Lab Report 1
Photon Counting with a Photomultiplier Tube (PMT)
Statistics of Light

Kirsten Howley [1]
Group 1
September 11, 2001

1. BACKGROUND & MOTIVATION

By analyzing data received from the photomultiplier tube, we are able to hypothesize about the statistics of photon counting. We can measure how variations in our parameters affect the probability of obtaining a"true" result. Our goal is to measure these random errors and compare them to their theoretical statistical counterparts in order to cultivate an understanding of the application of error analysis techniques.

This experiment also provides an introduction to using programs vital for analyzing and reporting on astrophysical data. This lab provides exposure to the programming languages of UNIX, IDL and LaTeX; all three of these programs will be essential in future lab projects.

2. EXPERIMENT & METHOD

Our experiment began by obtaining data from the photomultiplier tube. Our photometer consisted of a H5773 photomultiplier module, photon statistics demo unit, pulse amplifier, discriminator, squawker, counter, and CIO-CTR05 (counter card). A server program on the IBM PC (located in the node *pulsar.ugastro*) read out the counters. The limit on the sample rate was set at 5000Hz by an internal quartz crystal clock. This corresponds to a minimum period of 0.2 milliseconds per sample.[2]

In reporting data, the Rule for Stating Uncertainties and the Rule for Stating Answers was invoked. That is, "Experimental uncertainties should almost always be rounded to one significant figure," and, "The last significant figure in any stated answer should usually be of the same order of magnitude (in the same decimal position as the uncertainity," respectively.[3] Graphs and tables were also employed to help demonstrate the relationships between different parameters.

---

[1] E-mail: kirsten@ugastro.berkeley.edu

[2] See www.ugastro.berkeley.edu/optilab/photometer/index.html for further information

[3] Taylor, John R.

3. DATA COLLECTION & ANALYSIS

The first step in gathering data involved running a program in UNIX to communicate to the PMT to collect counts. Once that process was completed, it was a simple procedure to import the information into IDL and manipulate it into an understandable format.

Although the flux of photons into the PMT could be manually controlled, the two main variables in this experiment were the sample rate and sample size. (We will see in the subsequent sections how varying these parameters affected the size of the random errors in our results.) First, six consecutive data collection sets were run, each requiring a tenth of a second to complete.[4] In each set 100 samples at a sample rate of 1000Hz were taken. Figure 1 demonstrates the number of counts per millisecond for each acquisition.

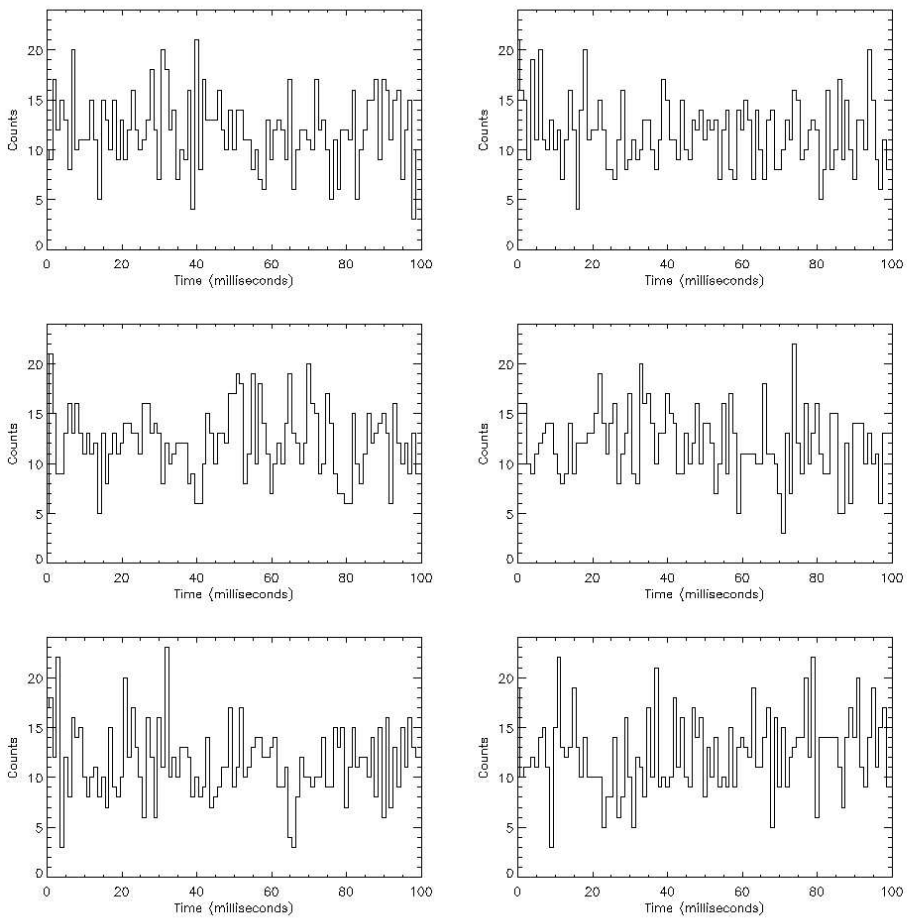

Fig. 1.— Plots of Photon Counts. For each plot, 100 samples were taken at a rate of 1000Hz.

---

[4]Neglecting communication time through the cord and processing time of the system

As can be seen, the average count rate lies somewhere between ten and fifteen counts. The histogram in Figure 2 provides a better view of the distribution of count rates.

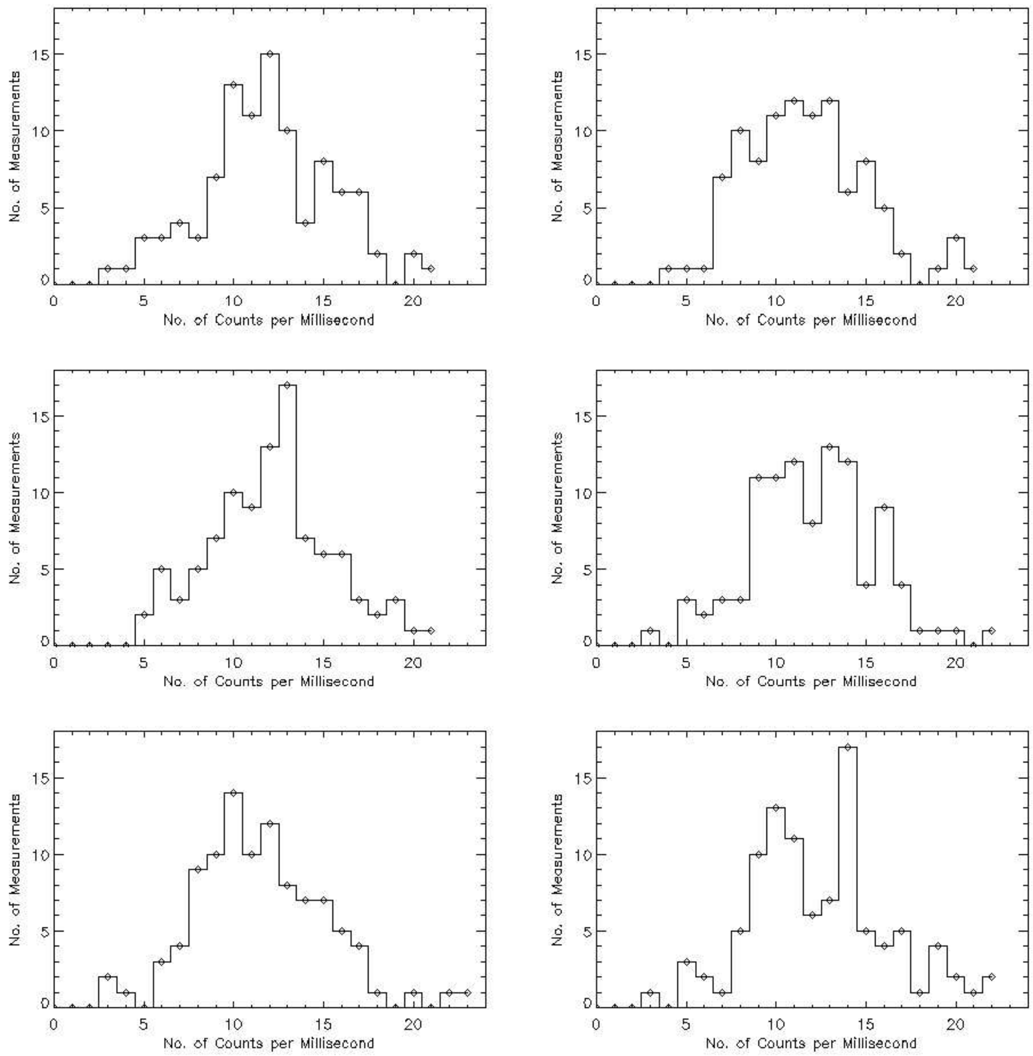

Fig. 2.— Histogram Plots of Photon Counting with PMT. For each plot, 100 samples were taken at a rate of 1000Hz. Although the specific structure of each histogram vaires, notice that the number of counts per millisecond for each collection set lies within a similar range.

Since we acquired all of our data in less than a second, ensuring nearly indentical environments for each data collection, we can attribute fluctuations in our data to random errors. This assumption is further supported by the high precision in our collection, that is, by the consistency in the range of our data. Determining the accuracy of the experiment requires an analysis of the design of our experiment. Because the focus of this lab is statistical analysis, we will neglect possible miscalibration of the equipment, since these are not easily measured by statistical analysis.

In order to determine the best estimate of the number of photon counts per millisecond, we take the mean. We can calculate the mean of each data set by

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

The average uncertainty of each measurement, x_i, is estimated by the standard deviation. This tells how much each measurement differs from the mean. The smaller the deviation, the more precise each measurement. We can calculate the standard deviation of each data set by

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

The tablulated results of the first data collections are listed in Table 1:

Table 1: Means and Standard Deviations: tabulated results of first data collection

Collection Set No	Sample Size	Sample Rate (Hz)	Mean (counts)	Std Deviation (counts)
1	100	1000	11.9	3.6
2	100	1000	11.6	3.4
3	100	1000	12.1	3.5
4	100	1000	12.0	3.4
5	100	1000	11.5	3.6
6	100	1000	12.4	3.9

As you can see, we did not get the same mean count rate for each data collection. However, we would not expect to since we did not vary our sample rate or sample size (the two variables). The variation in the values of the means indicate random errors in the data. By calculating the mean of the mean (MOM), we can determine how much our original means deviated from one another:

$$\bar{x}_{\bar{x}} = 11.9 \text{ counts} \qquad \text{s}_{\bar{x}} = 0.3 \text{ counts}$$

The same experiment was repeated again for a larger sample size. We increased the sample size by a factor of ten (for a sample size of 1000), computed $\bar{x}_{\bar{x}}$ and the standard deviation of the means, $\text{s}_{\bar{x}}$, and found:

$$\bar{x}_{\bar{x}} = 12.0 \text{ counts} \qquad \text{s}_{\bar{x}} = 0.1 \text{ counts}$$

Although the mean count rates were relatively the same, the standard deviation of the means decreased significantly. This demonstrates that increasing the number of measurements in this experiment (sample size) increased the accuracy of the mean of the mean, $\bar{x}_{\bar{x}}$.

Using the same procedure as in the first data collection, additional data was collected to determine how sample size and sample rate affect the random errors produced. The following tables list the mean count rates and standard deviations for various collections.[5] These results will be discussed further in the following section.

Table 2: Means and Standard Deviations: tabulated results of varying sample rates

No of Collection Sets	Sample Size	Sample Rate (Hz)	Mean (counts)	Std Deviation (counts)
1	100	5000	2.3	1.4
2	100	3333	4.0	1.7
3	100	2500	4.8	2.2
4	100	1666	7.3	2.8
5	100	1250	9.8	3.4
6	100	1111	10.8	3.3
7	100	1000	12.1	3.5
8	100	833	14.4	4.2
9	100	714	17.4	4.6
10	100	666	18.0	4.6
11	100	625	19.4	4.9
12	100	555	21.3	4.9
13	100	500	23.8	5.6
14	100	454	27.2	5.9
15	100	416	28.9	6.9
16	100	384	30.5	6.6
17	100	370	32.5	7.2

The results from the table above demonstrate that as the sample rate decreases, the standard deviation increases, suggesting an inverse relationship. However, notice that the percent error (deviation) for each data collection actually *decreases* with sample rate.

Table 3: Means & Standard Deviations: tabulated results of high and low count rates

No of Collection Sets	Sample Size	Sample Rate (Hz)	Mean (counts)	Std Dev (counts)
1	1000	5000	2.4	
1	1000	111	421.3	20.2

[5]The flux of photons was altered between tables, and between collection sets in Table 3.

Table 4: Means & Standard Deviations of the Means: tabulated results of various sample sizes

No of Collection Sets	Sample Size	Sample Rate (Hz)	$\bar{x}_{\bar{x}}$ (counts)	$s_{\bar{x}}$ (counts)
10	2	1000	59.6	5.7
10	4	1000	58.9	3.1
10	8	1000	58.1	3.0
10	16	1000	58.5	2.4
10	32	1000	59.0	1.6
10	64	1000	59.0	1.0
10	128	1000	59.3	0.6
10	256	1000	58.6	0.7
10	512	1000	58.4	0.3
10	1024	1000	58.6	0.2
10	2048	1000	58.6	0.1

The results from the table above show that as the sample size is increased, there is no change in the mean; however, there is a significant decrease in the standard deviation, implying an inverse relationship between the standard deviation and the sample size.

4. COMPARISON TO THEORY

4.1. Poisson Statistics

Poisson statistics claims that the square of the standard deviation, or the variance, is equal to the mean:

$$s^2 = variance = \bar{x} \tag{3}$$

Upon comparison with the data obtained in Table 2, it can be seen that the predicted linear plot of mean vs variance is a good approximation to the actual plot of mean vs (standard deviation)^2. We can simply plot the line, mean vs mean, to show this correlation. This relationship is shown in Figure 3.

The Poisson Distribution predicts the probability of obtaining a specific count rate for a discrete set of values; negative values of x are not allowed. It is antisymmetric about the mean and relies on only one parameter, the mean. The reason for this is that Poisson statistics defined the standard deviation to be a function of the mean, thus negating the need for two variables in the calculation of the distribution. The distribution function is given by:

$$P_{\bar{x}}(x;\bar{x}) = \frac{(\bar{x})^x \cdot e^{-\bar{x}}}{x!} \tag{4}$$

Plotting the data from Table 2 and the corresponding Poisson Distribution shows that the Poisson Distribution provides an excellent prediction in the low count limit.

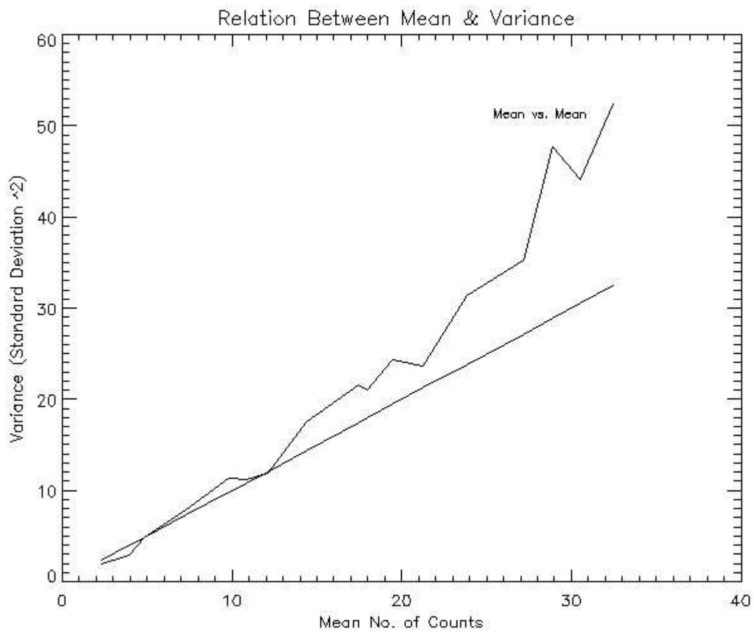

Fig. 3.— Mean vs (Standard Deviation)^2. Notice how Poisson statistics predicts a line that approximately fits the distribution.

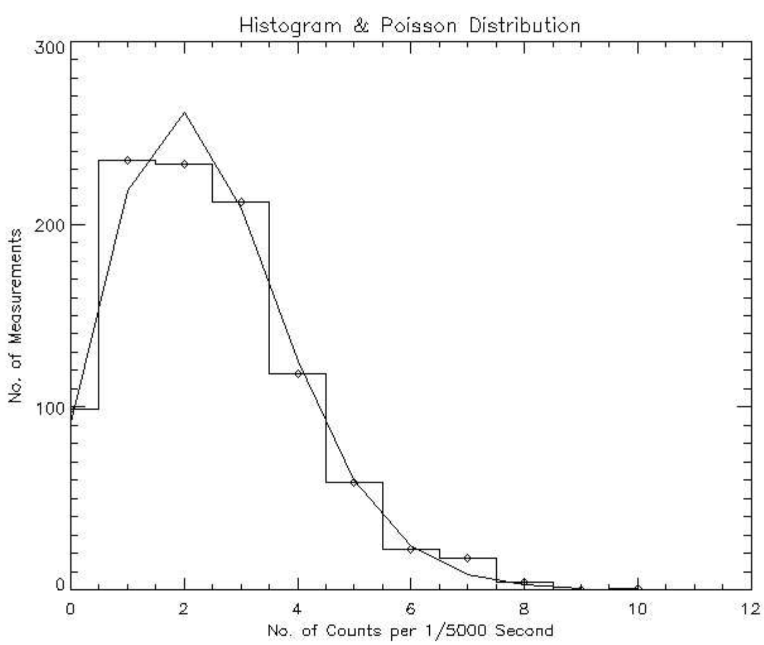

Fig. 4.— Poisson Distribution plotted over Histogram

Since the Poisson Distribution is a calculation of the probability, the sum of all of its components is 1. Thus, it is necessary to multiple your distribution by your sample size when comparing it with the corresponding histogram. Furthermore, if the binsize of the histogram is not one unit, the variables in the equation must be adjusted by $\frac{1}{binsize}$ to compensate for this scaling.

4.2. Gaussian Distribution

The Gaussian Distribution provides a symmetric, "bell shaped" curve as a prediction of distribution. Unlike the Poisson Distribution, the Gaussian Distribution incorporates both positive and negative values. It is a good approximation in the high photon count limit, where the Poisson distribution becomes extremely difficult to solve (even by computer), due to the factorial in the denominator. It is fortunate though that, in the high count limit, the Poisson Distribution approaches the Gaussian Distribution. Of course, the setback of this method is that two parameters are need to calculate the Gaussian Distribution: the mean and the standard deviation. The mean determines the location of the peak, and the standard deviation dictates the width of the curve. The distribution function is given by:

$$G_{\bar{x},s}(x;\bar{x},s) = \frac{e^{(-\frac{1}{2}\cdot(\frac{x-\bar{x}}{s})^2)}}{s\cdot\sqrt{2\pi}} \tag{5}$$

Plotting a Gaussian Distribution over the high count data from Table 3 yields the following result:

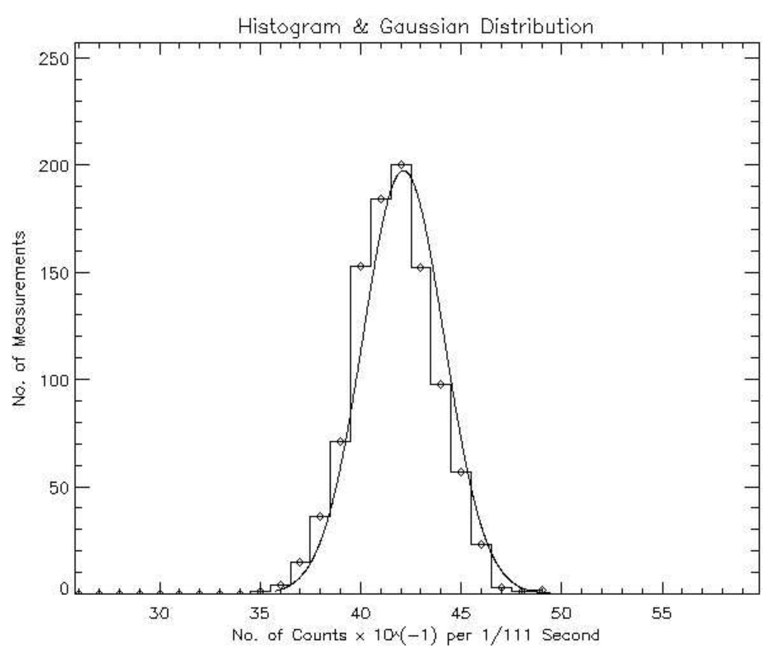

Fig. 5.— Gaussian Distribution plotted over Histogram

4.3. Affects of Sample Size

It has already been demonstrated in Table 4 that increasing sample size has little affect of the mean of the mean number of counts. Poisson statistics provides the key to predicting how the mean of the mean, $\bar{x}_{\bar{x}}$, as well as the standard deviation of the mean, $s_{\bar{x}}$, might vary with sample size. Based on Equation 1, it would be expected that the mean of the means would closely approximate the values from which it was constructed. Using the data from Table 4 and plotting the relation between sample size and $\bar{x}_{\bar{x}}$ confirms the prediction:

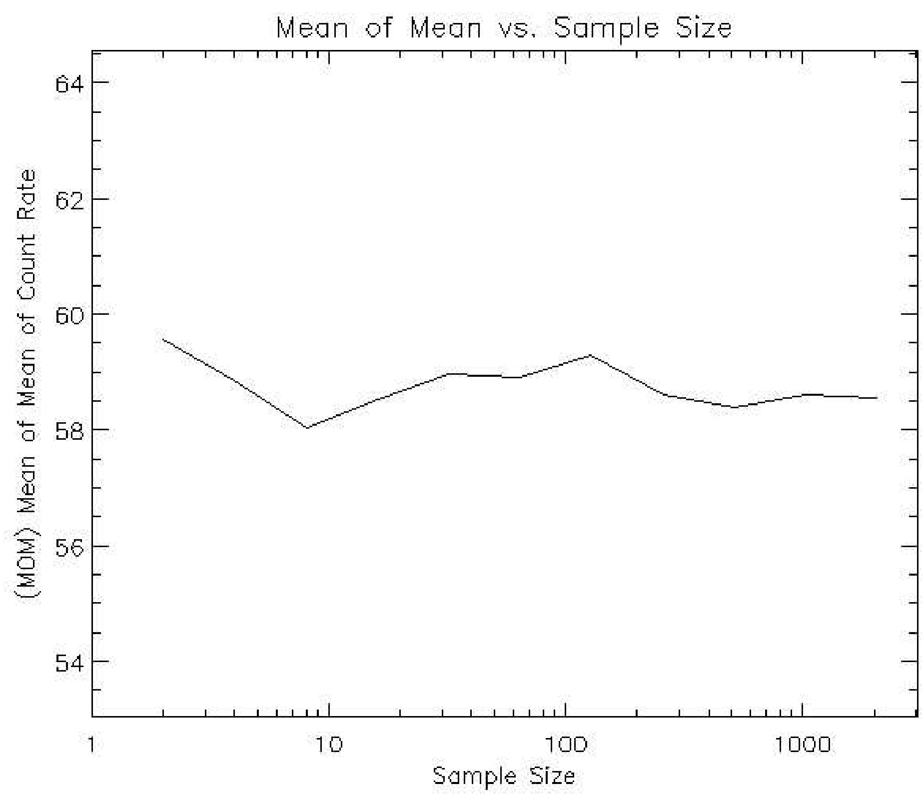

Fig. 6.— Mean of Mean vs Sample Size

It was also seen from Table 4 that increasing the sample size results in a more accurate measurement of the number of counts per sample. This means that the strandard deviation of the mean decreases as sample size increases. Poisson statistics predicts such a result. From Equations 2 & 3, the following conclusion can be drawn:

$$s_{\bar{x}} = \sqrt{\frac{\bar{x}}{N}} \tag{6}$$

Graphing the data from Table 4 and comparing the curve predicted by Poisson statistics results in almost a perfect match, as shown in Figure 7:

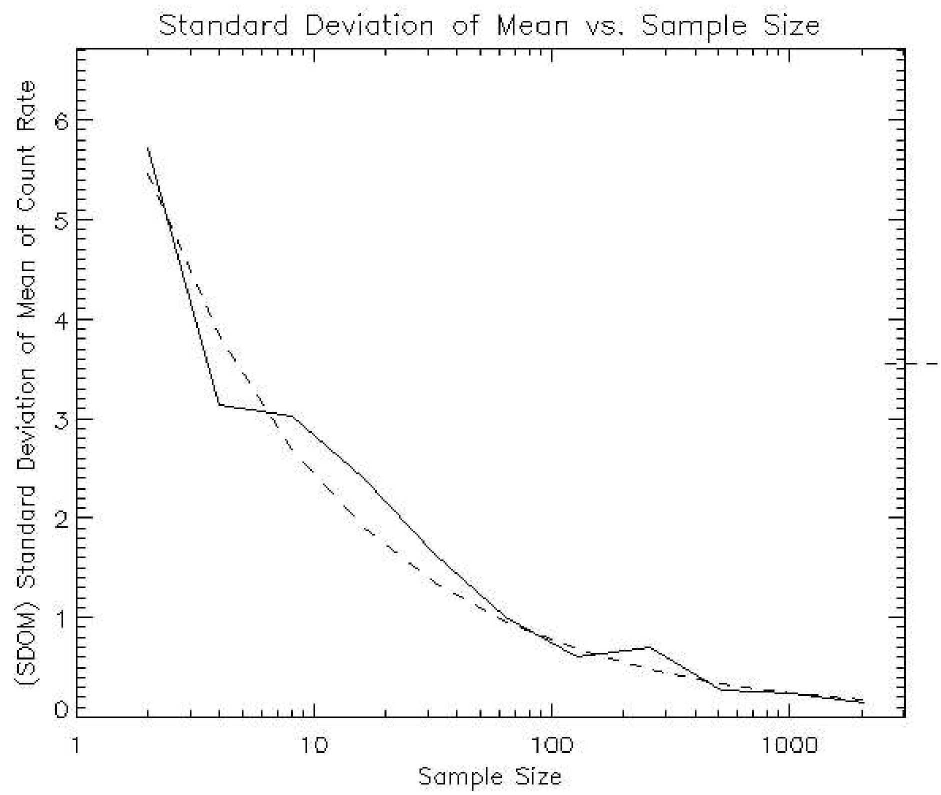

Fig. 7.— Standard Deviation of Mean vs Sample Size

5. CONCLUSIONS

To be able to predict random errors in our data allows us determine how precise our measurements are. Since it is impossible to take an infinite number of measurements, understanding how different parameters affect our results is key to reducing errors.

There are many practical applications for such statistical analysis. For example, in astrophysics, data is collected via photons. For this reason, it is essential to be able to determine how accurately and to what precision we are able to detect individual photons. Morever, given a distribution of photons, what would be our expected deviations? The ability to predict such distributions provides additional tools in the analysis of light. Ultimately, deviations in photon counts from predicted distributions provide useful information about occurences in the cosmos, sparking exploration into their sources.

6. REFERENCES

Taylor, John R. *An Introduction to Error Analysis.* Sausalito: University Science Books, 1997.

Graham, James R. *Errors & Statistics.* AY120/122 Lab notes.

7. ACKNOWLEDGEMENTS

I would like to thank Eric especially, for patiently explaining how to make things "go" in IDL. I'd also like to thank Jim for showing me the xv program to get those plots into the LaTeX document. I would also like to thank Christina, for being there to talk about the project and go with me to grab tea and cookies! I'd also like to thank Andy for the great proooofreading job, my lab partner Lee for being there to bounce ideas off of. And, of course, Lauren for wearing those cute, inspiring pink slippers. I'd also like to thank 'Lon' for those tasty M&M's which saved me from starvation. An additional special thanks goes out to the whole crew who deprived themselves of sleep along with me Monday night, I look forward to doing that again (but let's time let's get three pizzas and two liters of caffeine). And last, but certainly not least, thanks to Professor Graham & Nate, for those life saving tips and explanations. See you all next project!