Statistical Analysis of Data

Exponential Decay and Poisson Processes

Poisson processes

A *Poisson process* is one in which occurrences are randomly distributed in time, space or some other variable with the number of occurrences in any non-overlapping intervals statistically independent. For example, naturally occurring gamma rays detected in a scintillation detector are randomly distributed in time, or chocolate chips in a cookie dough are randomly distributed in volume. For simplicity, we will limit our discussion to occurrences or "events" randomly distributed in time.

A *homogeneous* Poisson process is one in which the long-term average event rate is constant. The average rate will be denoted Γ and in any interval Δt the expected number of events is

$$\mu = \Gamma \Delta t \tag{1}$$

A *nonhomogeneous* Poisson process is one in which the average rate of events changes and so might be expressed as some function $\Gamma(t)$ of time. The number of events expected in an interval from t_1 to t_2 would then be the integral

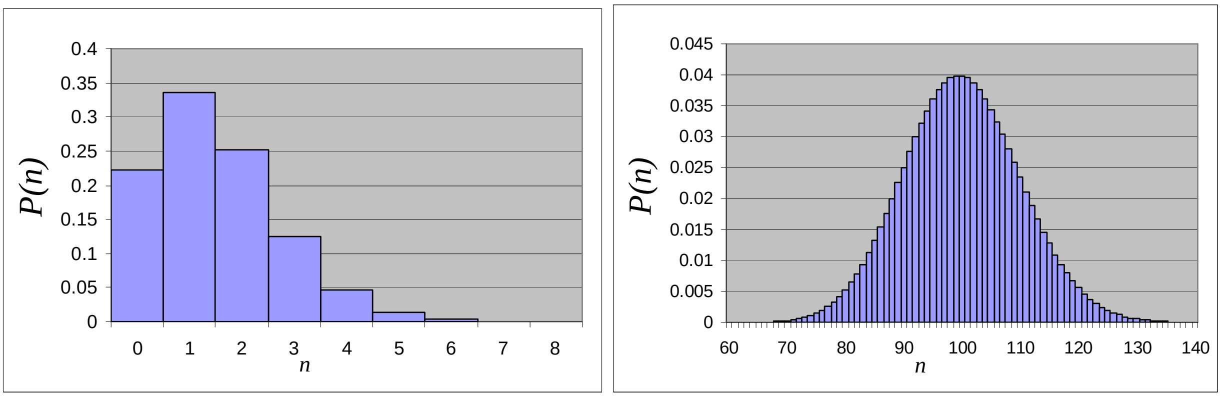
$$\mu = \int_{t_1}^{t_2} \Gamma(t) dt \tag{2}$$

While the expected number of events μ for a given experiment need not be an integer, the number of events n actually observed must be. Moreover, due to the randomness of the events, n may be more or less than μ with the probability for a given n depending only on μ and given by

$$P(n) = e^{-\mu} \frac{\mu^n}{n!} \tag{3}$$

This is the Poisson distribution, a discrete probability distribution, with each $P(n)$ giving the probability for that particular n to occur.

Two Poisson distributions for $\mu = 1.5$ and $\mu = 100$ are shown in Fig. 1. This figure shows the parent distributions. Real sample distributions would be expected to vary somewhat from the parent, getting closer to the parent as the sample size N increases.



(a) μ = 1.5 (b) μ = 100

Figure 1: Poisson probabilities for means of 1.5 and 100

The exponential probability density function

A better way of describing Γ is as a probability per unit time that an event will occur. That is

$$dP = \Gamma dt \tag{4}$$

where dP is the differential probability that an event will occur in the infinitesimal time interval dt. Of course, some care must be taken when translating a rate to a probability per unit time. For example, if $\Gamma = 10/\text{s}$, it is obviously not true that the probability is 10 that an event will occur in any particular second. However, if that same rate is expressed $\Gamma = 0.01/\text{ms}$ it is roughly true that the probability is 0.01 that an event will happen in any particular millisecond. Eq. 4 only becomes exact in the limit of infinitesimal dt.

Equation 4 also fundamentally describes the decay process of an excited state of an atom, nuclei, or subatomic particle. In these cases, $dP = \Gamma dt$ is the probability for the excited state to decay in the *next* time interval dt and Γ is called the decay rate for the excited state rather than an event rate.

Equation 4 can be shown to lead directly to the Poisson probability distribution. The first step is to see how it leads to the exponential probability density function (pdf) giving the probability $dP_e(t)$ that the *next* Poisson event (or the decay of an excited state) will occur in the interval from t to $t + dt$.[1]

If the probability of no event (or survival of the excited state) to a time t is denoted $P(0; t)$, then the probability of no event (or survival) to $t + dt$ would be the product of this probability with the probability of no event (or no decay) in the interval dt following t. Since the probability of an event (or decay) in this interval is Γdt, the probability of no event (or no

[1] Equation 4 is equivalent to $dP_e(0) = \Gamma dt$ and must be the $t = 0$ limiting case for the general solution.

decay) in this interval is $1 - \Gamma dt$ and thus:

$$P(0; t + dt) = P(0; t)(1 - \Gamma dt) \tag{5}$$

Rearranging and substituting $(P(0; t + dt) - P(0; t))/dt = dP(0; t)/dt$ gives

$$\frac{dP(0; t)}{dt} = -\Gamma P(0; t) \tag{6}$$

which has the general solution $P(0; t) = Ae^{-\Gamma t}$. Because we must start with no event (or no decay) at $t = 0$, $P(0; 0) = 1$ and so $A = 1$ giving

$$P(0; t) = e^{-\Gamma t} \tag{7}$$

Then, the differential probability $dP_e(t)$ for the next event (or decay) to occur in the interval from t to $t + dt$ is given by the probability of no event (or no decay) in the interval from 0 to t followed by an event (or a decay) in the next interval dt. The former has a probability $P(0; t) = e^{-\Gamma t}$ and the later has a probability Γdt. Thus

$$dP_e(t) = \Gamma e^{-\Gamma t} dt \tag{8}$$

Equations 8 is a continuous probability density function (pdf). It is properly normalized, i.e., the integral over all times from 0 to ∞ is unity as required. It also has the very reasonable property that the expectation value for the random variable t—the time to the next event (or to the decay)—is given by

$$\begin{aligned} \langle t \rangle &= \int_0^\infty t\Gamma e^{-\Gamma t} dt \\ &= \frac{1}{\Gamma} \end{aligned} \tag{9}$$

In the case of decay, the expectation value $\langle t \rangle$, henceforth denoted τ_e, is called the *lifetime* of the excited state. Thus, Γ and τ_e are equivalent ways to quantify the decay process. If the decay rate is 1000/s, the lifetime is 0.001 s. Moreover, Eq. 8 is often expressed in terms of the lifetime rather than the decay rate.

$$dP_e(t) = \frac{1}{\tau_e} e^{-t/\tau_e} dt \tag{10}$$

The probability for decay in a time τ_e is found by integrating Eq. 8 from 0 to τ_e and gives the value $1/e$. Thus, for a large sample of excited states at $t = 0$, the fraction $1/e$ of them will have decayed by τ_e. The time it would take for half the sample to decay is called the halflife $\tau_{1/2}$ and is easily shown to be $\tau_e \ln 2$.

The Poisson probability distribution

There are several possible derivations of the Poisson probability distribution. It is often derived as a limiting case of the binomial probability distribution. The derivation to follow relies on Eq. 4 and begins with Eq. 7 for the probability $P(0;t)$ that there will be no events in some finite interval t.

Next, a recursion relation is derived for the probability, denoted $P(n+1;t)$, for there to be $n+1$ events in a time t, which will be based on the probability $P(n;t)$ of one less event. For there to be $n+1$ events in t, three independent events must happen in the following order (their probabilities given in parentheses).

- There must be n events up to some point t' in the interval from 0 to t ($P(n,t')$ by definition).
- An event must occur in the infinitesimal interval from t' to $t'+dt'$ ($\Gamma dt'$ by Eq. 4).
- There must be no events in the interval from t' to t ($P(0,t-t')$ by definition).

The probability of $n+1$ events in the interval from 0 to t would be the product of the three probabilities above integrated over all t' from 0 to t to take into account that the last event may occur at any time in the interval. That is,

$$P(n+1;t) = \int_0^t P(n;t')\Gamma dt' P(0;t-t') \tag{11}$$

From Eq. 7 we already have $P(0;t-t') = e^{-\Gamma(t-t')}$ and substituting the following definition:

$$P(n;t) = e^{-\Gamma t}\overline{P}(n;t) \tag{12}$$

Eq. 11 becomes (after canceling $e^{-\Gamma t}$ from both sides):

$$\overline{P}(n+1;t) = \Gamma \int_0^t \overline{P}(n;t')dt' \tag{13}$$

From Eqs. 7 and 12, $\overline{P}(0;t) = 1$ and then $\overline{P}(1;t)$ can be found from an application of Eq. 13

$$\begin{aligned}\overline{P}(1,t) &= \Gamma \int_0^t \overline{P}(0,t')dt' \\ &= \Gamma \int_0^t dt' \\ &= \Gamma t \end{aligned} \tag{14}$$

Applying Eq. 13 for the next few terms

$$\begin{aligned}\overline{P}(2,t) &= \Gamma \int_0^t \overline{P}(1;t')dt' \\ &= \Gamma \int_0^t \Gamma t' dt' \\ &= \frac{\Gamma^2 t^2}{2} \end{aligned} \tag{15}$$

$$\begin{aligned}\overline{P}(3,t) &= \Gamma\int_0^t \overline{P}(2;t')dt' \\ &= \Gamma\int_0^t \frac{\Gamma^2 t'^2}{2}dt' \\ &= \frac{\Gamma^3 t^3}{2\cdot 3}\end{aligned} \tag{16}$$

The pattern clearly emerges that

$$\overline{P}(n;t) = \frac{(\Gamma t)^n}{n!} \tag{17}$$

And thus with Eq. 12, the Poisson probabilities result

$$P(n;t) = e^{-\Gamma t}\frac{(\Gamma t)^n}{n!} \tag{18}$$

Note that, as expected, the right side depends only on the combination $\mu = \Gamma t$, allowing us to write

$$P(n) = e^{-\mu}\frac{\mu^n}{n!} \tag{19}$$

where the implicit dependence on t (or μ) has been dropped from the notation on the left hand side of the equation.

Although the Poisson probabilities were derived assuming a homogeneous process, they are also correct for nonhomogeneous processes with the appropriate value of μ (Eq. 2).

Several important properties of the Poisson distribution are easily investigated. For example, the normalization condition — that some value of n from 0 to infinity must occur — translates to

$$\sum_{n=0}^{\infty} P(n) = 1 \tag{20}$$

and is easily verified from the series expansion of the exponential function

$$e^{\mu} = \sum_{n=0}^{\infty}\frac{\mu^n}{n!} \tag{21}$$

The expected number of events is found from the following weighted average

$$\langle n\rangle = \sum_{n=0}^{\infty} nP(n) \tag{22}$$

and is evaluated as follows:

$$\begin{aligned}\langle n\rangle &= \sum_{n=1}^{\infty} ne^{-\mu}\frac{\mu^n}{n!} \\ &= \mu\sum_{n=1}^{\infty} e^{-\mu}\frac{\mu^{n-1}}{(n-1)!} \\ &= \mu\sum_{m=0}^{\infty} e^{-\mu}\frac{\mu^m}{m!} \\ &= \mu\end{aligned} \tag{23}$$

In the first line, the explicit form for $P(n)$ is used and the first term $n = 0$ is explicitly dropped as it does not contribute to the sum. In the second line, the numerator's n is canceled with the one in the denominator's $n!$ and one μ is also factored out in front of the sum. In the third line, $m = n - 1$ is substituted, forcing a change in the indexing from $n = 1...\infty$ to $m = 0...\infty$. And in the last line, the normalization property is used.

Thus, the Poisson probability distribution gives the required result that the expectation value (or parent average) of n is equal to μ.

We should also want to investigate the standard deviation of the Poisson distribution which would be evaluated from the expectation value

$$\begin{aligned} \sigma^2 &= \langle (n-\mu)^2 \rangle \\ &= \langle n^2 \rangle - \mu^2 \end{aligned} \tag{24}$$

The expectation value $\langle n^2 \rangle$ is evaluated as follows:

$$\begin{aligned} \langle n^2 \rangle &= \sum_{n=0}^{\infty} n^2 P(n) \\ &= \sum_{n=1}^{\infty} n^2 e^{-\mu} \frac{\mu^n}{n!} \\ &= \mu \sum_{n=1}^{\infty} n e^{-\mu} \frac{\mu^{n-1}}{(n-1)!} \\ &= \mu \sum_{m=0}^{\infty} (m+1) e^{-\mu} \frac{\mu^m}{m!} \\ &= \mu \left[\sum_{m=0}^{\infty} m e^{-\mu} \frac{\mu^m}{m!} + \sum_{m=0}^{\infty} e^{-\mu} \frac{\mu^n}{n!} \right] \\ &= \mu \left[\mu + 1 \right] \\ &= \mu^2 + \mu \end{aligned} \tag{25}$$

In the second line, the form for $P(n)$ is substituted and the first term $n = 0$ is dropped from the sum as it does not contribute. In the third line, one power of μ is factored out of the sum and one n is canceled against one in the $n!$. The indexing and lower limit of the sum is adjusted in the fourth line using $m = n - 1$. In the fifth line, the $m + 1$ term is separated into two terms, which are are evaluated separately in the sixth line; the first term is just $\langle m \rangle$ and evaluates to μ by Eq. 23 and the second term evaluates to 1 based on the normalization condition—Eq. 20.

Now combining Eq. 25 with Eq. 24 gives the result

$$\sigma^2 = \mu \tag{26}$$

implying that the parent variance is equal to the mean, i.e., that the standard deviation is given by $\sigma = \sqrt{\mu}$.

Counting statistics

When μ is large enough (greater than 10 or so), the Poisson probability for a given n

$$P_p(n) = e^{-\mu}\frac{\mu^n}{n!} \tag{27}$$

is very nearly that of a Gaussian pdf having the same mean μ and standard deviation $\sigma = \sqrt{\mu}$ integrated over an interval of $\pm 1/2$ about that value of n.

$$P_g(n) = \int_{n-1/2}^{n+1/2} \frac{1}{\sqrt{2\pi\mu}} \exp\left(-\frac{(x-\mu)^2}{2\mu}\right) dx \tag{28}$$

To a similar accuracy the integrand can be taken as constant over the narrow interval giving the simpler result[2]

$$P_g(n) = \frac{1}{\sqrt{2\pi\mu}} \exp\left(-\frac{(n-\mu)^2}{2\mu}\right) \tag{29}$$

Typically, the true mean μ (of the Poisson distribution from which n is a sample) is unknown and, based on the principle of maximum likelihood, the measured value n is taken as a best estimate of this quantity.[3] The standard deviation of the distribution from which n is a sample is, by convention, the uncertainty in this estimate. This standard deviation is $\sqrt{\mu}$ and if the true μ is unknown and n is used in its place, the best estimate of this uncertainty would be $\sqrt{n}$.

The large-n behavior of the Poisson probabilities is the basis for what is typically called "square root statistics" or "counting statistics." That is, that a measured count obtained from a counting experiment can be considered to be a sample from a Gaussian distribution with a standard deviation equal to the square root of that count.

Exponential decay

There are two experiments in our laboratory investigating decay processes. In Experiment GA, excited ^{137}Ba nuclei are monitored while they decay with the emission of a gamma ray. It is the decreasing number of such gamma rays with time that is measured and compared with the prediction of Eq. 8. For a sample of N_0 nuclei at $t = 0$, the predicted number decaying in the interval from t to $t + dt$ is given by $dN(t) = N_0 dP_e(t) = N_0 \Gamma e^{-\Gamma t} dt$. If the gamma rays emitted during the decay are detected with a probability ϵ, the number of detected gamma rays $dG(t)$ in the interval from t to $t + dt$ is predicted to be

$$dG(t) = \epsilon N_0 \Gamma e^{-\Gamma t} dt \tag{30}$$

---

[2] It would be interesting to study how such different looking functions as Eqs. 27 and 29 can agree so well. Their near equivalence (for $\mu = n$) leads to Stirling's formula, $n! \sim \sqrt{2\pi} e^{-n} n^{n+1/2}$.

[3] It is simple to show that using $\mu = n$ will maximize the probability $P_p(n)$.

Background gamma rays and detector noise pulses are expected to occur with a constant average rate λ. Taking these into account gives

$$dG(t) = (\epsilon N_0 \Gamma e^{-\Gamma t} + \lambda) dt. \tag{31}$$

The quantity in parentheses is a nonhomogeneous Poisson rate. The lifetime for ^{137}Ba is around 3.5 minutes and in the experiment, gamma rays are counted in 15-second intervals throughout a 30 minutes period. The analysis consists of fitting the number of gammas detected in each interval to the integral of Eq. 31 over the appropriate time interval and taking into account dead time corrections.

In Experiment MU, cosmic ray muons occasionally stop in a scintillation detector and, with a lifetime of a few μs, decay into an electron and two neutrinos.[4] After an initial electronic discrimination step, the detector produces identical logic pulses with various efficiencies (or probabilities) for various processes. We will distinguish pulses arising under three different conditions:

Capture pulses are produced when a muon stops in the detector.

Decay pulses are produced when the muon decays in the detector.

Non-capture pulses are produced from the passage of a muon through the detector, from other natural background radiation, and from other processes such as detector noise.

Keep in mind that the pulses are identical. Their origin is distinguished for theoretical purposes only—in order to build a model for their relative timing. In fact, we will further distinguish between events in which both the capture and the decay process produce a detector pulse and those in which only one of the two produces a detector pulse. Events in which two pulses are produced will be called capture/decay events and associated variables will be subscripted with a c. Pulses from muon events in which only one pulse occurs are indistinguishable from and can be grouped with the non-capture pulses, all of which will be called non-paired pulses and their associated variables will be subscripted with an n.

Capture/decay pulse pairs in our apparatus are rare—occurring at a rate $R_c \approx 0.01$/s (one pulse pair every hundred seconds or so). Non-paired pulses occur at a much higher rate $R_n \approx 10 - 100$/s.

These pulses result in one of two possible experimental outcomes.

Doubles are events in which two pulses follow in rapid succession—within a short timeout period of 20 μs or so for the muon decay experiment.

Singles are all pulses that are not doubles.

[4]Our theoretical model will assume a single scintillator is used rather than four scintillators as in the actual experiment.

The measurement and analysis of time intervals between the pulses of doubles are used to "discover" the muon decays. To measure the interval, a high speed clock is started on any detector pulse and it is stopped on any second pulse occurring within a 20 μs timeout period. If a second pulse does not occur within the timeout, the clock is rearmed and ready for another start. If a second pulse occurs within the timeout, the measured time interval is saved to a computer. Most starts are not stopped within the timeout. These are the singles.

After every start pulse, the apparatus is "dead" to another start pulse until a stop pulse or the timeout occurs. This *dead time* leads to difference between the true rate R at which particular pulses occur and the lower rate R' at which they would occur as start pulses. (Dead time is relatively unimportant for the muon decay measurements, but the prime symbol will be added to any rate when it is the rate at which start pulses occur.)

Even when a non-paired pulse starts the timer, a double may result if a second pulse—by random chance—just happens to occur before the timeout period. Doubles having a non-paired start pulse are considered *accidentals* because their time distribution can be predicted based on the probability that two unrelated pulses just happen to follow one another closely in time. This case is discussed next.

Non-paired start pulses can be stopped by either another non-paired pulse or the capture pulse of a capture/decay pair. The latter would be relatively rare because capture/decay pairs are rare, but are included for completeness. Non-paired pulses and capture/decay pulse pairs are homogeneous Poisson events and occur at a combined rate $R_n + R_c$. Consequently, the probability for the next pulse to occur between t and $t + dt$ is given by the exponential distribution, Eq. 8, for this combined rate.

$$dP_{sn}(t) = (R_n + R_c)e^{-(R_n+R_c)t}dt \tag{32}$$

The rate of non-capture pulses stopped in the interval from t to $t + dt$ is then the product of the rate of non-paired start pulses R'_n and the probability dP_{sn}

$$dR_{sn}(t) = R'_n(R_n + R_c)e^{-(R_n+R_c)t}dt \tag{33}$$

Start pulses arising from the capture pulse of capture/decay pairs are considered next. The theoretical muon decay model is that the decay (and its pulse) occurs with a probability per unit time equal to the muon decay rate Γ. Again, for completeness we should also consider the possibility that the stop pulse will be from a non-paired pulse (less likely, with a probability per unit time R_n) or by the capture event of a different capture/decay pair (even less likely, with a probability per unit time R_c). All three of these are Poisson processes and whichever one comes first will stop the clock. The net probability per unit time for any of the three to occur is their sum $\Gamma + R_n + R_c$ and thus the probability dP_{sc} that the stop pulse will occur between t and $t + dt$ is given by

$$dP_{sc}(t) = (\Gamma + R_n + R_c)e^{-(\Gamma+R_n+R_c)t}dt \tag{34}$$

The rate of stopped capture events in the interval from t to $t + dt$ is then the product of the rate of capture/decay start pulses R'_c times the probability $dP_{sc}(t)$

$$dR_{sc}(t) = R'_c(\Gamma + R_n + R_c)e^{-(\Gamma+R_n+R_c)t}dt \tag{35}$$

Finally, the total rate dR_s of events with stops between t and $t+dt$ is the sum of Eqs. 33 and 35.

$$dR_s(t) = \left[R'_n(R_n + R_c)e^{-(R_n+R_c)t} + R'_c(\Gamma + R_n + R_c)e^{-(\Gamma+R_n+R_c)t}\right] dt \tag{36}$$

The electronics sort each stop time t into bins of uniform size τ, which can be considered to be the period of a high speed clock. The clock is started on a start event, stopped on a stop event, and the number of clock ticks between these events determines which bin the stop event is sorted into. A stop occurring in bin 0 (before one clock tick) would correspond to t between 0 and one clock period τ. A stop occurring in bin 1 (after 1 clock tick has passed) would correspond to t between τ and 2τ, etc.

Thus, the differential rate $dR_s(t)$ becomes a finite rate by integration over one clock period. And the rate R_i of stop events in bin i becomes

$$R_i = \int_{i\tau}^{(i+1)\tau} \left[R'_n(R_n + R_c)e^{-(R_n+R_c)t} + R'_c(\Gamma + R_n + R_c)e^{-(\Gamma+R_n+R_c)t}\right] dt \tag{37}$$

The bin size or clock period in the muon decay experiment is 20 ns and small enough that the integrand above does not change significantly over an integration period. Taking the integrand as constant at its value at the midpoint of the interval gives

$$R_i = \left[R'_n(R_n + R_c)e^{-(R_n+R_c)t_i} + R'_c(\Gamma + R_n + R_c)e^{-(\Gamma+R_n+R_c)t_i}\right] \tau \tag{38}$$

where $t_i = (i + 1/2)\tau$ is the midtime for the interval

Of course, the histogram bins continue filling according to the rate R_i and how long one collects data. Thus, the product of the rate and the data collection time Δt

$$\mu_i = R_i \Delta t \tag{39}$$

is the expected number of counts in bin i. Keep in mind that the bin filling process is a homogeneous Poisson process and the actual counts occurring in bin i will be a Poisson random variable for the mean μ_i.

For the muon experiment, there are at least two orders of magnitude between each of the rates:

$$\Gamma \gg R_n \gg R_c \tag{40}$$

Thus, to better than 1% only the largest need be kept when several are added together. Moreover, the first exponential term, which decays at the rate $R_n \approx 10 - 100$/s stays very nearly constant throughout the 20 μs timeout period. Consequently, the dependence of μ_i on t_i is very nearly given by

$$\mu_i = \alpha_1 + \alpha_2 e^{-\Gamma t_i} \tag{41}$$

This equation will be useful in fitting the accumulated muon data to determine Γ.

Fitting with Poisson random variables

Recall that the normal fitting procedure is to choose parameters of a fitting function $F(x_i)$ that will minimize the chi-square:

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - F(x_i))^2}{\sigma_i^2} \tag{42}$$

where y_i is the measured value for the point i and $F(x_i)$ is the fitted value for that point. This least squares principle is based on the principle of maximum likelihood and the assumption that each y_i is a random variable from a Gaussian distribution of mean $F(x_i)$ and standard deviation σ_i.

For a Poisson-distributed event-counting experiment, the data set is represented $\{n_i\}$ or $n_i, i = 1...N$ where each n_i is the measured number of events for bin i. In the fit, μ_i would be the equivalent of the $F(x_i)$[5] and n_i would be the equivalent of y_i. Assuming the validity of "counting statistics," $\sqrt{n_i}$ would be the standard deviation (the equivalent of σ_i) and the probability of the entire data set would be

$$P_g(\{n_i\}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi n_i}} \exp\left(\frac{-(n_i - \mu_i)^2}{2n_i}\right) \tag{43}$$

The fitting parameters appearing in μ_i would then be determined by maximizing $P_g(\{n_i\})$, or equivalently, by minimizing the chi-square

$$\chi^2 = \sum_{i=1}^{N} \frac{(n_i - \mu_i)^2}{n_i} \tag{44}$$

Counting statistics are not expected to be valid for low values of μ (or n). In cases where μ_i is expected to be small for many points in the data set, extra care must be taken to avoid errors in the fitting procedure. For example, if μ_i is around 2-3, roughly 5-15% of the time n_i will be zero. Using counting statistics and setting $\sigma_i = \sqrt{n_i} = 0$ is obviously going to cause problems with the fit (divide by zero error). In fact, using $\sigma_i = \sqrt{n_i}$ for any bins with just a few events can cause systematic errors in the fitting parameters. In such cases, a modified chi-square method could be employed using the fitting function μ_i instead of n for σ_i^2 in the chi-square denominator.

$$\chi^2 = \sum_{i=1}^{N} \frac{(n_i - \mu_i)^2}{\mu_i} \tag{45}$$

For this chi-square, special care would need to be taken to ensure that the fitting function μ_i remains non-zero and positive throughout the fitting procedure. Using Eq. 45 can lessen the systematic error, but it is not the best approach.

[5] For example, Eq. 41 for the muon decay experiment.

The best approach to handling data sets with low counts is to return to the method of maximum likelihood, using the Poisson distribution to describe the probability of each bin. The probability of the entire data set $\{n_i\}$ is then given by

$$P_p(\{n_i\}) = \prod_{i=1}^{N} e^{-\mu_i} \frac{\mu_i^{n_i}}{n_i!} \tag{46}$$

Noting that the chi-square of Eq. 44 can be written as:

$$\chi^2_{\text{Gauss}} = -2 \ln P_g(\{n_i\}) + C \tag{47}$$

where C is a constant, we define an effective chi-square statistic from the Poisson likelihood in an analogous manner

$$\begin{aligned} \chi^2_{\text{Poisson}} &= -2 \ln P_p(\{n_i\}) \\ &= 2 \sum_{i=1}^{N} (\mu_i - n_i \ln \mu_i + \ln n_i!) \end{aligned} \tag{48}$$

Then, the process of maximizing the Poisson likelihood is equivalent to minimizing the Poisson chi-square function above.

We can simplify the Poisson chi-square expression even more. Since only μ_i will change during the minimization procedure and not the data points n_i, the last term in the sum can be dropped and we need only minimize

$$\chi^2_{\text{Poisson}} = 2 \sum_{i=1}^{N} (\mu_i - n_i \ln \mu_i) \tag{49}$$

Moreover, because this effective chi-square for Poisson-distributed data was constructed in analogy to the Gaussian chi-square (aside from a constant offset), the uncertainty in any fitted parameter can be determined in a manner analogous to the method used for Gaussian-distributed data. That is, while holding a single fitting parameter fixed and slightly offset from its optimized value, all the other parameters are then reoptimized. The amount the fixed parameter must be changed to cause the Poisson chi-square to increase by 1 is that parameter's uncertainty.

Because of the constant offsets, the Poisson chi-square cannot be used for evaluating the goodness of the fit. Goodness of fit can be checked in a subsequent step by forming the reduced chi-square statistic

$$\chi^2_\nu = \frac{1}{N-M} \sum_{i=1}^{N} \frac{(n_i - \mu_i)^2}{\mu_i} \tag{50}$$

where M is the number of fitting parameters. If the data follow the prediction for μ_i, this statistic should occur with probabilities governed by the standard reduced chi-square variable with $N - M$ degrees of freedom.

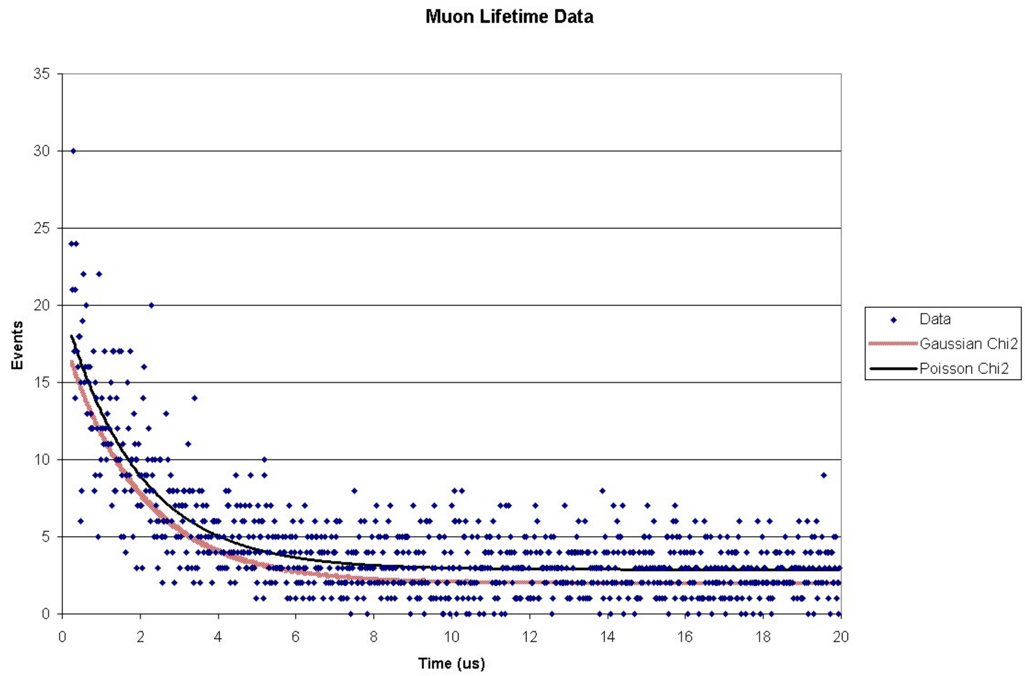


Figure 2: Example data and fit for muon lifetime measurements.

Figure 2 shows fitting results for a 2-day run of the muon decay experiment. Each data point is the number of events n_i (on the vertical axis) for which the measured time between pulses of a "double" was in a 20 ns wide interval starting at t_i (on the horizontal axis). Two fits were performed to the hypothesis of Eq. 41. One fit uses the modified chi-square Eq. 45, and the other uses the Poisson chi-square Eq. 49. For this 2-day run, both fits give a lifetime of $1/\Gamma = 1.95 \pm 0.1\,\mu$s, which compares well to the known lifetime of 2.2 μs. But when data from only one day is used and the number of bins with 0 or 1 entries increases, the Gaussian method shows a bias toward lower lifetimes (typically 1.7 ms), whereas the Poisson method is unaffected except for a larger statistical error. Also, notice in the figure that the two methods disagree on the size of the constant background even for 2 days of running: 2 events per bin for the Gaussian method, and 2.9 events per bin for the Poisson method. As a check, the background can be accurately estimated by taking the average of n_i number of events for time intervals longer than 10 ms, where the contribution from the exponential is negligible. The result is 2.9 entries per bin, in agreement with the Poisson method.