Poisson Statistics

Jason Gross*
Student at MIT
(Dated: October 9, 2011)

I present the theory behind the use of Poisson statistics; Poisson statistics arise from the counting of random, statistically independent processes, examples of which abound in nature. We investigate background radiation counts with a scintillation counter, and we do not reject the null hypothesis that they are Poisson-distributed. On the other hand, we reject the hypothesis that they are normally distributed for $\mu \leq 10$ with confidence $> 99\%$.

I. THE THEORY OF INDEPENDENT RANDOM VARIABLES AND POISSON STATISTICS

I.1. Random Variables

It is commonly assumed, at least implicitly, that every measurement taken is a measurement of a random variable. Sometimes, the set of possible values is discrete; for example, we may look at a light and observe it to be "on". Other times, the set of possible values seems continuous; for example, we may measure the width of this sheet of paper in centimeters.

Regardless of what the details of a measurement are, we assume that our measurements are repeatable in the limit, and that our particular set of measurements constitute a representative sample of all the set of all such possible measurements. If, for example, you have a small field of apple trees, you might estimate the number of apples that you have by counting how many trees you have, counting how many apples are on some particular tree, and multiplying your two counts. In science, we assume that such methods of estimation are reasonable.

More formally, we assume that every sequence of measurements is a *random variable*; this amounts, I think, to assuming that frequentist definition of probability is self-consistent and that the Law of Large Numbers holds for every sequence of measurements.[1]

The assumption that every measurement, both in theory and in practice, is of a random variable, allows us to talk meaningfully and systematically about models for measurement, which in turn allow us to reason about models for the underlying universe, and how experimental results correspond to these models. It remains to establish that the displays of our measurement devices are actually measurements of the underlying quantities. To me, this seems a remarkable fact, which I hypothesize to be a result of the central limit theorem and the fact that variances add in quadrature. A complete theoretical analysis of the nature of measurement is either far beyond the scope of this paper or far beyond my understanding (or both).

I.2. Poisson Statistics: A Limit of the Binomial Distribution

Consider a measurement, such as flipping a coin, with two possible outcomes, with probabilities p and $q := 1-p$; call the outcomes "success" and "failure", respectively. If you take n such measurements, all statistically independent, then the probability distribution on the number of successes x is

$$P(x; n, p) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}. \qquad (1)$$

Straightforward calculation yields that the mean of this distribution is $\mu = np$.

Consider now a measurement, such as the number of radioactive decays in a minute of an atom with a particular half-life, that is believed to occur with binomial distribution with known mean μ and an unknown but large number of potential events ($n \gg \mu$; or $p \ll 1$). The distribution on the measured number of successes of such a measurement is given by

$$P(x; \mu) = \frac{\mu^x}{x!} e^{-\mu}. \qquad (2)$$

The mean and deviation of this distribution are both equal to μ. A full derivation and discussion of this distribution can be found in an introductory statistics textbooks such as Bevington [3].

* jgross@mit.edu

[1] The law of large numbers states that the sample average of identical statistically independent random measurements almost always converges to the expected value of a given measurement. The formal definition of "almost always" is non-trivial and beyond the scope of this paper. Any introductory formal statistics or probability theory textbook should explain it; alternatively, see MathWorld[1] or Wikipedia[2].

This may be combined with the frequentist definition of probability to note that the proportion of measurements in which we measure a fixed particular outcome is itself a random variable subject to the law of large numbers, and so the law of large numbers permits us to assert that the average of distributions of measurements will match the probability distribution in the large n limit.

II. MOTIVATION AND METHODOLOGY

II.1. Motivation

Poisson statistics are everywhere. Examples of Poisson-distributed measurements include the rate of radioactive emissions of a radioisotope, the number of drops of water per minute that fall in to a bucket of water during a rainstorm, and the number of particles per unit area in an approximately uniform beam of particles.

We set out to observe a particular example of Poisson statistics—the level of background radiation—in the lab. We tested the hypothesis that the background γ-ray count, at various energy levels (not measured), is well-described by a Poisson distribution.

II.2. Methodology

We measured the level of background radiation (in particular, γ-rays) using a scintillation counter. A scintillation counter is composed of a scintillator and a photomultiplier. The scintillator is a material that fluoresces (releases photons) when struck with charged particles. The entrance of the photomultiplier is metallic; the metal releases electrons when struck by light, via the photoelectric effect. The electrons are then accelerated to progressively higher voltage dynodes, which are made of materials that emit many electrons when struck with [few] electrons. This cascade of electron emissions creates a signal that can be measured, or amplified by a traditional amplifier. The signal is approximately proportional to the energy of incoming radiation.

III. PROCEDURE

As shown in Figure 1, we hooked the NaI scintillation counter to a high voltage source, and to a counter through a preamplifier and amplifier. The counter was coupled with a [built-in] discriminator which dropped signals below a particular voltage.

We wanted to measure the distribution at mean per-second counts of approximately 1, 4, 10, and 100 γ-rays; we set the amplifier and discriminator to the following settings:

μ	Amplifier (ratio)	Discriminator (V)
≈ 1	26.10 ± 0.01	9.7 ± 0.2
≈ 4	26.10 ± 0.01	6.0 ± 0.2
≈ 10	46.10 ± 0.01	9.7 ± 0.2
≈ 100	46.10 ± 0.01	2.0 ± 0.2

For each setting, we recorded the number of events in 100 one-second intervals, and then in one 100 second interval. The timer on the counter was determined to be accurate to one part in a million with a pulse generator.[2]

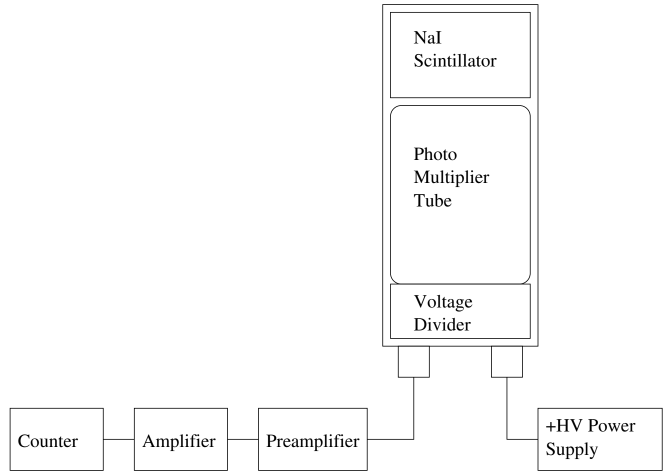


FIG. 1. The setup for measuring the number of incident γ-rays of background radiation in a given time interval. Figure shamelessly stolen (and slightly modified) from the lab guide.[4]

IV. RESULTS AND CONCLUSIONS

IV.1. Analysis and Results

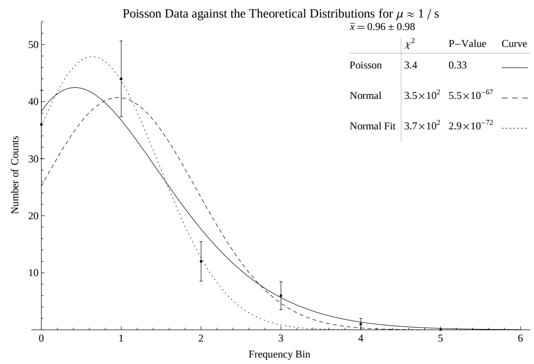
We found that background γ-ray radiation was Poisson distributed.

The data for the one-second trials were plotted against the [continuous] theoretical Poisson and normal distributions in Figure 2, using the average of the data as the mean and the square root of that value as the standard deviation. Plots of the best-fit normal curves are also included for $\mu \approx 1$ and $\mu \approx 4$;[3] no significant differences were found between best-fit curves and curves based only on the averages, and the visual differences between best-fit Poisson curves and average-based Poisson curves was not worth showing. Cumulative average plots are included in Figure 3 to show that the error is inversely proportional to the square root of the number of measurements. The Pearson χ^2 test was performed to find goodness-of-fit. I do not reject the null hypothesis that γ-ray counts are Poisson distributed. I reject the null hypothesis that γ-ray counts are normally distributed for $\mu \leq 10\,{}^1/\mathrm{s}$, with a confidence of $> 99\%$.

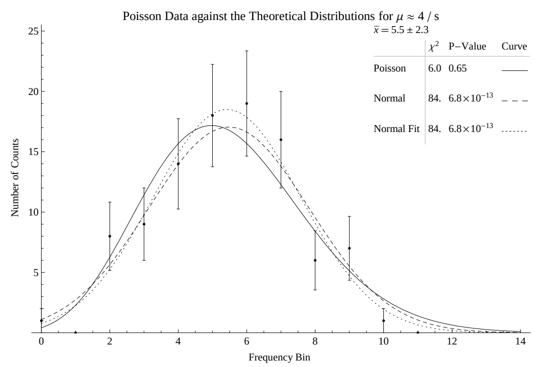
---

[2] We also found a systematic offset of 12 µs, but I suspect this was due to a confusing setting of our pulse generator (called `ToPer`) which we did not discover until the electromagnetic pulses lab.

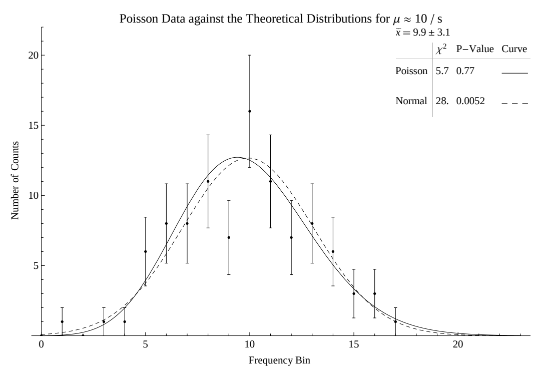
[3] I was having trouble with the larger values of μ, presumably because Mathematica decided that χ^2 had no gradient at around $\mu \approx 0$, where I'm guessing it started. I minimized the χ^2 values myself, though, and found no appreciable differences.



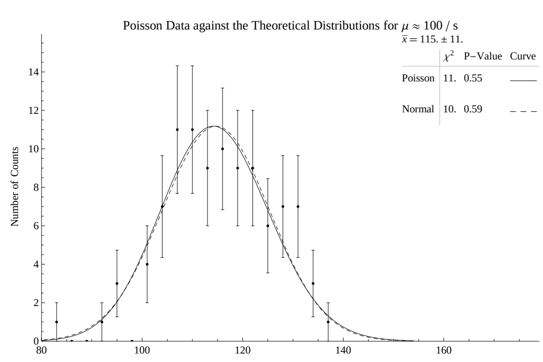
(a)



(b)

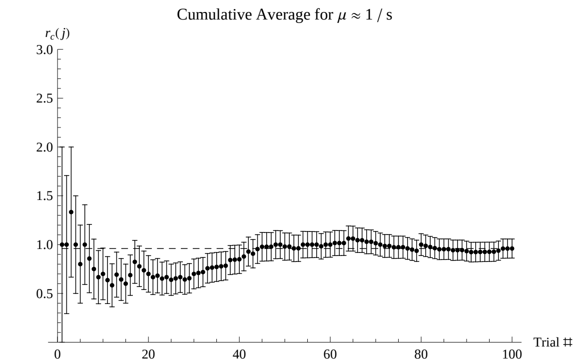


(c)

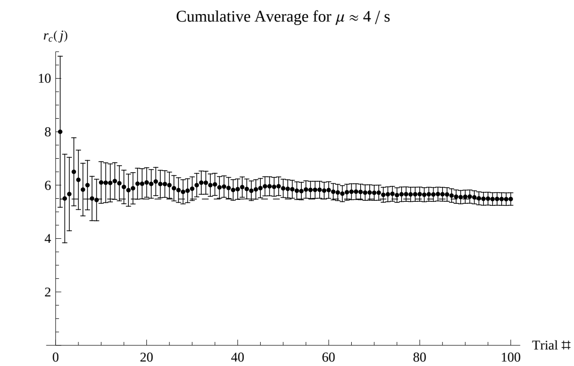


(d)Unlike the other plots, a bin count of 3 was used here.

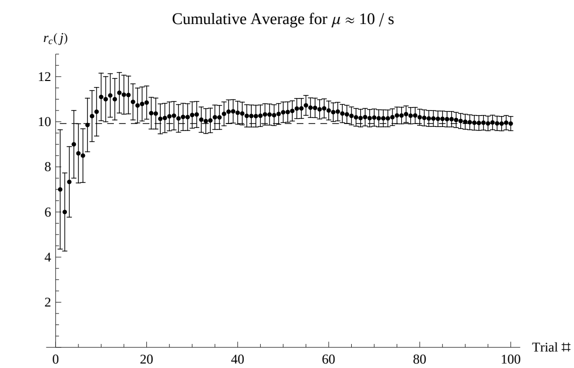
FIG. 2. Plots of experimental data against the theoretical curves. The fitted normal curves have two degrees of freedom, and the other theoretical curves have one.



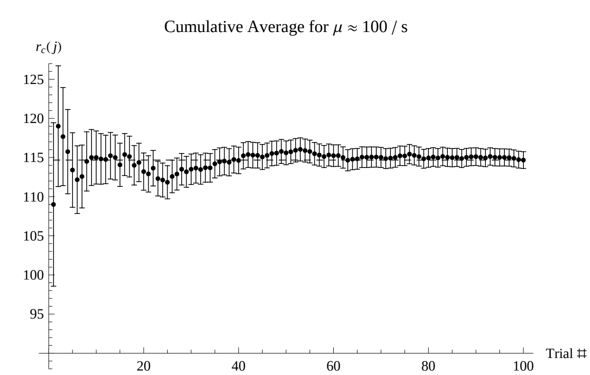
(a)



(b)



(c)



(d)

FIG. 3. Plots of the cumulative average with error bars.

IV.2. Discussion and Conclusion

Unsurprisingly, background γ-ray radiation counts at the energies that we probed were found to be Poisson distributed. This supports our hypothesis that radioactive decays are random, statistically independent processes.[4]

[1] J. Renze and E. W. Weisstein, "Law of large numbers," http://mathworld.wolfram.com/LawofLargeNumbers.html.

[2] "Law of large numbers," http://en.wikipedia.org/wiki/Law_of_large_numbers (2011).

[3] P. Bevington and D. Robinson, *Data Reduction and Error Analysis for the Physical Sciences* (McGraw-Hill, 2003).

[4] S. P. Robinson, "Poisson statistics," (2011).

ACKNOWLEDGMENTS

The author gratefully acknowledges his lab partner Thomas Vandermeulen for help collecting the data, and the J-Lab staff for their help in experimentation.

[4] Actually, it supports the hypothesis that background γ radiation comes from random, statistically independent processes. There is an additional hypothesis connecting these two hypotheses, which is roughly that there is no process that alters the character of the distribution of γ-rays between emission and detection.