time between each event in the experiment without using the dtype= 'int32', that is without specifying the range of the data type. By default the np.loadtxt command will load numbers as 16 bit floats. Our data values are beyond the range of 16 bits therefore all the numbers are translated as 0

[2] numpy is a module which is imported onto the Python programming language and contain the np.loadtxt command which was used to import our data



(a) Sample Data

(b) Initial data plotted

Figure 3: Figure3a depicts the interval between subsequent events vs event number. The graph allows us to see the time elapse between different events. We observe that the longest interval runs to approximately 8×10^7. Note: Although time elapse is recorded for each event, we are not able to make a clear distinction between events because of the format of the graph.Figure3b was modified to see individual events. we observe that events are clustered around 1×10^6 Interval 'Clock ticks'.

In order to make the data more readily useful we took the time difference between subsequent events and plotted them on Figure3a. From the plot we could observe the time elapse in clock ticks between events; however, events with short interval clock ticks are clustered near the bottom of the graph, we use Figure3b to illustrate individual events.

II. Statistical Mean

By analyzing Figure3 we observe that even though events occur at random[3], there must be some average time between events, thus a cluster of events at short 'Clock Ticks'. We can calculate the average time between events by applying the following formula (see Figure4):

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

where $\bar{x}$ represents the mean time for N events.

In Figure4 we divided the sample size of 100,000 into smaller sample sizes chunks of 1000 events. we observer that for large N the mean starts to approach a true mean,which is displayed in Figure5.

$$\bar{x} = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} x_i \tag{2}$$

Note that our data has a finite number of events ,N. However, since N is large compared to the difference between 'clock ticks' we can make an approximation to the true mean. The plot on Figure5 appears to be approaching 70.7×10^4 Interval [clock ticks].

[3] Random events imply that the detection of photons are independent of one another, that is, detecting one photon does not affect the probability of detecting another.

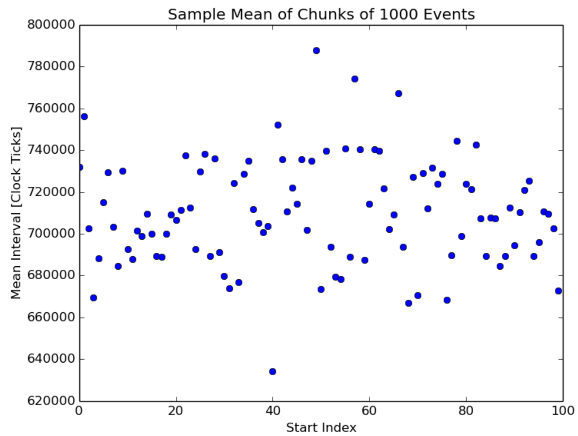


Figure 4: Each point on The plot represents the mean for 1000 events, therefore, the initial 100,000 events was divided into 100 chunks of 1000 events. The plot has a span of 1.5×10^5. The mean of the means interval [Clock Ticks] is 7.1×10^5 ticks and the standard deviation is 2.5×10^4 ticks

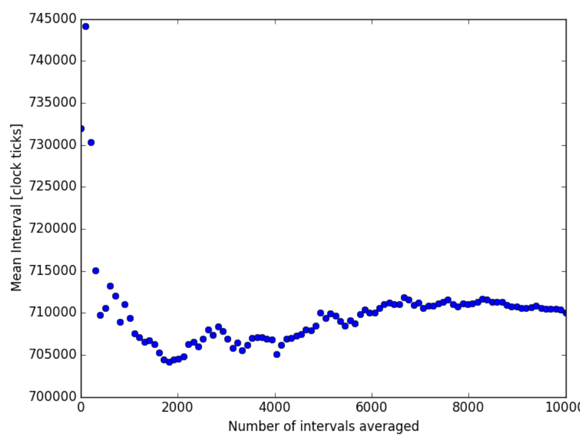


Figure 5: this graph is demonstrating how applying a limit to the mean interval [Clock Ticks] will cause the mean to converge to a true mean. In our experiment the sample size was large compared to time difference between events.

III. Standard Deviation of the Mean

Our sample size or chunks in Figure4 were chosen arbitrarily and therefore the standard deviation of the mean, SDOM, may vary depending on the chunk size. To illustrate this Figure6 was constructed using a chunk size of 100 events, i.e. there is 1000 chunks of 100 events. We notice the that the chunk size causes the SDOM to change to 8.4×10^4 clock ticks. This was the result of an increase in scattering.

In order to investigate how the SDOM varies with sample size we constructed Figure7. The SDOM is plotted as a function of number of events averaged. The graph illustrates how the SDOM drops as N increases. This is because the standard deviation of the mean is equal to $\frac{s}{\sqrt{N}}$ where s is the standard deviation of the sample and N is the chunk size, used to compute the mean. Figure8 demonstrates how the SDOM follows the theoretical expectation.

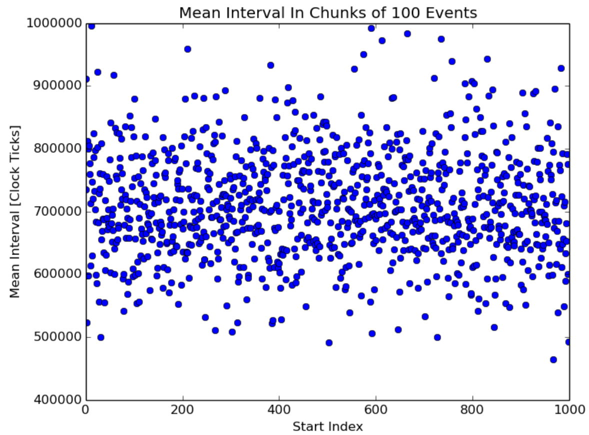


Figure 6: The plot is the mean intervals for 1000 chunks of 100 events taken from Figure 3. The means is calculated by taking the mean of 100 events and then plotting the result until all 10,000 exhausted. Note: that smaller chunk size caused the SDOM and the span to change. The span a the data is about 5×10^6 clock ticks, with a mean of 7.1×10^5 clock ticks and a standard deviation of 8.4×10^4 clock ticks

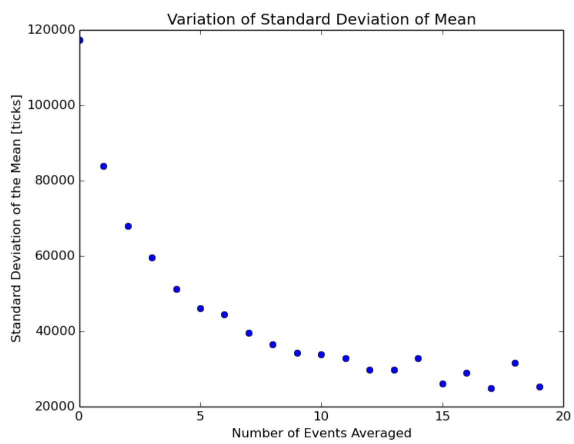


Figure 7: A plot of variation of the standard deviation of the mean with the size of the data chunk averaged. the graph appears to decay by $1/\sqrt{N}$ which is what we expect.

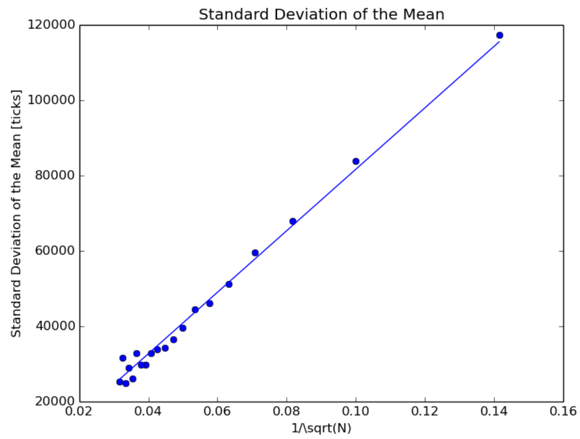


Figure 8: The plot is the standard deviations of the means (SDOM) of our data vs $\frac{1}{\sqrt{N}}$, where N represents the sample size used to calculate the distribution of means. The line drawn is the theoretical expectation with our standard deviation of the mean $\frac{s}{\sqrt{N}}$, where s is the sample standard deviation. we see that the standard deviation of the mean fits the theoretical prediction.

IV. Model discussion

I. Histograms Analysis

The previous section allowed us to calculate the central tendency and the spread of the photons collected, however, in order to extract distribution of the measured values we analyzed the data using a histogram. The histogram shown in Figure9 allows us to view the distribution of detected photons within interval ticks. The key feature of the histogram are: it does not resemble a Gaussian distribution, there is a spike at short interval ticks, and there is a long tail at the long interval ticks.

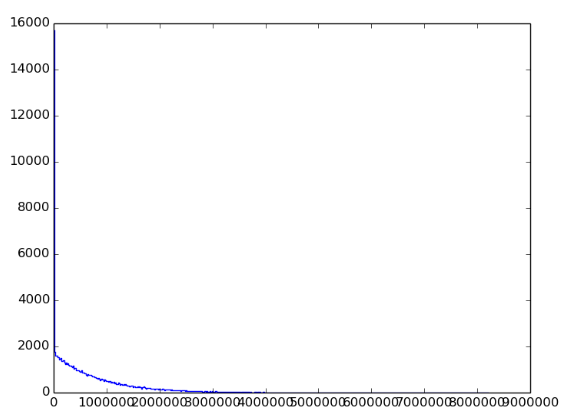


Figure 9: Figure9 is a histogram of the events (collected photons) for a given intervals tick. in Figure 3. The large spike at short interval tick is a result of after pulsing and the decaying feature indicates that photons are less likely to be detected at large interval ticks.

The histogram was magnified on Figure10 in order to investigate the large spike. We observe that the large spike occurs at small interval ticks. This indicates that the spike is a result of after pulse[4] The decaying feature indicates that the probability of detecting a photon at small interval ticks is greater than the probability of detecting it a large interval ticks.

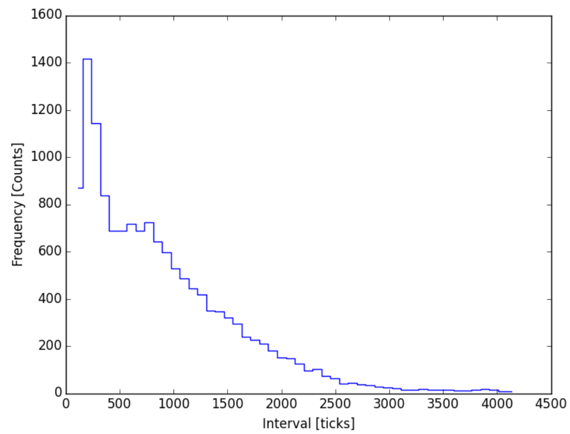
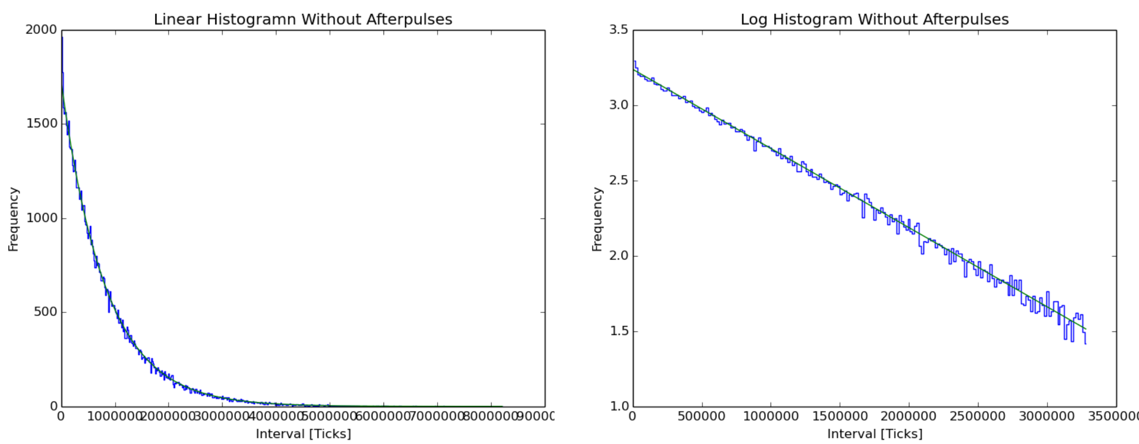


Figure 10: A zoomed in histogram of the large spike in Figure 10. This spike shows the effects of after-pulsing. The after-pulse decays around the 3,000 interval tick, so we remove this intervals.



(a) Sample Data

(b) Initial data plotted

Figure 11: Figure11a is the corrected histogram after after pulsing is removed, plotted in linear scale. Figure11b is the histogram plotted in log scale. The line drawn is the theoretical exponential distribution $p(\Delta t) = \frac{1}{\tau} e^{\left(\frac{-\Delta t}{\tau}\right)}$. Since the exponential distribution is a single parameter function, the mean and the standard deviation are equal to $\tau = 8.3 \times 10^5$clock ticks.

We use a technique called grating, in which we remove or veto these after-pulses. Figure11

[4] An after-pulse is a pulse detected by our detector, but this pulse is a result from ion in our apparatus which was caused form our electron ionizing other atoms, creating a cascade of pulses that follow our photon.

show the corrected histogram in linear and log scale. The corrected graph exhibits an exponential form, which we should expect for a Poisson process.

$$p(\Delta t) = \frac{1}{\tau} e^{\left(\frac{-\Delta t}{\tau}\right)} \tag{3}$$

where τ is the mean interval and Δt is the intervals of time. The solid line which runs along the histogram graph represents the theoretical exponential distribution.

II. LED Brightness Analysis

The previous data was for one data set at constant LED brightness.In Figure12 we analyze how count rate of events and mean interval clock ticks between events varies with brightness. We note that photon count rate will be higher and mean interval will be shorter for increasing brightness.We observe that photon counting has statistical properties that are governed by exponential distribution

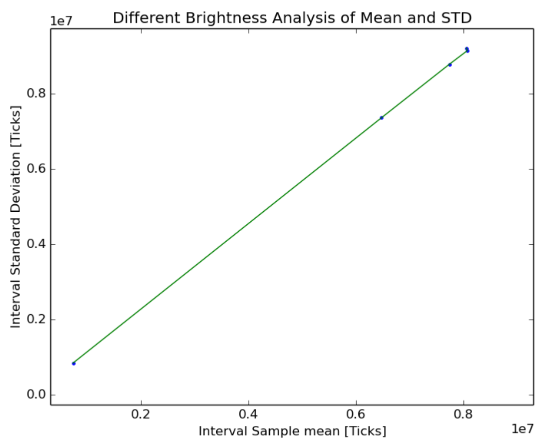


Figure 12: A plot of five different brightness level for the LED. Each point on the graph has it's mean and standard deviation plotted. Despite having five different brightness, two of the points overlap each other. This is a result of the maximum amount photons the detector is able to collect.

We can now examine brightness as a function of time. We do this by examining the counting rate and comparing it with the time. This result is found in Figure13a. We observe that brightness remains constant overtime and the arrival time of photons increases linearly. Figure13b shows a record of the light source brightness via counts per bin vs elapse time. Figure13b displays the random arrival time for the photons. Figure 14c shows a histogram of counts per time bin for the time series. This plot shows the characteristic skewed shape of the Poisson probability distribution

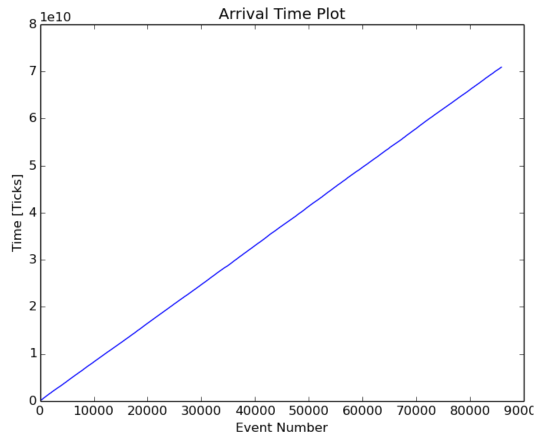
$$p(x; \mu) = \frac{e^{-\mu} \mu^x}{x!} \tag{4}$$

where x is the counts of bins that we use in the histogram, μ is the mean number of counts per bin.

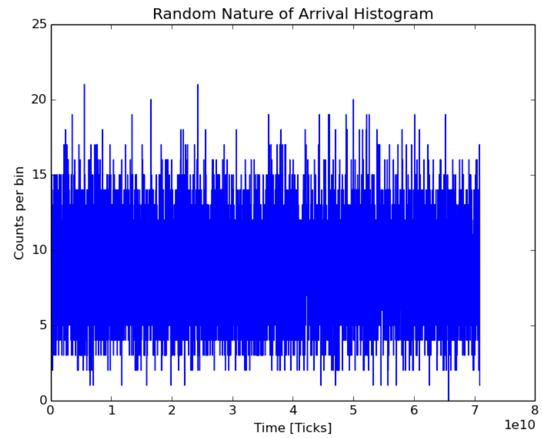
Similar to the exponential distribution, there is a unique relationship between the mean number of counts per bin and the standard deviation of counts per bin

$$\sigma^2 = \mu \tag{5}$$

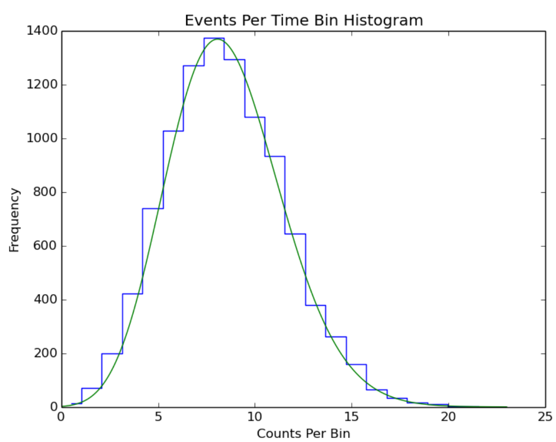
where σ is the standard deviation. Thus for a Poisson distribution, the variance equals to the mean.



(a) Sample Data



(b) Initial data plotted



(c) Sample Data

Figure 13: Figure13a shows the arrival time of photons, finding the slope leads to the number of photons detected per time. The linearity implies that photons arrive at regular intervals. Figure (b) shows the random nature of arrival, in which it counts the number of events in 1000 spaced time bins. Figure (c) shows the counts per time bin for Figure (b). Figure (c) is also fitted with a Poisson probability distribution $p(x;\mu) = \frac{e^{-\mu}\mu^x}{x!}$

V. Conclusion

In this experiment we explored the physical limitation on the detection of light. We also investigated the precision of brightness. We were able to conclude that photons are detected at a mean clock tick rate even though events occur independently of one another. for our experiment the mean was 7.1×10^5 Interval clock ticks. we also observed that SDOM varies with sample size, however, overall it exhibits an exponential behavior, which fits the theoretical prediction. The histogram plots allowed us to analyze the distribution of photon arrival time. Ignoring the initial after pulse at the beginning of the experiment, we observe that the distribution follows the expected poison process. we also analyzed how the standard deviation varied with brightness and concluded that it exhibits a linear behavior. Finally, we examined the brightness as a function of time and noted the poison distribution, the variance, is equal to the mean number of counts per bin.