Lab 1: Photon Counting with a PMT
The Statistics of Light

Joey Cheung - jcheung@ugastro.berkeley.edu
Lab Partners: Elizabeth Niespolo, Paul Higgins, Jeff Yen
September 12, 2006

Abstract

Data acquisition and error analysis are integral parts of any quantitative experiment or survey. In our introductory lab assignment we were required to analyze the statistical properties of a computerized experiment used to count photons from an LED. The end objective was to examine how errors from varying the parameters of the experiment, mainly the number of samples and the count rate, could be analyzed using statistical methods. We found that Poisson statistics provided the best fit to our data. By overlaying the Poisson distribution of the estimated parent population over the sample population, we were able to see the correlation between the samples acquired during the experiment and the theoretical Poisson probability distribution function. The Gaussian also provided a good fit to the data as we increased the number of counts that we were receiving. The Gaussian approximation got increasingly better as the number of counts went up. These approximations provided us with a way to quantify some of the physical limitations of our experiment. We found that our approximations improved by a factor of $\sqrt{N}$ where N is the number of experiments we ran for each data set.

1 Introduction

The statistical properties of photons are important to understand when studying astronomy. We know that there is a certain quantum mechanical probability that a photon will or will not be released from an excited atom. In this experiment, we attempt to analyze this probability from a statistical point of view. We will do this by examining the error and variability of photon counts as we vary certain parameters of the experiment.

2 Equipment and Methods

For this lab we used a Photomultiplier Tube (PMT), a dark box, and an LED light. The PMT itself is a very sensitive instrument that is used to detect individual photons from a light source. The light source, in this case, was an LED light. We were able to vary the intensity of the LED to control the average number of photons being sent into the PMT. Due to the sensitivity of the PMT, the device needed to be enclosed in a dark box along with other light reducing materials to ensure minimal exposure to outside light sources that could damage the device. The PMT works on a basis similar to that of the photo-electric effect. The incoming photons strike a photoemissive diode which emits electrons due to the photo-electric effect. These electrons are then accelerated through additional electrodes which affectively amplify the initially low signal, allowing for it to be collected at the final anode to be measured. All of this is attached to a local SUN workstation which provides us access to the digitized PMT data.

Various samples of data were gathered from the PMT at different rates. We had control over how bright the LED source was, how many samples we were collecting for each data set, and the rate at which the PMT would collect data[1]. We can calculate the time interval, in seconds, that the PMT was active for during each experiment by taking the inverse of the rate ($t = \frac{1}{rate}$). For example, a rate of 1000 Hz corresponds to 1 millisecond. The data acquired was then read into IDL where we created histogram plots. We used these plots to analyze some of the statistical properties of the light we were receiving from the LED. The data for this lab was gathered by myself on two separate days. The first set of data was gathered on September 4, 2006. The second set was gathered on September 8, 2006.

3 Statistics

In order to understand the statistical properties of light, we need to first understand the differences between the empirical and theoretical approaches to studying our data.

3.1 Parent Distribution and Sample Populations

The parent distribution is obtained from taking the limiting values of the sample as the number of experiments go to ∞. This is important because the parent population tells us the exact distribution of the data points. This, in turn, gives us the chance to examine the error associated with making measurements. Experimentally, the best estimation we can get for the mean of the parent population μ is the mean $\bar{x}$ of the data with the highest number of samples and the best estimation for the standard deviation σ is the deviation s obtained from the same data. The following equations[2] are for a discrete distribution.

For the parent population:

$$\mu = \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N} x_i \qquad (1)$$

[1] The rates in all of the experiments were taken in Hz

[2] from Statistics handout equations 2.5, 2.4, 2.10, 2.11

$$\sigma^2 \equiv \langle x^2 \rangle - \langle x \rangle^2 = \lim_{N\to\infty} [\frac{1}{N}\sum_{i=0}^{N}(x_i - \mu)^2] \tag{2}$$

For the Sample Population:

$$\bar{x} = \frac{1}{N}\sum_{i=0}^{N} x_i \tag{3}$$

$$s^2 = \frac{1}{N-1}\sum_{i=0}^{N}(x_i - \bar{x})^2 \tag{4}$$

3.2 Poisson Statistics

One type of theoretical distribution function that is of interest to us is the Poisson distribution. The Poisson distribution is derived from the Binomial distribution by taking the approximation that the average number of "successes" is very small when compared to the number of experiments.[3] In other words, $\mu \ll N$. Another way to think about it is that the probability of "success" for any one experiment is very small ($p \ll 1$). One of the most common examples of this is radioactive decay. The probability of any single atom decaying is very small. However, when we are looking at a very large number of atoms, we will be able to see some atoms decaying. The probability of this happening can be modeled by the Poisson distribution function.

$$P(x,\mu) = \frac{\mu^x}{x!\mathrm{e}^{\mu}} \tag{5}$$

Note that the standard deviation for the Poisson distribution is equal to the square root of the mean. This is shown in equation 6:

$$\sigma^2 = \langle (x-\mu)^2 \rangle = \sum_{x=0}^{\infty}[(x-\mu)^2 \frac{\mu^x}{x!}\mathrm{e}^{-\mu}] = \mu \tag{6}$$

3.3 The Gaussian

Another type of theoretical distribution function is the Gaussian distribution. The Gaussian is a normal distribution function that is symmetric about its center, unlike the Poisson distribution. It also extends to ∞ in both directions. This means that the Gaussian states that there is a finite probability, albeit very low for some values, that we could get a measurement of any value between $-\infty$ to ∞. Below is the equation for the Gaussian[4]:

$$P(x,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \tag{7}$$

[3]For more information on the Binomial distribution and the derivation of the Poisson Distribution refer to the Statistics lab handout (p.25-31).

[4]Taylor p. 121-135

4 Results

4.1 Experimental Errors

As I examined some of the data, I noticed that there were several plots that contained one or two isolated points at a great distance away from the mean. This occurred when I got data with a high number of photon counts (i.e. low rates). Most of these points were much lower than the mean value. I turned to statistics to analyze whether these counts were feasible or not. Statistics tells us that for a normalized distribution, 99.994% of the data lies within 4 standard deviations (σ) from the mean. For a non-normalized distribution, the same percentage lies within about 7 standard deviations from the mean. I ended up choosing 7 standard deviations to be the limit for whether I should include the counts or not in my data. Therefore, anything that fell outside of this range I interpreted as dropped counts due to experimental error or error from the equipment.

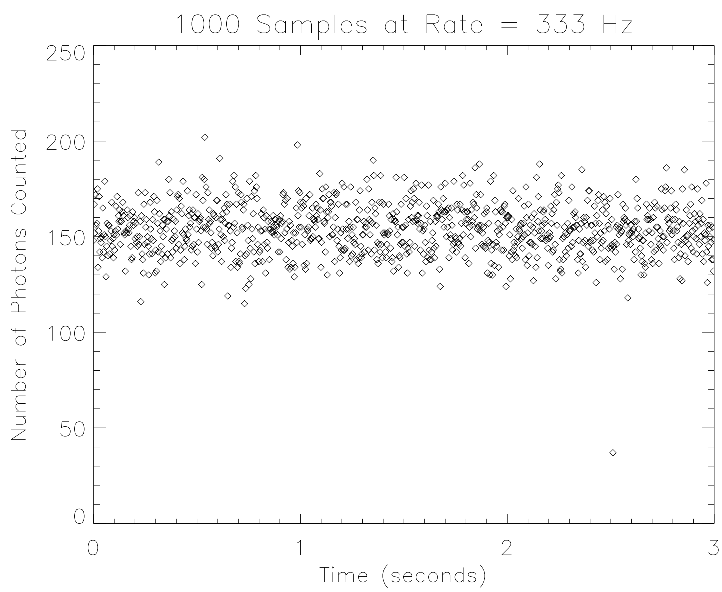
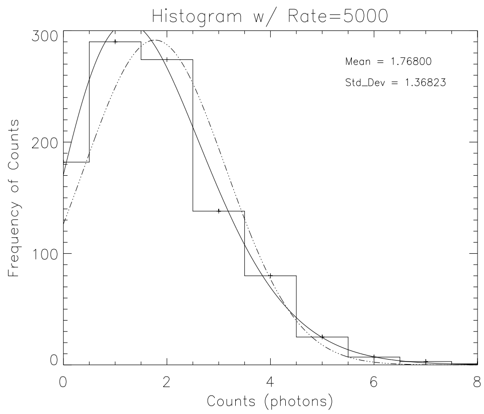


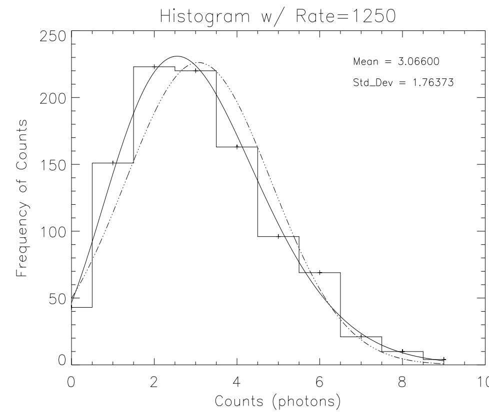
Figure 1: Here is an example of one of the plots where I saw a dropped count. You can see one data point at about 2.5 seconds after the start of the experiment that is isolated from the rest of the counts.

4.2 Histograms

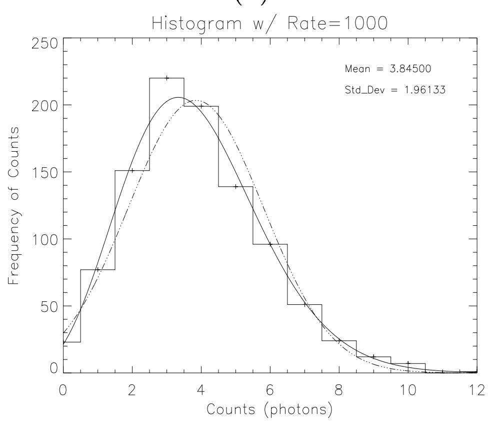
Figure 2 shows the histogram plots of some of my data ranging from a high rate to a low rate. Plotted on top of the histograms are the Poisson and Gaussian distributions of the data to show the frequency of the number of counts that we should be getting based on these statistical approximations. One problem that we encountered was how to scale the normalized theoretical distribution



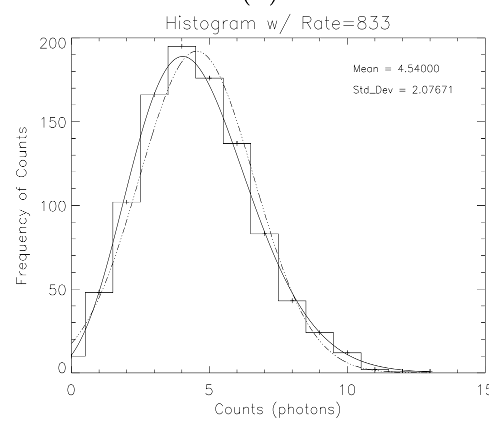
(a)



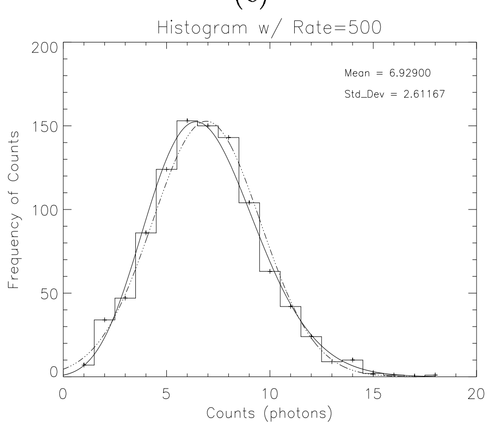
(b)



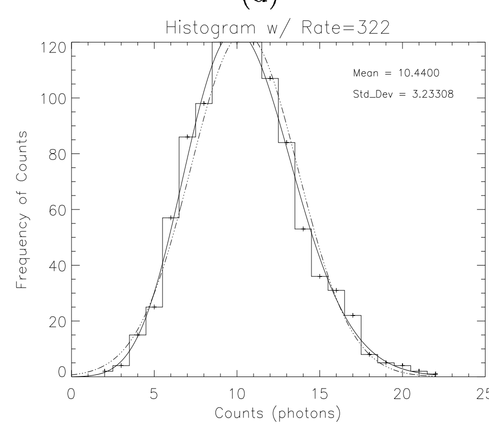
(c)



(d)



(e)



(f)

Figure 2: All of the above Histograms contain 1000 samples of data at different rates. They are also all overlayed with a Gaussian and Poisson model of the data. The Gaussian is the dotted line and the Poisson is the solid line. (a)—Rate of 5000Hz; (b)—Rate of 1250 Hz; (c)—Rate of 1000 Hz; (d)—Rate of 833 Hz; (e)—Rate of 500 Hz; (f)—Rate of 322 Hz.

functions to reflect the measured values. To solve this, we had to understand that the normalized Poisson and Gaussian curves tell us the probability that we would get a certain measurement if we did one experiment. In order to scale it up to reflect how many times we should be getting a certain measurement after doing N experiments, we had to multiply the theoretical curves by N.

As you can see from figure 2, the Poisson distribution works very well for experiments in which we get low counts. In the case of our 5000 Hz experiment, we have about 2 photons for our mean. We know that inside the LED there are a lot of atoms (on the order of Avogadro's number 10^{23}) that can emit photons. However, only a small fraction of these will actually emit photons. Thus, in the regime of our experiments, Poisson statistics matches very well with our sampled data. Additionally, you can see the asymmetry in the data that is a signature of low μ Poisson statistics. As we go higher, the sample distribution becomes less asymmetric and we can see that the Gaussian very quickly becomes a reasonable fit to the data.

One problem with the Poisson distribution, however, is how quickly it becomes difficult to calculate. By just examining the equation directly (equation 5), we can see that the μ^x part of the numerator becomes very large very quickly as we increase μ. Thus, we run into computational problems for any data sets where we get a high mean. This is where the Gaussian distribution becomes useful. However, we have to remember to be careful whenever we use the Gaussian for any real data because of the fact that the Gaussian gives a finite possibility of getting any measurement (including negative values). For this photon experiment, there is zero possibility that we can get a negative measurement. However, for a high enough μ, the probability of getting negative measurements based on the Gaussian becomes so small that it is basically zero.

4.3 More Evidence of Experimental Errors

We can further see the errors in our experiment when we examine the mean versus variance of our data as we receive an increasing number of photon counts. Figure 3 reveals a deviation of the variance from the mean as we increase the mean[5]. We know that the Poisson distribution should fit well with our data because the number of atoms within our LED is much greater than the number of photons that we are receiving. We also know that for Poisson statistics, $\mu = \sigma^2$. Figure 3 shows that experimentally this relationship slowly falls apart as we increase the mean. We also saw from above in my discussion in section 4.1 that at lower rates (i.e. high number of counts) we got more experimental errors from dropped counts and other kinds of erroneous data.

[5] The data for the Mean vs. Variance plot is different from the original histogram plots. Here I turned up the LED intensity so that the mean becomes larger more quickly as I decreased the rate between each measurement. This better shows the divergence of the variance away from the mean vs. mean line as we increase the number of counts we are receiving.

These errors serve to increase the variance of our data, so that would be the most likely explanation of the deviation of the variance from the mean. Another explanation could be that the variance is actually statistically deviating from the mean and that the Poisson distribution is actually becoming a worse approximation for the sample distribution. However, this is unlikely because even if we are getting photon counts on the order of a million, the photon counts are still much smaller than the number of atoms in the semiconductor diode of the LED that is emitting the light[6].

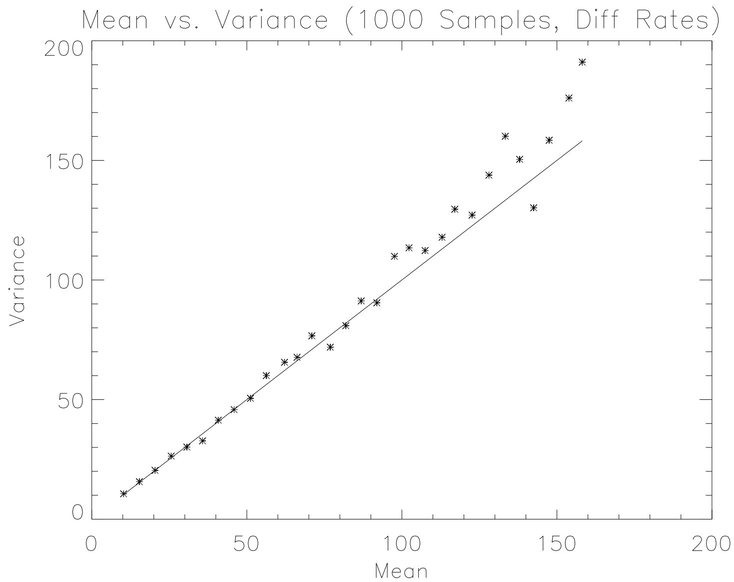


Figure 3: The plotted points are the mean vs. variance of a specific rate. The solid line is the mean vs. mean line so we can visualize the deviation of the variance from the mean as the mean increases.

4.4 Standard Deviation of the Mean

It's intuitive that the more samples we take the more accurate we can get our mean μ. What's not intuitive, however, is how to quantify by how much our accuracy improves as we increase the number of samples (N) we collect. In order to quantify this data, we examined how error on the mean changes as we increase the number of samples. For my experiment I decided to vary the number of samples taken by taking increasing exponentials of 2 (i.e. $2^1, 2^2, 2^3, 2^4, 2^5 ... 2^{12}$). Firstly, we can see how accurate the calculated mean values are for each data set as we increase the number of samples collected by comparing the mean of the means for all of the different sample sizes. So after taking the mean of each individual data set, I took the mean of all 10 data sets with the same number of samples. This is plotted in figure 4. As

[6] The number of atoms in a single drop of water is on the order of 10^{23}

you can see in the figure, the mean of the means seem to be approaching a constant as the number of samples increases. This makes sense because as you take more samples, the estimation of the mean approaches that of the parent population (equation 1).

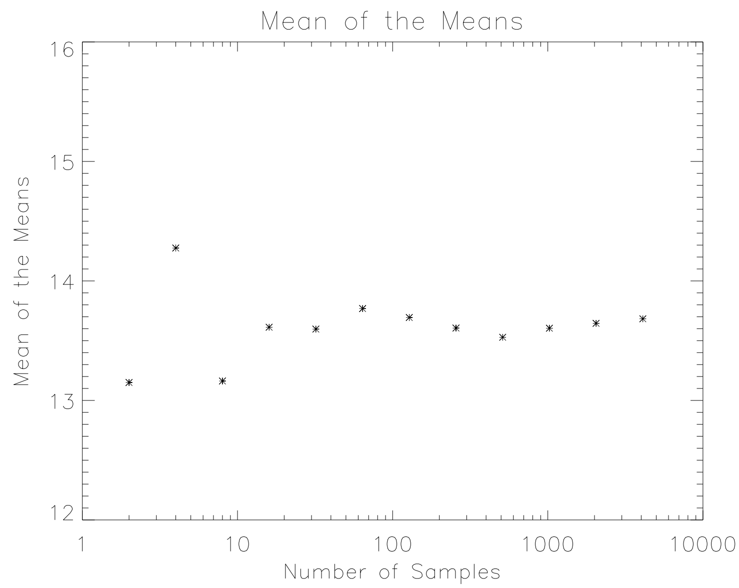


Figure 4: This is a figure of the "Mean of the Means". As you can see here, the mean of the means are, on average, constant. However, you can see that for a lower number of samples, the mean of the means is not as consistent as the ones for higher numbers of samples. It appears that as the number of samples tends towards ∞, the means converge toward a specific number.

The accuracy of the data can be further examined by looking at the standard deviation of the means (SDOM). After calculating the mean of each set of data, the deviation between the means of the data can be calculated by taking the standard deviation of all the mean values. If the accuracy in measuring the counts per sample improves as we increase the number of samples for each data set, we would hope that the standard deviation of the means will decrease as we take more samples. I have plotted the standard deviation of the means in figure 5. We can see from the figure that the deviation does indeed decrease as we take more samples of data. The solid line plotted in figure 5 is the theoretical prediction of how the standard deviation of the means should behave based on the data. The theoretical prediction is based on equation 8[7].

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}} \tag{8}$$

Equation 8 is the theoretical equation for the standard deviation of the mean. This, in effect, tells us how accurate our means are by giving us an error bound

[7]see Taylor p.102-103

for the actual μ of the parent population. So for my data, the best estimate I have for the mean of the parent population is the mean of the data set with $2^{12} = 4096$ number of samples. This has an SDOM of 0.0581. Therefore, we can state the accuracy of this mean value to be 13.6832 ± 0.0581. This error goes to zero as N approaches ∞. Additionally, we can see from the equation that in order to improve the accuracy of our mean by, say, a factor of 2, we need to take 4 times more samples. This is due to the inverse square root dependency on N for the SDOM. Therefore, if we have η more samples of data, our error (SDOM) decreases by a factor of $\sqrt{\eta}$.

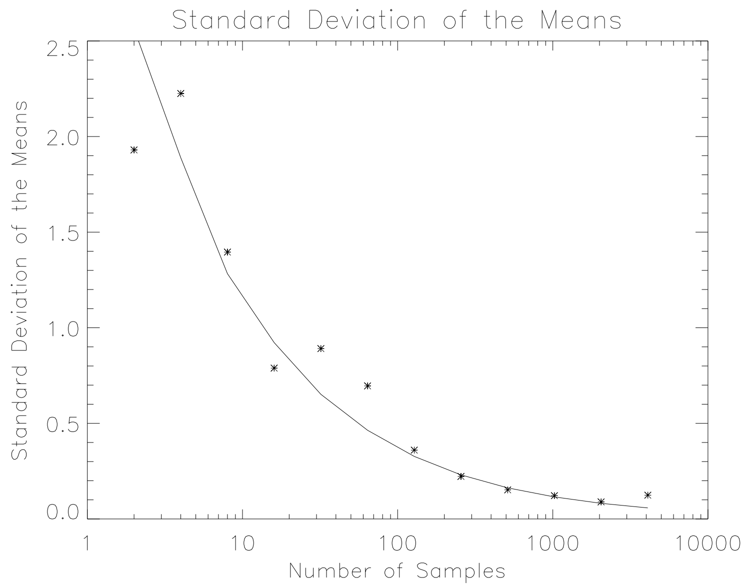


Figure 5: This is a plot of the "Standard Deviation of the Means". The points are the standard deviation of the means based on the data. The line is the theoretical prediction of how the standard deviation of the means should behave based on the data.

This brings about the question of whether or not we can construct a light source that would not show any variations in the count rate. There is not any possible way to do this because the variations come from the quantum mechanical description of the system stating that there is a probability for whether or not a photon will be emitted from the atom. This is independent of how many samples we try to take. Therefore, even if we take an infinite amount of samples, the individual variations between the photon counts will not be zero. What will be zero, however, would be the deviation between the average means between each data set. If we can take an infinite amount of samples for each data set, each data set as a whole should give us the same number of photon counts as an average.

5 Summary of Conclusions

We found that our sample distribution matched fairly well with Poisson statistics. Furthermore, the Gaussian distribution of the data very quickly became a good fit to the data. Both experimental and statistical errors were apparent in our experiment as well. Experimental errors came from the equipment dropping counts in some cases. This could be somewhat minimized by limiting the range of our data to a few standard deviations from the mean to eliminate some of the more erroneous data. Even after doing this, however, there was still enough experimental error that could be seen from figure 3. Statistical errors, on the other hand, provided us ways to quantify the error in our data. By calculating and examining the standard deviations, we could analyze how the error changes as we vary different parameters of the experiment. In the end, we found that decreasing the rate or increasing the brightness of the LED created more experimental errors, but did not change the statistical errors very much. The Poisson and Gaussian were very good fits for all of our data sets. Increasing the number of samples for each data set, on the other hand did show a change in the error of our approximations. We found that as we increased the number of samples for each data set, the error that we can state for our mean would decrease by a factor of $\sqrt{N}$. Therefore, to increase the accuracy of our data by a factor of 2, we had to take 4 times the number of samples for each data set.