Counting statistics of random events: A tutorial

This tutorial presents derivations of some results from the theory of the statistics of random events that are studied in the muon counting experiment. Specifically, it uses two basic rules of probability to derive a formula for the rate of random coincidences between two pulse trains and gives a derivation of the probability density function of time intervals between n counts in a random pulse train.

Introduction: random pulse trains

Many experiments in the physics 433 laboratory are built around counting pulses generated by random events. Such events may be produced by radioactive decay of a nearby nuclide, such as a ^{22}Na gamma-ray source, or they may come from the passage of cosmic rays—mainly muons—through a scintillator paddle. In counting experiments, the detector pulses which typically have a variable height (maximum voltage) are sent to a discriminator which responds only to those pulses above a particular threshold; the output of the discriminator is a sequence of regular-shaped "digital" pulses of fixed voltage height and duration or "width".

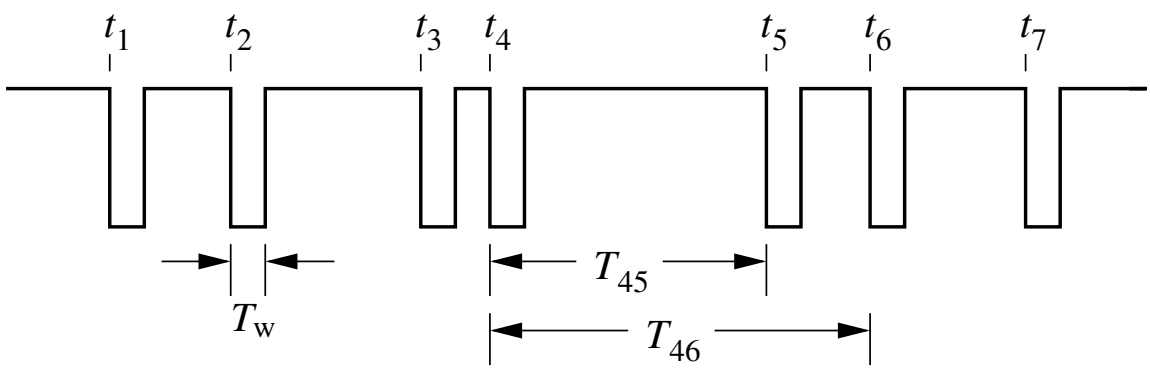

Figure 1: A random pulse train. This example shows negative-going ("fast NIM") pulses typical of those used in nuclear counting experiments.

An example of such a signal, what we will call a "random pulse train", is shown in Fig. 1 as it might appear on an oscilloscope screen. The pulse width T_w is important only insofar as it determines the maximum rate of pulses that may be represented by the pulse train, since pulses which occur more frequently than $1/T_w$ cannot be resolved. The occurrence time of each pulse, denoted as t_1, t_2, etc., as measured from some (arbitrary) starting time, is marked at the leading edge of each pulse.

If the pulse train in question is being created by a long-lived radioactive nuclide or by cosmic rays, experiments show that although the time between specific pulses T_{ij} is completely unpredictable, the time intervals do converge to a well-defined average τ in the following sense: If one counts very many pulses (i.e., thousands) N over a long time interval T, then $\tau \approx T/N$. More specifically, if one estimates τ by counting N pulses over number of trials, the variation of the calculated τ's from these trials will cluster about each other with a fractional standard deviation of $1/\sqrt{N}$. (Later we

will strengthen this assertion.) We can define an average *rate* of pulses r by the relation $r = 1/\tau$.

We are interested in two problems:

1. If two independent pulse trains are compared, what is the likelihood that a pulse from each train will occur simultaneously? This is the problem of *coincidence*.
2. What is the *distribution of intervals* between successive pulses in a given pulse train? More generally, what is the distribution of *scaled intervals*: intervals between two pulses separated by an arbitrary number of pulses in between?

We will answer these questions by making a basic assumption and invoking two rules. The basic assumption is

> Any infinitesimally small interval of time dt is equally likely to contain a pulse. The probability of finding a pulse in dt is given by $r\,dt$.

The two rules are from the theory of probabilities [1]. Assume that two events of a similar type A and B follow from the same proximate cause. The probabilities of these events $P(A)$ and $P(B)$ are subject to these two rules:

> **I.** *The rule of compound probabilities.* If the two events A and B are mutually exclusive, then the probability of having either event $P(A\,\text{or}\,B)$ is given by the sum $P(A)+P(B)$.
>
> **II.** *The rule of independent probabilities.* If the two events A and B are independent of each other, then the probability of having both events occur $P(A\,\text{and}\,B)$ is given by the product $P(A)\times P(B)$.

Simple examples of these two rules are illustrated by considering rolls of dice. A die has six numbered sides, and when it is thrown any one number will be equally likely to land face up: the probability of any number is 1/6. If we ask, "What is the probability to roll an even number?", we are invoking rule **I**, since the roll of any number excludes the roll of any other number. Thus,

$$P(\text{even}) = P(2) + P(4) + P(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}\ .$$

If we ask, "What is the probability to roll two 1s in a row?", we are invoking rule **II**, since each successive roll of the die is independent of any previous roll. Thus

$$P(1\,\&\,1) = P(1)\times P(1) = \frac{1}{6}\times\frac{1}{6} = \frac{1}{36}\ .$$

A bit more complex example that will be relevant to our task is illustrated by this question, "What is the probability of rolling only one 1 in two rolls of the die?" This is asking for the probability of a 1 on the first roll *and* not 1 on the second *or* a 1 on the second roll *and* not on the first. We can write this formally by rules **I** and **II** as

$$\begin{aligned} P(\text{only one 1 in 2 rolls}) &= P(1;1\text{st})\times P^*(1;2\text{nd}) \;+\; P^*(1;1\text{st})\times P(1;2\text{nd}) \\ &= \frac{1}{6}\times\frac{5}{6} \;+\; \frac{5}{6}\times\frac{1}{6} \\ &= \frac{10}{36}\ , \end{aligned}$$

where P^* is the compliment of P, that is, it is the probability of *not* obtaining the specified outcome. This result agrees with the simple reckoning which says that out of 36 possible combinations there are five ways to get a 1 on the first roll but not on the second, and five ways to get a 1 on the second roll but not on the first, or a total of 10/36 possibilities that give the result.

Before turning to our two main problems, we will use these rules to answer a preliminary question: Given a random pulse train of average pulse rate r, what is the probability of finding *no pulses* in an interval T, $P_0(T)$? Our basic assumption concerns the probability of finding at least one pulse in an infitesimal interval dt, so we will approach this problem by a somewhat circuitous path. Let $P_0^*(T)$ be the compliment of $P_0(T)$, that is, let it be probability of finding a pulse in T. Consider the following construction. Let the interval T be subdivided into a series of infinitesimal intervals of length dt, and note that by our assumption a pulse could occur in any one of these intervals. If we define $dP_0^*(t)$ as the probability of finding a pulse in the interval dt contained between t and $t + dt$ under the condition that *there were no pulses up to time* t, then by rule **II**

$$dP_0^*(t) = \begin{pmatrix} \text{Probability of finding} \\ \text{zero pulses between 0} \\ \text{and } t \end{pmatrix} \times \begin{pmatrix} \text{Probability of finding} \\ \text{a pulse between } t \text{ and} \\ t + dt \end{pmatrix} ,$$

because the probability of finding no pulses in a previous interval is independent of the probability of finding a pulse in a subsequent interval. Symbolically, our construction reads

$$dP_0^*(t) = P_0(t) \times r\, dt \ .$$

This differential relates P_0^* to P_0, and right now we don't know either. But, since any interval must either contain *no* pulses or *one or more* pulses, we have by rule **I**, $P_0(T) + P_0^*(T) = 1$. If we take differentials of this, we get $dP_0^*(t) = -dP_0(t)$; in other words, as the differential probability of finding a pulse increases, the differential probability of not finding a pulse decreases by the same amount. So

$$dP_0(t) = -P_0(t) r\, dt \ . \tag{1}$$

Moreover, by rule **I**, to find the probability that no pulses occur in the large interval T, we must *add up* the differential probabilities $dP_0(t)$ pertaining to each subsequent differential dt (because we want the probability of no pulse in the first dt interval *or* the second dt interval *or* the third dt interval, etc.). In other words, we integrate Eq. (1):

$$\int_0^{P_1(T)} \frac{dP_1(t)}{P_1(t)} = -\int_0^T r\, dt \ ,$$

which works out to

$$P_0(T) = e^{-rT} \ . \tag{2}$$

This gives for the probability of finding one (or more) pulses in the interval T

$$P_0^*(T) = 1 - e^{-rT} \ . \tag{3}$$

With these results, we are ready to attack our two problems.

Rate of random coincidences

The use of coincidence counting is pervasive in elementary particle experiments for one big reason: noise. For example, if you set up a scintillator paddle to count the passage of cosmic-ray muons,

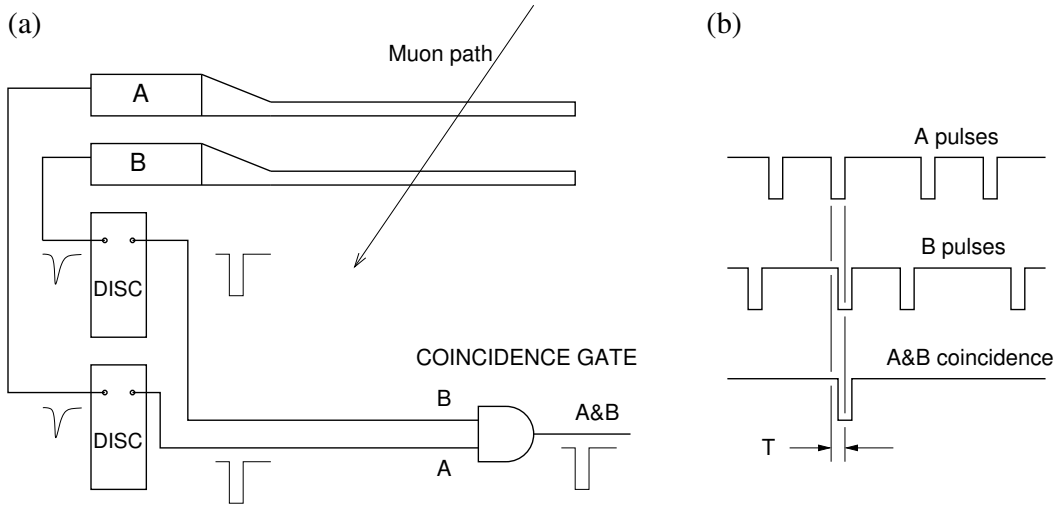

Figure 2: (a) A two-paddle coincidence setup. Variable height detector pulses are discriminated and turned into digital pulses. The coincidence gate produces a pulse when two input pulses overlap within a gate time T. (b) Example of two pulse trains showing one coincidence.

most of the counts you record will be due to non-muon events such as electrical noise, low-level background radiation, etc. Fortunately, muons have fairly high energy, so that if they pass through a thin scintillator paddle they will continue out the other side and can be detected by a second paddle nearby. Such an event produces two pulses, one in each detector, at (nearly) the same time. By using a logic gate, one can force the counting electronics to record only those pulses which occur in coincidence. A schematic of this setup is shown in Fig. 2. In order for the coincidence technique to work profitably, the rate of random coincidences must be appreciably lower than the rate of "true" coincidences. We can use our results to predict this rate.

Let the rate of random pulses recorded by detector A be be r_A and the rate of random pulses recorded by detector B be r_B. Each time a pulse from either A or B enters the gate, the gate is triggered: it "opens" for a short time interval T called the *resolving time*, meaning that if a pulse from the other detector occurs during T, then a coincidence pulse is produced at the output. Since r_A and r_B are the average rates of random pulses, the rate of coincidences will also be random (in the absence of any true coincidence events). Let R be this rate; more specifically, we let $R\,dt$ be the probability that a random coincidence event will occur within the infinitesimal time interval dt.

By rules **I** and **II** $R\,dt$ is given by the following construction:

$$R\,dt = \begin{pmatrix} \text{Probability of the gate} \\ \text{being triggered by} \\ \text{detector } A \text{ in } dt. \end{pmatrix} \times \begin{pmatrix} \text{Probability of detector} \\ B \text{ delivering a pulse} \\ \text{within resolving time} \\ T \text{ of the trigger time.} \end{pmatrix}$$

$$+ \begin{pmatrix} \text{Probability of the gate} \\ \text{being triggered by} \\ \text{detector } B \text{ in } dt. \end{pmatrix} \times \begin{pmatrix} \text{Probability of detector} \\ A \text{ delivering a pulse} \\ \text{within resolving time} \\ T \text{ of the trigger time.} \end{pmatrix} .$$

In symbolic form, this equation reads

$$R\,dt = r_A\,dt \times P_0^*(T;B) + r_B\,dt \times P_0^*(T;A)\ .$$

Typically the resolving time T is much shorter than the average rates r_A and r_B so that $rT \ll 1$. In this case, the exponential in Eq. (3) may be expanded to first order:

$$P_0^*(T) = 1 - e^{-rT} \approx 1 - (1 - rT) = rT \ .$$

Thus

$$\begin{aligned} R\,dt &= r_A\,dt \times r_B T + r_B\,dt \times r_A T \\ &= 2 r_A r_B T\,dt \ . \end{aligned}$$

So the rate of random coincidences R is given by

$$R = 2 r_A r_B T \ . \tag{4}$$

Exercise 1 *Show that if you added a third paddle C to the setup, the rate of random coincidences would be equal to $3 r_A r_B r_C T^2$.*

Distribution of time intervals

Before we turn to the theoretical problem of predicting a distribution of time intervals, let's look at the experimental situation. Look again at Fig. 1, which represents a portion of a measurable pulse train. The time interval between the fourth and fifth pulse is marked T_{45}, and other intervals could be calculated according to $T_{ij} = t_j - t_i$. Consider the question, "What is the experimental distribution of time intervals in a given pulse train?" What this means is this: if we were to record a long pulse train and make a list of all of the time intervals between successive pulses $T_{i,i+1}$, how many of these would be, say, between 0 and 1 microseconds long, how many would be between 1 and 2 microseconds long, how many would be between 2 and 3 microseconds long, and so forth. In this example, the "bin width" Δt is one microsecond. We can define an experimental distribution function $N_1(t; \Delta t)$ as a function that gives the number of intervals in our pulse train which fall between time t and time $t + \Delta t$. The subscript "1" indicates that the distribution refers to 1-pulse intervals, i.e., the intervals between pulse i and $i + 1$. We could also calculate the distributions of 2-pulse intervals, 3-pulse intervals, and so forth by going through our list and adding together the 1st and 2nd interval, the 2nd and 3rd interval, etc., to get $N_2(t; \Delta t)$, or adding the 1st, 2nd, and 3rd intervals, the 2nd, 3rd, and 4th intervals, etc., to get $N_3(t; \Delta t)$.

The operations just mentioned are precisely what the LabVIEW program that collects the data in the muon counting experiment does. Each time a pulse is sensed by the digital counter card, the time between it and the previous pulse is recorded in a list. Then the list is passed to a routine that sums successive intervals according to the requested number n of pulses per interval and creates a new list of n-pulse intervals. This list is passed through a histogram program to sort the list into a distribution. Instead of defining a preset bin width Δt, the program calculates Δt from the longest recorded interval $T_{\max}$ in the set and the specified number of bins to use: $\Delta t = T_{\max}/(\text{no. of bins})$.

A couple of points should be clear upon inspection. First, the distributions $N_n(t; \Delta t)$ as defined are *not normalized*, since they will depend on the total number of pulses in the pulse train being examined. Second, there is a trade-off between the bin width Δt and the value of $N_n(t; \Delta t)$, since bins of wider width will collect more intervals per bin than bins of narrower width (for a given data set). One can resolve the normalization problem by using the facts that the total number of intervals in the list is equal to the sum of the numbers of intervals in each bin (regardless of the

bin width) and the fact that the *area* of the histogram "curve" is equal to the bin width times the total number of intervals in the list. Mathematically, we can write these facts as

$$\sum_i N_n(t_i; \Delta t) = M_n \ ,$$

and

$$\frac{\sum_i N_n(t_i; \Delta t)\Delta t}{M_n \, \Delta t} = 1 \ ,$$

where the sum is over the total number of bins in the histogram and M_n is the total number of n-pulse intervals in the list. The normalization factor $M_n \, \Delta t$ will have to be applied to the theoretical distribution in order to compare it to the experimental one.

Theoretical n-pulse distribution

The theoretical distribution function is continuous and normalized; it is a *probability density* function which we will define as follows: $I_n(t)\, dt$ is the probability of finding an n-pulse interval of length between t and $t + dt$. The expectation is that for long enough experimental runs $N_n(t; \Delta t) \approx I_n(t) M_n \, \Delta t$.

The expression $I_1(t)\, dt$ describes the probability of finding a pulse during the infinitesimal time interval dt after a time interval of length t during which there have been no pulses, following any particular pulse in the pulse train. This situation is illustrated by Fig. 3. If we pick a pulse in the

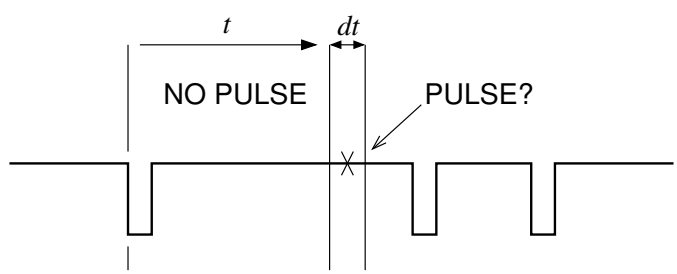

Figure 3: The construction leading to $I_1(t)\, dt$ asks how likely is a pulse to occur in dt after there have been no pulses after a previous pulse.

train, and assign its time as $t = 0$, then we obtain $I_1(t)$ by the following construction:

$$\begin{aligned} I_1(t)\, dt &= \left(\begin{array}{l} \text{Probability of finding} \\ \text{zero pulses between 0} \\ \text{and } t \end{array} \right) \times \left(\begin{array}{l} \text{Probability of finding} \\ \text{a pulse between } t \text{ and} \\ t + dt \end{array} \right) \\ &= P_0(t) \times r\, dt \\ &= re^{-rt}\, dt \ . \end{aligned}$$

Thus

$$I_1(t) = re^{-rt} \ . \tag{5}$$

This result looks familiar—it is the same expression as we gave for $dP_0^*(t)$, which is not surprising since it was derived from similar considerations.

Exercise 2 *Show that $I_1(t)$ is normalized, that is, prove $\int_0^\infty I_1(t)\, dt = 1$.*

Exercise 3 *Show that the average 1-pulse interval τ_1 is equal to $1/r$. The average interval is given by $\int_0^\infty tI_1(t)\,dt$.*

The situation for the distribution of 2-pulse intervals is a bit more complicated. What we want to know is the likelihood of finding an interval of a particular length between two pulses that contains a single pulse anywhere inside. One such possibility is shown in Fig. 1, where the 2-pulse interval T_{46} is shown by a single pulse occurring at t_5 between the pulses at t_4 and t_6. The (differential) probability of finding an interval like T_{46} in a pulse train is given by rule **II**:

$$
\begin{aligned}
dP(T_{46}) &= \\
&\left(\begin{array}{l}\text{Probability of}\\ \text{finding 0}\\ \text{pulses between}\\ t_4 \text{ and } t_5\end{array}\right) \times \left(\begin{array}{l}\text{Probability of}\\ \text{finding a pulse}\\ \text{between } t_5 \text{ and}\\ t_5 + dt_5\end{array}\right) \times \left(\begin{array}{l}\text{Probability of}\\ \text{finding 0}\\ \text{pulses between}\\ t_5 \text{ and } t_6\end{array}\right) \times \left(\begin{array}{l}\text{Probability of}\\ \text{finding a pulse}\\ \text{between } t_6 \text{ and}\\ t_6 + dt_6\end{array}\right) \\
&= P_0(t_5 - t_4) \times r\,dt_5 \times P_0(t_6 - t_5) \times r\,dt_6 \ .
\end{aligned}
$$

This construction is nothing more than the differential probability of finding two 1-pulse intervals of particular lengths. In other words, We can write the differential probability of finding a *particular* two pulse interval of length t with one pulse occurring at t' where $0 < t' < t$ as

$$dP_2(t;t') = I_1(t')\,dt' \times I_1(t-t')\,dt \ .$$

However, what we want is the differential probability of a two pulse interval where the intermediate pulse happens anywhere in the interval. Thus by rule **I**, we must sum the individual probabilities over all values of t' between 0 and t. Hence, we see that

$$
\begin{aligned}
I_2(t)\,dt &= \left(\int_0^t I_1(t')I_1(t-t')\,dt'\right)\,dt \\
&= \left(\int_0^t re^{-rt'}\,re^{-r(t-t')}\,dt'\right)\,dt \\
&= \left(r^2e^{rt}\int_0^t dt'\right)\,dt \\
&= r(rt)e^{-rt}\,dt \ .
\end{aligned}
$$

So the distribution function for 2-pulse intervals is

$$I_2(t) = r(rt)e^{-rt} \ . \tag{6}$$

What about 3-pulse intervals? We can follow the same prescription, namely, we find the differential probability of finding a 2-pulse interval of length t' followed by a 1-pulse interval of length $t - t'$, and then sum over all t' between 0 and t:

$$
\begin{aligned}
I_3(t)\,dt &= \left(\int_0^t I_2(t')I_1(t-t')\,dt'\right)\,dt \\
&= \left(\int_0^t r(rt')e^{-rt'}\,re^{-r(t-t')}\,dt'\right)\,dt \\
&= \left(r^2e^{rt}\int_0^t (rt')dt'\right)\,dt \\
&= r\frac{(rt)^2e^{-rt}}{2}\,dt \ .
\end{aligned}
$$

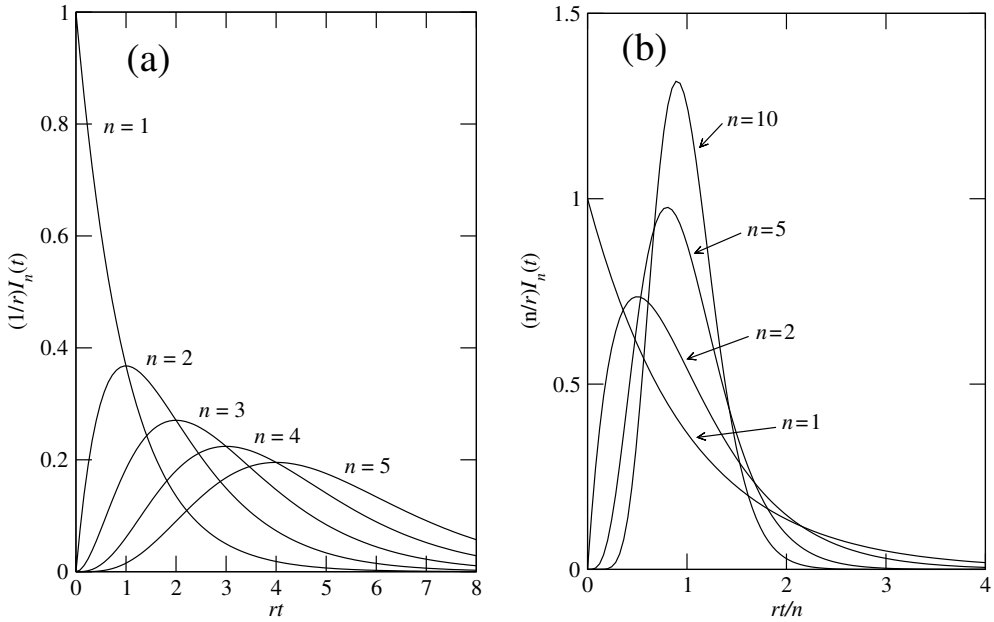

Figure 4: The n-pulse interval density functions, after Knoll [2, p. 98]. (a) $(1/r)I_n(t)$ for n from 1 to 5. Notice that the functions peak at $rt = n - 1$. (b) $(1/r)I_n(t)$ scaled by the transformations $t \to t/n$ and $(1/r)I_n(t) \to (n/r)I_n(t)$ to show that as n increases, the function becomes more sharply peaked around the mean value $\tau_n = n/r$.

Thus,

$$I_3(t) = r\frac{(rt)^2 e^{-rt}}{2} \ . \tag{7}$$

If we follow this prescription out for calculating the distribution of n-pulse intervals, we will obtain the following formula:

$$I_n(t) = r\frac{(rt)^{(n-1)} e^{-rt}}{(n-1)!} \ . \tag{8}$$

It should be noted that the fraction in Eq. (8) is none other than the Poisson distribution giving the probability of finding $n - 1$ events given that the average number of events is rt. This fact is used in the derivations of the interval distribution function given by Knoll [2] and Melissinos [3]. Most derivations of the Poisson distribution are based on taking a particular limit of the discrete binomial distribution. The derivation given here follows a more direct route from the basic rules of probability as applied to random pulse trains.

Exercise 4 *Convince yourself of the* $(n-1)!$ *factor in Eq. (8) by using* $I_3(t)$ *and* $I_1(t)$ *to show that*

$$I_4(t) = r\frac{(rt)^3 e^{-rt}}{1 \cdot 2 \cdot 3}$$

Exercise 5 *Show that the average n-pulse interval* τ_n *is equal to* n/r*, but that the most probable interval* $T_{n,\text{peak}}$ *is given by* $(n-1)/r$*. The most probable interval is the interval at the peak of the distribution, and is found by evaluating* $dI_n(t)/dt = 0$*. Hint:* $\int_0^\infty x^n e^{-x}\, dx = n!$*.*

Graphs of $I_n(t)$ are shown in Fig. 4(a) for a few values of n. A couple of things are notable. First, the probability of finding zero length intervals for $n > 1$ is zero. This makes sense in that there must be at least one pulse within each such n-pulse interval, and even a single pulse takes time. More interestingly, the relative width of the peak of $I_n(t)$ shrinks as n increases. This is shown most clearly by scaling $I_n(t)$ so that the functions all have the same mean value $(1/r)$ while maintaining the normalized (unit) area, as is shown in Fig. 4(b). What this means is that for larger n, the collection of intervals becomes more uniform.

Exercise 6 *Evaluate the variance σ_n^2 of $I_n(t)$. The variance is given by the formula*

$$\sigma_n^2 = \int_0^\infty (t^2 - \tau_n^2) I_n(t)\, dt \ .$$

Then show that the fractional standard deviation in τ_n, σ_n/τ_n is equal to $1/\sqrt{n}$.

References

[1] Wannier, Gregory H., *Statistical Physics,* (Dover, Mineola NY, 1987), p. 29.

[2] Knoll, Glenn F., *Radiation Detection and Measurement*, 2nd edition, (John Wiley, New York, 1989), pp. 96–99.

[3] Melissinos, Adrian C., and Jim Napolitano, *Experiments in Modern Physics*, 2nd edition (Academic Press, New York, 2003), pp. 401–404, 470–473.

Prepared D. Pengra

`tutorial.tex -- Updated 15 March 2006`