An Investigation Of The Performance Of Poisson Statistics In Approximating Radioactive Decay Counts With Uncertainty.

Melissa Githinji, GTHMEL001

Submitted 26 April 2024

1 Abstract

The aim of this experiment is to use decay count detection to investigate how well mean and variance obtained using Poisson statistics, from long spectrum population datasets, holds up to sampled short interval datasets. A cumulative average shows that over time, the approximation of the mean μ gets better. Therefore, long spectra provide a population mean for the Poisson distribution that has a variance that in the worst case is 5.8% away from the sample mean for μ= [0.83, 315]. In this way the Poisson performs within one standard deviation of the sample. The gaussian distribution is equivalent to the Poisson distribution for samples with $\mu > 30$ by a at least a 99% confidence. For samples with $\mu > 315$, the gaussian performs like a Poisson with at least 99.5% confidence level.

2 Method

A 2" x 2" NaI scintillation detector was used in the setup of the apparatus with a 700 V power supply. The detector was connected directly to an UCS30 acquisition unit and placed on top of a fixed metal stand at the end of a ruler. The ruler had a metal track with another stand which could be detached and placed at a variable position along the ruler, with prestick on top to secure the radioactive source ^{137}Cs in front of the circular surface of the detector.

The source was initially placed around 10 cm away from the face of the detector and the USX acquisition software was opened to calibrate the apparatus. The mode was set to PHA-PreAmp In by default. After adjusting the course and fine gains of the power supply, the 0.662 MeV photopeak was centered at approximately channel 256 and a full pulse height spectrum was measured for the radioactive decay of ^{137}Cs for 600 seconds, saved as a .csv.

Thereafter the acquisition was windowed to record the pulse height spectrum about 50 channels either side of 256 which which came out to channels 205 - 310. The spectrum was taken for 600 seconds and saved.

The mode was changed to MCS - Internal to record samples of events that fall within the set pulse height window for a chosen time interval known as the dwell time. At the fixed dwell time of 20 ms, a MCS measurement was made and USX showed the number of counts recorded over the channels as a scatter plot. Using that graph, an estimation was made for the average number of counts per interval (CPI).

Using this initial measurement, the distance of the source from the detector was adjusted and new measurements were made until approximately 1 CPI was observed. The MCS measurement at this distance was recorded and the mode was switched back to PHA-PreAmp to record a 600 second pulse height spectrum (with the window still on). After switching back to MCS mode, this process was repeated for approximately(within 25% of) 4, 10 and finally 30 CPI. With a MSC and PHA-PreAmp measurement saved as a .csv for each.

At the same source-detetcor distance used for the 30 counts per time interval run, the dwell time was doubled to 40ms and the MSC measurement was taken to observe approximately 60 CPI. The dwell time was increased to 100ms, and then to 200ms and the .csv of the MSC measurements were saved.

3 Data and Results

Data was processed and graphed by a self-authored Python program.

To analyse the statistical information from the data, the following mathematical equations are used to extract known quantities using different models. This will form the basis of discussion and comparison of different methods of calculating a value like the mean, μ.

The cumulative average, r, over an expression of time, j, is:

$$r_c(j) = \sum_{i=1}^{j} C_i / \sum_{i=1}^{j} i \tag{1}$$

The Poisson distribution is a probability mass function given by:

$$P(r) = \frac{(\lambda \Delta t)^r e^{-\lambda \Delta t}}{r!} \tag{2}$$

This discrete function is an expression of probability and therefore to get the predicted outcomes as a frequency it is further scaled by multiplying the probability by the total number of events. This scaling factor is equal to the 1024 (number of events = number of channels of acquisition).

And the Gaussian distribution is:

$$f(x) = Ae^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3}$$

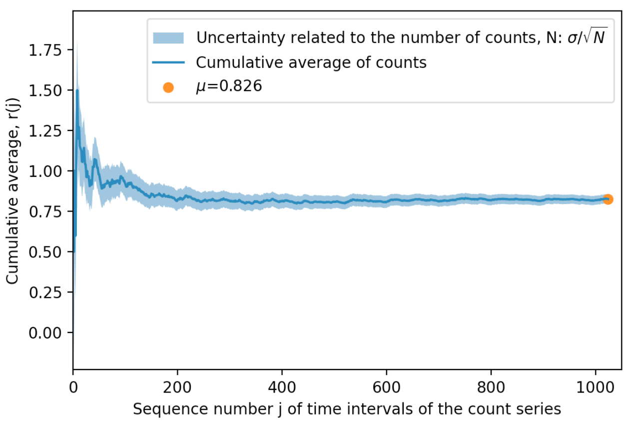

FIG 1

A graph showing the cumulative average given by equation (1) for a MCS mode measurement at a distance of 22.0cm for a dwell time of 20ms. This run was for CPI ≈ 1.

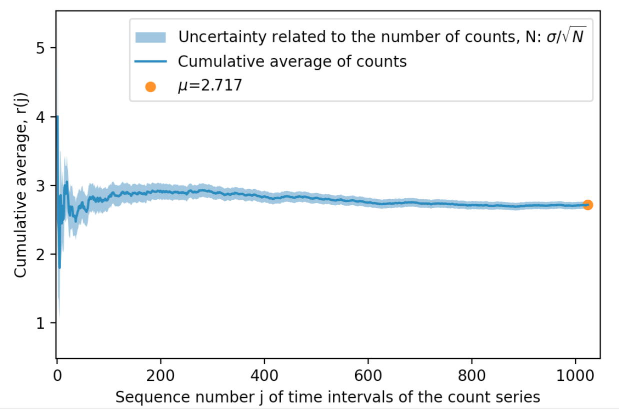

FIG 2

A graph showing the cumulative average given by equation(1) for a MCS mode measurement at a distance of 11.4cm for a dwell time of 20ms. This run was for CPI ≈ 4.

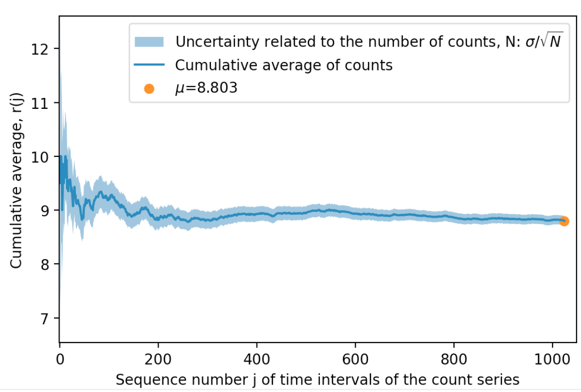

FIG 3

A graph showing the cumulative average given by equation(1) for a MCS mode measurement at a distance of 4.4cm for a dwell time of 20ms. This run was for CPI ≈ 10.

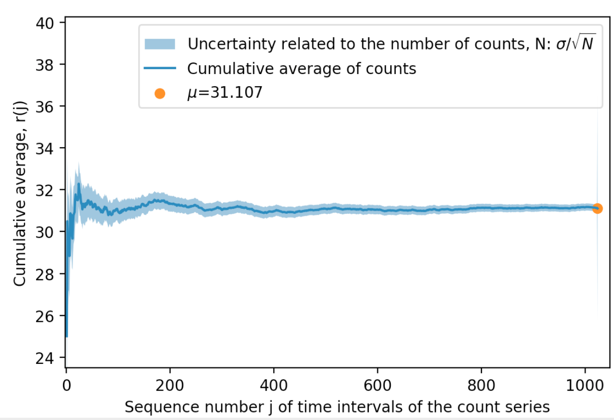

FIG 4

A graph showing the cumulative average given by equation(1) for a MCS mode measurement at a distance of 1.0cm for a dwell time of 20ms. This run was for CPI ≈ 30.

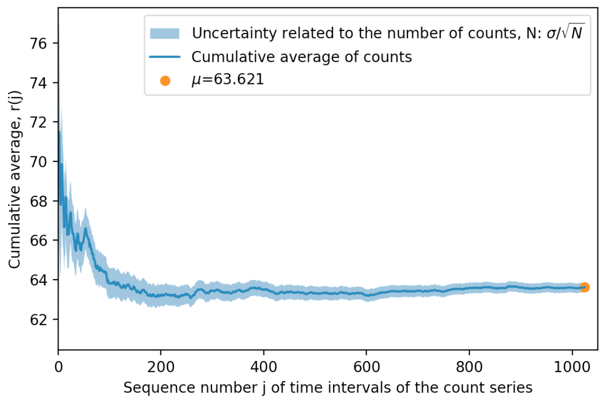

FIG 5

A graph showing the cumulative average given by equation(1) for a MCS mode measurement at a distance of 1.0cm for a dwell time of 40ms.

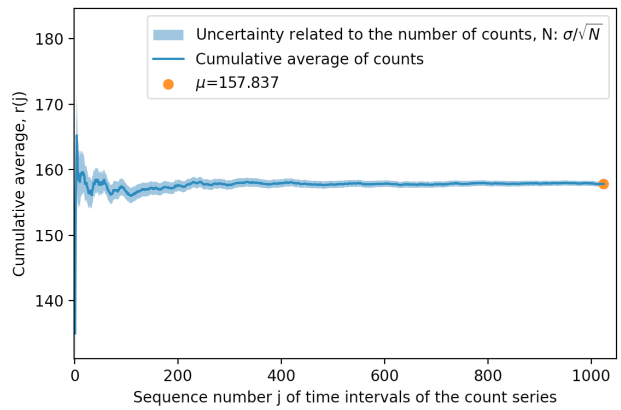

FIG 6

A graph showing the cumulative average given by equation(1) for a MCS mode measurement at a distance of 1.0cm for a dwell time of 1000ms.

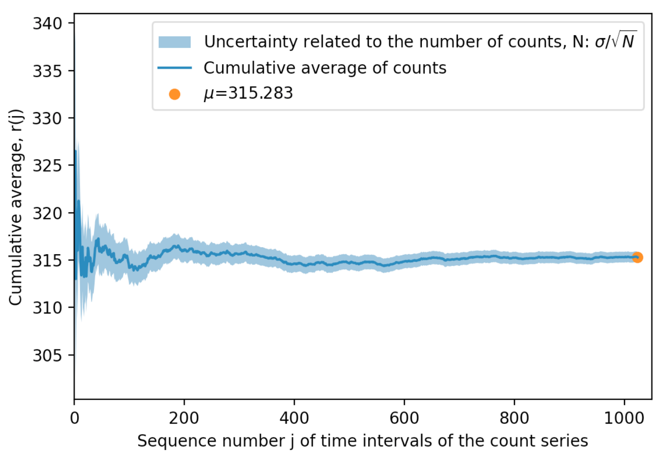

FIG 7

A graph showing the cumulative average given by equation(1) for a MCS mode measurement at a distance of 1.0cm for a dwell time of 1000ms.

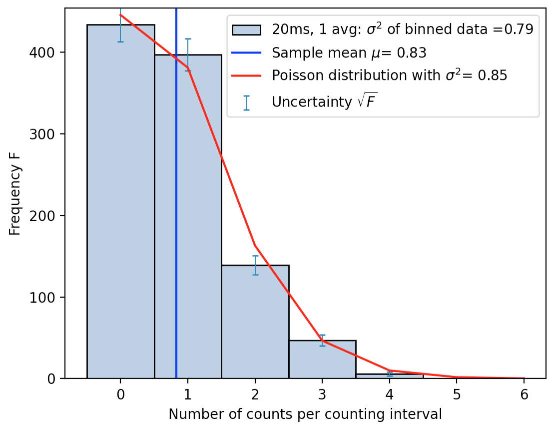

FIG 8

Graph showing a histogram of the sample MCS data in figure 1, showing the μ point from the cumulative average. The Poisson distribution(2) was plotted using the population mean from the long spectrum associated with this MCS dataset. The population mean used was time scaled from 600 seconds to 20ms by dividing number of counts recorded by 30 000.

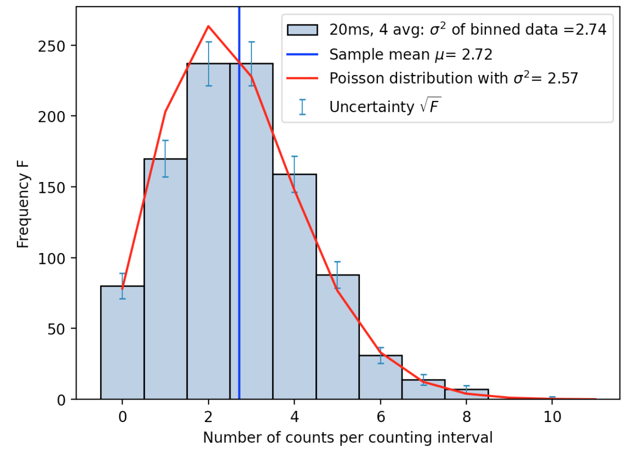

FIG 9

Graph showing a histogram of the sample MCS data in figure 2, showing the μ point from the cumulative average. The Poisson distribution(2) was plotted using the population mean from the long spectrum associated with this MCS dataset. The population mean used was time scaled from 600 seconds to 20ms by dividing number of counts recorded by 30 000.

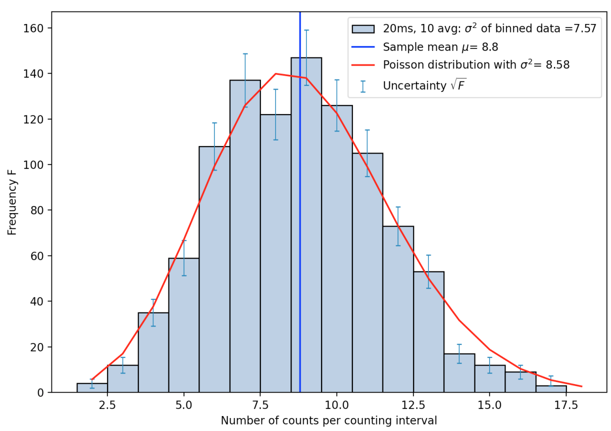

FIG 10

Graph showing a histogram of the MCS data in figure 3, showing the μ point from the cumulative average. The Poisson distribution(2) was plotted using the population mean from the long spectrum associated with this MCS dataset. The population mean used was time scaled from 600 seconds to 20ms by dividing number of counts recorded by 30 000.

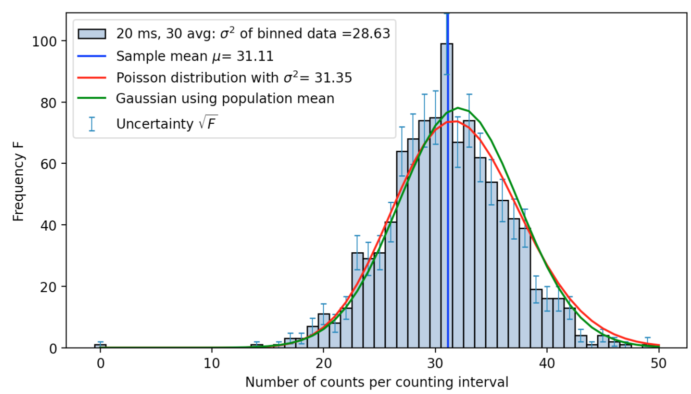

FIG 11

Graph showing a histogram of the sample MCS data with figure 4, showing the μ point from the cumulative average. The Poisson distribution(2) was plotted using the population mean from the long spectrum associated with this MCS dataset. The gaussian distribution(3) is plotted using the same population mean and associated standard deviation. The population mean used was time scaled from 600 seconds to 20ms by dividing number of counts recorded by 30 000.

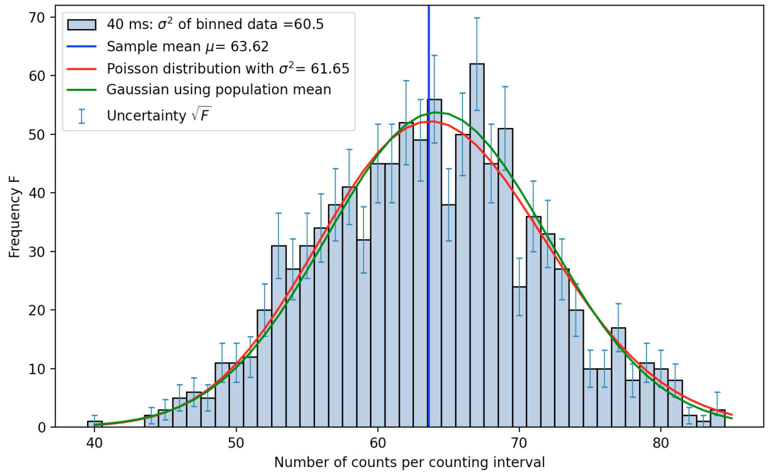

FIG 12

Graph showing a histogram of the MCS data in figure 5, showing the μ point from the cumulative average. The Poisson distribution(2) was plotted using the population mean from the long spectrum associated with this MCS dataset. The gaussian distribution(3) is plotted using the same population mean and associated standard deviation. The population mean used was time scaled from 600 seconds to 40ms by dividing number of counts recorded by 15 000.

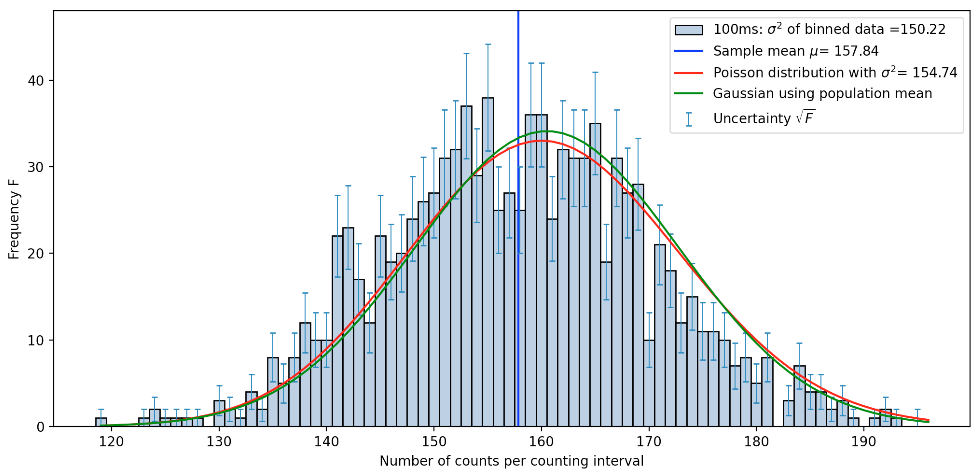

FIG 13

Graph showing a histogram of the MCS data in figure 6, showing the μ point from the cumulative average. The Poisson distribution(2) was plotted using the population mean from the long spectrum associated with this MCS dataset. The gaussian distribution(3) is plotted using the same population mean and associated standard deviation. The population mean used was time scaled from 600 seconds to 40ms by dividing number of counts recorded by 6000.

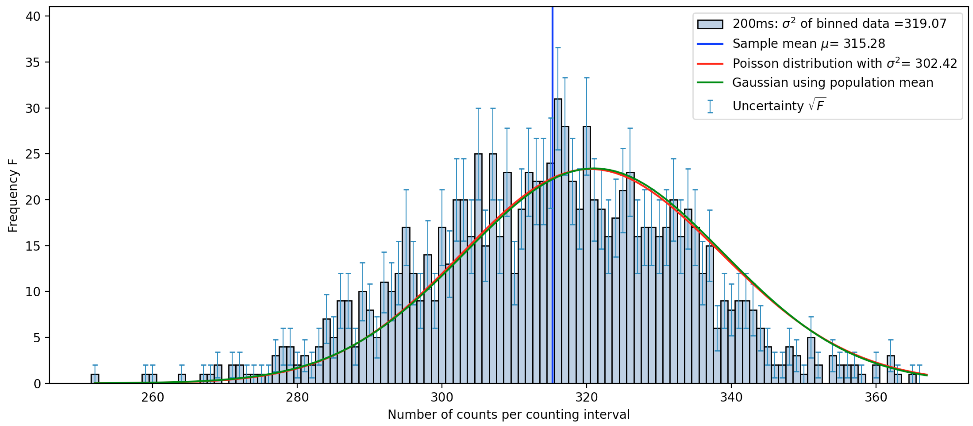

FIG 14

Graph showing a histogram of the MCS data in figure 7, showing the μ point from the cumulative average. The Poisson distribution(2) was plotted using the population mean from the long spectrum associated with this MCS dataset. The gaussian distribution(3) is plotted using the same population mean and associated standard deviation. The population mean used was time scaled from 600 seconds to 40ms by dividing number of counts recorded by 3000.

4 Analysis and Discussion

In figures 1-7, the cumulative frequencies and their uncertainties are seen to converge to the sample mean quoted in figures 8-14 respectively.

The Poisson frequency distribution is what was expected probabilistically by using the population mean from the significantly longer spectra after scaling as a stronger representation of an average over time. While the histograms in figures 8-14 are a sample of what happened in a small time interval.

The data set with the lowest mean counts per interval is the one shown in figure 8 with an average of 0.83 counts per interval. The data point with the highest deviation from the mean is when the detector measured 4 counts. There were 1024 data points and the frequency for 4 counts is 6 as read off of the graph. The event happened 6/1048 of the time and the Poisson distribution predicted 4 counts to occur with a probability of 10/1048. Considering the sample had a deviation of 0.382% (very small) at this point, it is justified to conclude that the detector was working properly and what was observed was ordinary experimental uncertainty.

In discussing confidence level in the Poisson result it can be seen in figures 8 - 10 that for small values of μ, the Poisson is within one standard uncertainty in most cases, hardly missing the peak by many.

The variance of the Poisson distribution in figures 8-13 is consistently equal to the mean of the sample. Furthermore, the Poisson variance is always closer to the mean than the variance of the binned dataset. The 2 variances are in agreement to each other as well. The Poisson distribution's variance performs well with good accuracy to the mean. Table 1 shows that performance isn't directly related to number of counts, and that the deviation from the sample mean as an uncertainty takes on varying percentages.

Table 1 Showing The Performance of The Poisson Distribution Variance To Approximating The Mean Of The Samples In Figures 8-14

Sample mean μ (N)	Uncertainty (%)
0.83	2.4
2.72	5.8
8.80	2.6
31.1	0.766
63.6	3.20
158	2.00
315	4.25

This goes to show that the lack of sample data can be reliably supplemented by a Poisson distribution with at least 94.2% confidence to obtain a mean. This is a very useful result and strengthens the accuracy of Poisson statistics.

Table 2 Showing The Percentage Deviation From The Entire Sample For The Poisson and Gaussian Distributions in Figures 11-14 In Order To Determine Overall Similarity Between The Two.

Sample mean μ (N)	Deviation of The Poisson Distribution From The Sample(%)	Deviation of The Gaussian Distribution From The Sample(%)
31.1	18.0	19.0
63.6	15.1	16.2
158	22.9	23.8
315	28.6	29.1

As shown in Table 2, the difference between the Poisson and the Gaussian for the 4 largest means, $\mu > 30$, is at most 1.0%. Furthermore, as μ increases, the difference between the Poisson and the Gaussian decreases steadily by 0.10% until for the largest mean, the Gaussian is only 0.50% away from the Poisson distribution. The Poisson distribution is still, evidently, a consistently better approximation.

Therefore, the Gaussian distribution can be used instead of the Poisson distribution with a confidence level of 100-1=99.0% for datasets in figures 11-14 where $\mu > 30$.

Looking deeper at the decay process, we know that for a "source containing N nuclei, that there is a certain probability p, related to a decay constant λ, that a particular nucleus will decay in a period Δt of measurement."[1] Thus a decay rate, R, can be found for the source and is given by:

$$R = \lambda \frac{C}{(1 - e^{-\lambda \Delta t})} \sum_{i=1}^{j} k_i \tag{4}$$

where k_i is a correction related to the uncertainty associated with detection.

Table 4: Uncertainty budget for R as expressed in (4)

Input variable into (4)	Uncertainty budget
λ	In previous labs, gamma spectroscopy was used to detect counts over time and from the decay rate, the decay constant was determined. Therefore $u(\lambda)$ is a propagation of C and other uncertainties and carries a Type B uncertainty. There is a statistical element in fitting photopeaks, however that isn't directly related to λ.
C	We have just used Poisson statistics to analyse the decay counts, therefore it has a Type A uncertainty.
k_i	Since k is a discrete, pre-determined correction, it has Type B evaluation of uncertainty since there is no statistical variation in the value of k.
Δt	Since Δt can be sampled by the Poisson distribution, it has a Type A uncertainty.

The equation that combines all the uncertainties in the budget is the following standard uncertainty:

$$u(R) = R\sqrt{\left(\frac{\Delta\lambda}{\lambda}\right)^2 + \left(\frac{\Delta C}{C}\right)^2 + \left(\frac{\Delta(\Delta t)}{\Delta t}\right)^2 + j\left(\frac{\Delta k}{k}\right)^2} \tag{5}$$

5 Conclusions

In conclusion, Poisson statistics provided a successful interpretation of sample MCS data by correctly achieving the shape and frequency for the sample data often within one standard uncertainty for smaller means. Poisson statistics approximates the mean of the sample using its variance with at least 94.2% confidence in the worst case and for large data sets, the gaussian can be used with accuracy which is, for all intents and purposes, equivalent to that of the Poisson distribution.

6 References

[1] Buffler, A., Poisson counting statistics: experiment, PHY3004W laboratory, University of Cape Town (2024).