Performing Statistical Models on Cs-137's Gamma Ray Emissions

Joonwoo Noh(260973555), Leo Kim(260914383)

McGill University Department of Physics

February 10, 2025

Abstract

Using Python and Arduino, we collected data on the number of gamma rays detected by a Geiger counter and applied statistical tests to compare the observed distributions with theoretical models: the Poisson and Gaussian distributions. A chi-squared χ^2 test was performed to evaluate the goodness-of-fit between the experimental data and these theoretical distributions. Our results indicate that as the reliability of the data increases, the observed distribution aligns well with the Poisson model, whereas the Gaussian distribution fails to provide an adequate fit. Additionally, we analyzed the time intervals between successive detection events, which are theoretically expected to follow an exponential distribution. The chi-squared test confirmed that our experimental interval data were consistent with this expectation.

Contents

1 Introduction

When each measurement is independent and the average rate remains constant over time, the probability of obtaining a given number of occurrences follows a Poisson distribution:

$$P(N;\bar{N}) = \frac{e^{-\bar{N}}\bar{N}^N}{N!}, \tag{1}$$

where N is the number of counts, and $\bar{N}$ is the average count given by:

$$\bar{N} = \lambda\tau, \tag{2}$$

where λ is the average count rate, and τ is the total measurement time. Since the Poisson distribution describes discrete events, it is well-suited for modeling radioactive decay. In this experiment, we used Cs-137 as a radioactive source and recorded the number of gamma rays emitted using a Geiger counter.

Furthermore, we can approximate the distribution with a Gaussian function, characterized by the mean $\bar{x}$ and the standard deviation $\bar{\sigma}$:

$$G(x;\bar{x},\sigma) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}, \tag{3}$$

where x is the observed variable. The main difference between these two probability distributions is that the Poisson distribution requires only one parameter $\bar{N}$, while the Gaussian distribution depends on both $\bar{x}$ and σ.

To assess how well our measurements align with these theoretical models, we can apply the reduced chi-squared (χ^2_ν) test:

$$\chi^2_\nu = \frac{1}{\nu}\sum_i \frac{(O_i - E_i)^2}{\sigma_i^2}, \tag{4}$$

where ν is the degrees of freedom, O_i is the observed value, E_i is the expected value, and σ_i is the standard deviation of the observations. This test is referred to as 'reduced' because the standard χ^2 is normalized by ν. A reasonable data set should yield $\chi^2_\nu \approx 1$ and the

corresponding $p - value \approx 0.5$ indicating that there's no reason to reject the null hypothesis and we can accept the fitted parameters of the hypothesized function and their uncertainties with a high degree of confidence [1].

Radioactive decay is also characterized by the time intervals between successive detection events. Since decay events are independent, the waiting time between them follows an exponential distribution:

$$P(t) = \lambda e^{-\lambda t}, \tag{5}$$

where λ is the decay rate, the inverse of the mean time between events. This arises from the memory-less property of Poisson processes, which states that the probability of an event occurring is independent of past events, a fundamental characteristic of spontaneous decay [2]. By analyzing these time intervals, we can verify whether the Geiger counter detections align with the expected exponential behavior. To assess and compare the behavior of our measurements with theoretical models, we applied the χ^2_ν test again.

2 Materials and Methods

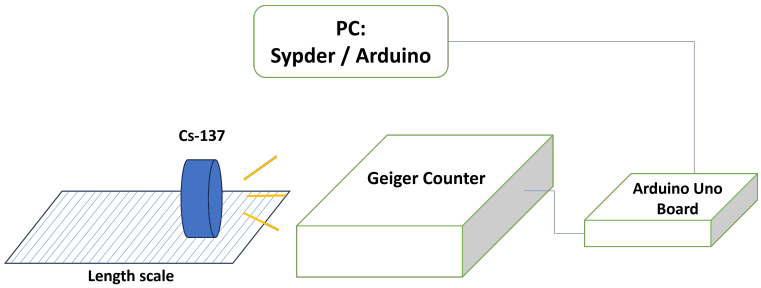


Figure 1: Experiment apparatus setup.

The statistical properties of radioactive decay were investigated using a Geiger counter connected to an Arduino controller and a Python-based data acquisition system. The primary components of the setup included a **Geiger counter** to detect gamma ray, a **Cs-137** as a radioactive source, an **Arduino board** which processed the Geiger counter signals, a **Python script** to receive and analyze data from the Arduino, and a **Millimeter-grid graph paper** for precise placement of the radiation source.

The Geiger counter detects radiation from the source and sends signals to the Arduino accordingly, which records the number of clicks per unit of time or the time interval between successive clicks. The data were then transmitted via serial communication to a Python script that processed and stored them in text files for further statistical analysis.

3 Part 1: Comparing replicas with standard models

3.1 Results

We divided the Geiger counter data into blocks of 100, 200, 400, and 800 counts. For example, with a total of 20,000 data, this results in 200 sets of 100-count blocks, which we refer to as 'replicas'. Figure (2), presents a representative to illustrate its structure.

For every block size, we computed the χ^2_ν values for all replicas using equation (4) and averaged them to assess how well the data align with the Poisson and Gaussian distributions. Additionally, we varied the distance of the Cs-137 from the Geiger counter (8cm, 2cm, and 1cm). As expected, decreasing the source-to-counter distance increased the average count rate per time interval, as summarized in Table (1).

3.2 Discussion

In Figure (2), the histogram becomes smoother as the block size increases. This is because larger data blocks ensure that all bins are adequately populated, reducing statistical fluctuations. Quantitatively, the standard deviations for each block size decreases with increasing block size, as noted in the caption of Figure (2).

Furthermore, as shown in Table (1), for every source distance, the average χ^2_ν for the Gaussian distribution increases with block size, resulting in excessively small p-values far below 0.5. This indicates that the null hypothesis, which assumes our data follow a Gaussian distribution, should be rejected. In contrast, the χ^2_ν values for the Poisson distribution approach 1, leading to reasonable p-values around 0.5. This suggests that our data are well-described by the Poisson distribution, consistent with the expected statistical behaviour of radioactive decay.

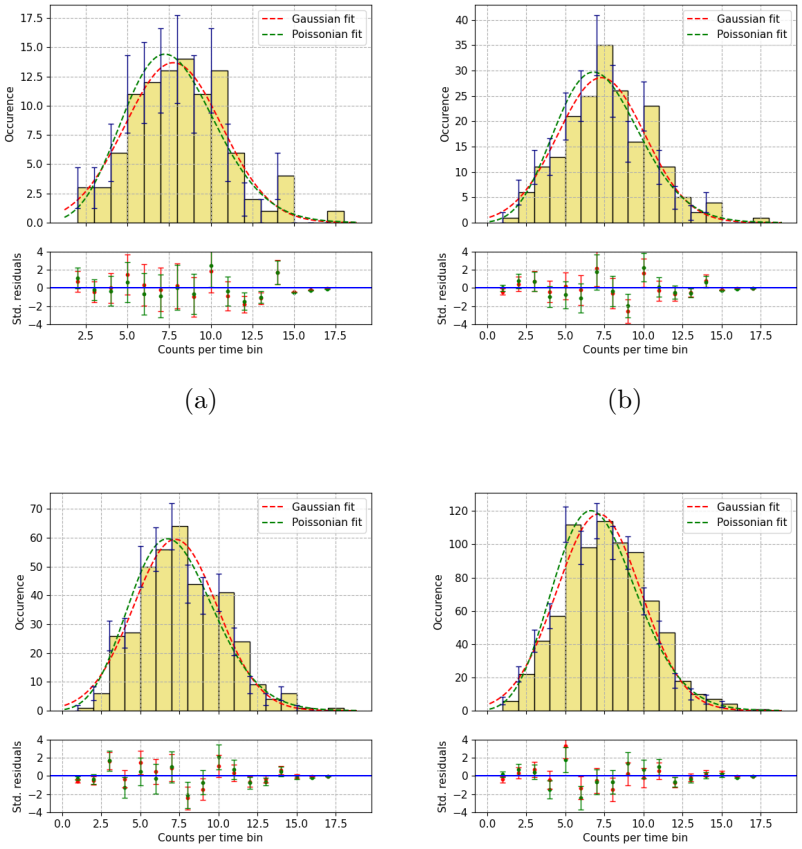


Figure 2: Each panel represents a different block size: (a) 100, (b) 200, (c) 400, and (d) 800, with the source-to-counter fixed at 2 cm. The fitted Gaussian distribution is overlaid in red, and Poisson in green. The average counts per time interval (2 seconds) are 7.8 ± 0.3, 7.3 ± 0.2, 7.2 ± 0.1, and 7.1 ± 0.1, and the standard deviations are 0.29, 0.20, 0.13, and 0.10, respectively.

4 Part 2: Integrating all the replicas

4.1 Results

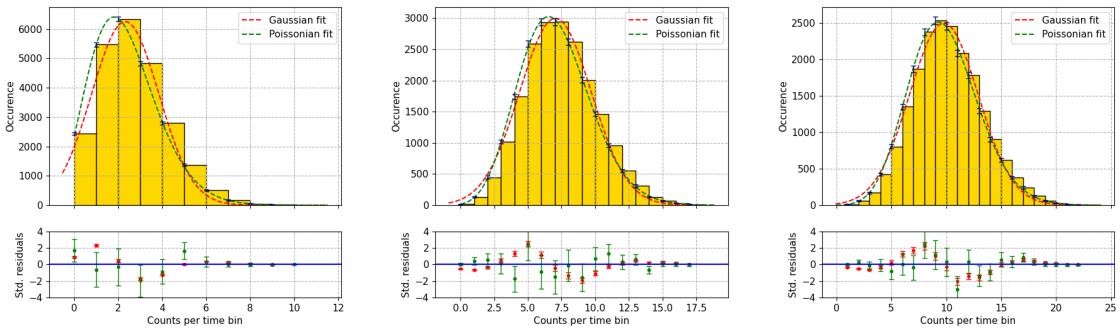
In Part 1, we analyzed individual replicas for different block sizes. To extend this analysis, we summed all replicas by adding the number of occurrences in each corresponding bin across all replicas. Since this integration accumulates all recorded data, the resulting plots remain the same regardless of block size. Therefore, we present only one plot for each source distance in Figure (3).

	Block of 100	**Block of 200**	**Block of 400**	**Block of 800**
Avg. Count Rate per 2s	2.302 ± 0.007	2.302 ± 0.006	2.301 ± 0.006	2.301 ± 0.006
Degrees of Freedom (G)	4	7	7	7
Degrees of Freedom (P)	5	8	8	8
Avg. χ^2_ν (G)	1.76	1.66	2.47	4.22
Avg. χ^2_ν (P)	0.96	0.81	0.78	0.80
p-value (G)	0.133	0.113	0.016	0.000
p-value (P)	0.311	0.484	0.515	0.497
Avg. Count Rate per 2s	7.04 ± 0.02	7.05 ± 0.02	7.04 ± 0.02	7.04 ± 0.02
Degrees of Freedom (G)	14	15	15	15
Degrees of Freedom (P)	15	16	16	16
Avg. χ^2_ν (G)	0.94	1.10	1.36	1.82
Avg. χ^2_ν (P)	0.83	0.93	0.93	0.95
p-value (G)	0.519	0.354	0.155	0.026
p-value (P)	0.572	0.458	0.461	0.434
Avg. Count Rate per 2s	9.78 ± 0.02	9.78 ± 0.02	9.78 ± 0.02	9.77 ± 0.02
Degrees of Freedom (G)	13	15	16	16
Degrees of Freedom (P)	14	16	17	17
Avg. χ^2_ν (G)	1.06	1.04	1.25	1.73
Avg. χ^2_ν (P)	0.95	0.91	0.92	1.03
p-value (G)	0.394	0.409	0.218	0.034
p-value (P)	0.423	0.482	0.474	0.349

Table 1: Statistical analysis results for the source 8cm, 2cm, and 1cm away in order. (G) refers to Gaussian and (P) refers to Poisson. We set our time interval to collect data to be 2 seconds. Degrees of freedom of Gaussian is one smaller than that of Poisson because Gaussian requires two parameters while Poisson requires one.

4.2 Discussion

As shown in Table (2), the χ^2_ν values for the Gaussian distribution are excessively high, leading to very low p-values. This confirms, as in Part 1, that the summed replicas do not follow a Gaussian distribution. In contrast, the χ^2_ν values for the Poisson distribution are slightly below 1 but remain very close, indicating an excellent fit to the data. This could be due to either an overestimation of uncertainties or the Poisson model being an exceptionally good representation of the data. The uncertainty for each bin in the histogram is given by $\sqrt{N}$, where N is bin count. Since the summed replicas have larger N values, the relative uncertainty per bin is lower than in Part 1. This suggests that uncertainties are not



(a) (b) (c)

Figure 3: Summed replicas for 8cm, 2cm, and 1cm in order.

Distance:	8 cm	2 cm	1 cm
Avg. Count per 2s	2.3 ± 0.1	7.0 ± 0.2	9.8 ± 0.2
Standard Deviation	1.53	2.66	3.13
Degrees of Freedom (G)	9	16	20
Degrees of Freedom (P)	10	17	21
χ^2_ν **(G)**	86.4	57.1	21.1
χ^2_ν **(P)**	0.64	0.60	0.75
p-value (G)	1.39×10^{-161}	3.00×10^{-184}	5.62×10^{-77}
p-value (P)	0.782	0.892	0.783

Table 2: Statistical analysis results for summed replicas in different distances.

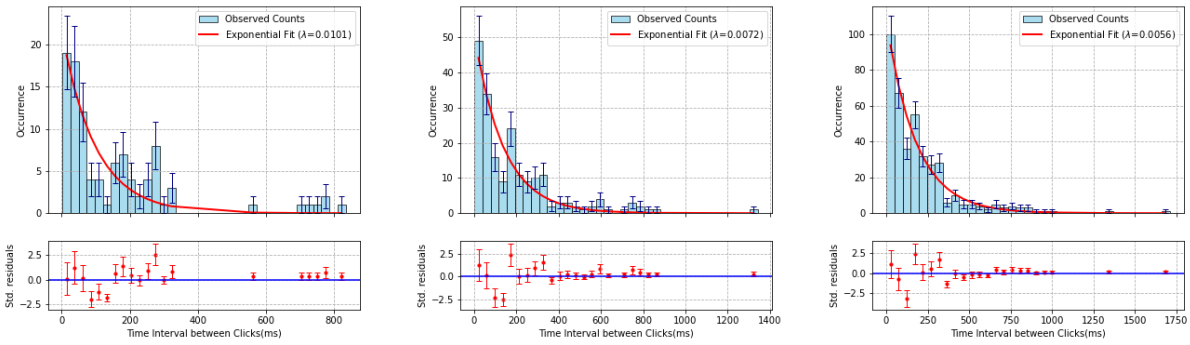
overestimated, but rather the data strongly aligns with the Poisson distribution.

Additionally, Figure (3) shows that as the Cs-137 source moves closer to the Geiger counter, the histogram shifts to the right. This occurs because a shorter distance increases the average count per time interval. Consequently, the fitted Poisson distribution becomes more symmetric and increasingly resembles a Gaussian distribution, as expected.

5 Part 3: Time-resolved Analysis of Gamma Decay

5.1 Results

The time intervals between successive radiation events were analyzed by partitioning the dataset into block sizes of 100, 200, and 400. For each block size, an exponential distribution was fitted to the data, and statistical parameters, including the mean decay rate (λ), reduced



(a) (b) (c)

Figure 4: Measurements of an interval between successive clicks. Blocks of size 100, 200, and 400 in order. Distance between the source and the counter fixed at 2cm.

chi-squared, and p-values, were computed to assess the goodness-of-fit.

For all block sizes, we initially set 35 bins for the histogram. However, for each histogram, bins with zero counts were excluded from chi-squared calculations to prevent division errors in uncertainty estimation. This exclusion reduces the degrees of freedom, but ensures statistical validity by preventing artificially inflated chi-squared values from underpopulated bins. .

The expected λ values were determined through nonlinear least squares fitting in Python, rather than being directly taken from the theoretical decay rate of Cs-137. While the decay rate for an individual Cs-137 nucleus is well-documented ($\approx$ 30 years half-life) [3], the decay rate of the entire radioactive source used in the experiment depends on its total activity, which was not explicitly provided. Therefore, instead of assuming an arbitrary value, we opted to extract λ empirically from the data, allowing us to focus on evaluating whether the detected events follow the expected exponential behavior. This approach ensures that our analysis remains independent of any assumptions about the source strength while still verifying the underlying statistical nature of radioactive decay.

Figure (4) shows the histogram of time intervals for a single replica in each block size along with the fitted exponential distribution. The fits capture the expected decay trend, with deviations at lower time intervals due to fluctuations in bin counts. Standard residual plots beneath each histogram indicate a relatively even scatter around zero, further confirming the adequacy of the fit.

5.2 Discussion

	Block of 100	**Block of 200**	**Block of 400**
Avg. Lambda Parameter (s^{-1})	0.0058 ± 0.0001	0.0058 ± 0.0001	0.0058 ± 0.0001
Degrees of Freedom	20.7 ± 0.3	21.4 ± 0.3	21.8 ± 0.3
Avg. χ^2_ν	0.98 ± 0.04	0.97 ± 0.03	0.99 ± 0.03
p-value	0.49	0.51	0.49

Table 3: Statistical analysis results for the time interval between clicks. The distance between the source and the counter was fixed at 2cm.

Table (3) summarizes the statistical results for each block size with uncertainties represented by standard errors rather than standard deviations, as they quantify the uncertainty in the mean decay rate over multiple block-size replicas. The measured decay rate remained consistent across block sizes, with $\lambda = 0.0058 \pm 0.0001 s^{-1}$. The degrees of freedom increased slightly with block size from 20.7 ± 0.3 for block size 100 to 21.8 ± 0.3 for block size 400. The reduced chi-squared values remained close to 1, ranging from 0.97 ± 0.03 to 0.99 ± 0.03, indicating a strong agreement between the observed data and the exponential model. The p-values remained around 0.49 to 0.51, suggesting that the null hypothesis that the data follows an exponential distribution cannot be rejected at any block size used.

Uncertainties arise from systematic error and statistical errors. Systematic errors include the timing resolution of the Arduino and potential geometric biases in radiation detection due to the source not behaving as a perfect point emitter. Statistical errors stem from variations in sample sizes within histogram bins, which influence degrees of freedom and chi-squared calculations. However, the small uncertainties in both χ^2_ν and λ suggest that statistical fluctuations are well-controlled. Since increasing block size does not significantly affect any of the statistical values, the model already describes the data well at smaller block sizes. A larger data set could further improve precision.

Overall, the statistical results and visual agreement between observed data and the expected model support the assumption that radioactive decay follows an exponential time distribution. The consistency of the reduced chi-squared values and p-values across different block sizes reinforces the robustness of this statistical model.

6 Conclusions

Throughout the analysis, we successfully rejected the null hypothesis that the number of counts from radioactive decay follows a Gaussian distribution, while failing to reject the hypothesis that it follows a Poisson distribution. This conclusion was supported by χ^2 tests: as the block size increases, χ^2_ν for the Gaussian distribution grew unreasonably large, whereas χ^2_ν for the Poisson distribution approached 1. Since a null hypothesis can only be rejected or fail to be rejected, we cannot definitively conclude that our data follows the Poisson distribution, but the results strongly suggests a high likelihood.

Additionally, we confirmed that the time intervals between successive radioactive decays follow an exponential distribution, independent of block size. The fitted decay rate λ remained stable, with χ^2 values near 1 and p-values close to 0.5, indicating a good fit. This consistency suggests that the memory-less property of radioactive decay holds across different statistical samplings. In conclusion, we failed to reject the null hypothesis, reinforcing that radioactive decays follow an exponential probability distribution.

References

[1] I. Hughes and T. Hase, "Measurements and their uncertainties: A practical guide to modern error analysis," vol. 1, pp. 24, 28, 29, 2010. 2

[2] H. Pishro-Nik, *Introduction to Probability, Statistics, and Random Processes.* Kappa Research, LLC, 2014. 2

[3] K. Lavelle *et al.*, "A reference material for evaluation of 137 cs radio chronometric measurements." *J Radioanal Nucl Chem.*, vol. 318, pp. 195–208, 2018. 7