Randomness at the root of things 2: Poisson sequences

Jon Ogborn[1], Simon Collins[2] and Mick Brown[3]

[1] Institute of Education, University of London, UK
[2] Stowe School, Buckingham, UK
[3] Cavendish Laboratory, University of Cambridge, UK

Abstract
This article is the second of a pair of articles about randomness in physics.
In the preceding article we have shown how simple random walk arguments
can give the mean and standard deviation of a number of random counts.
In this article we go a little deeper, and offer a simple and novel teaching
approach to the idea of random events. Taking the Poisson sequence as
fundamental, very simple arguments show that the time intervals between
random events are distributed exponentially. The mathematics needed is
merely the familiar mathematics of exponential change, making a good
opportunity for revision. Further arguments give the form of the Poisson
distribution, on the basis of very elementary assumptions. Several
interesting experiments are suggested.

A random sequence of events

It is of the essence of quantum phenomena that quantum events occur at random. Only their probability can be predicted. It would be nice to be able to show that real quantum events, such as radioactive decay, really do arrive at random. But what is it to be random? How would one tell that a given behaviour is really random?

Poisson events represent a very fundamental kind of random behaviour. An example is the distribution of stars in the sky. Are they distributed randomly, with equal probability per unit area? Nineteenth-century astronomers found that they are not. For one thing, close pairs of stars (binaries) are far too common. So it can be very important to know how to detect random or non-random behaviour.

A Poisson sequence describes the arrival of events randomly in time. Let's spell out carefully the fundamental assumptions behind saying that the arrivals are random:

(1) the probability p that an event occurs in a short time interval Δt is proportional to Δt. That is, $p = \lambda \Delta t$. This is the *linearity* condition.

(2) the constant of proportionality λ is independent of time. This is the *no memory* condition[1]. The arrival of one event (or not) has no influence whatever on the arrival of a later one. λ is called the rate of the sequence.

(3) there is essentially no chance that two events occur at the same time. Thus in time Δt there are only two possibilities: arrival of an event or non-arrival. The probability of one or the other is therefore 1, that is, a certainty. So the probability of *no* event in time Δt is $(1 - \lambda \Delta t)$. This is the *instantaneous event* condition: the events happen 'in a flash', in a time much less than any time interval involved in measurement.

[1] This is the reason we prefer the term 'Poisson sequence' to the common alternative 'Poisson process'. The word 'process' suggests a dependence of what happens next on what happens now, which is exactly what the 'no memory' condition denies.

From just these principles we can obtain many useful and indeed surprising consequences.

How long between counts, and how many to expect?

Switch on your Geiger counter. It 'clicks' irregularly. How many counts can you expect in a given time t? And how long a time should you expect between counts? The number of counts to be expected in time t is easy to write down. It is just:

$$\text{expected number of counts } N = \lambda t.$$

What about the expected time interval τ between counts? It is just the total time t divided by the expected number of counts:

expected time interval between counts,

$$\tau = \frac{t}{N} = \frac{1}{\lambda}.$$

For example, if 50 counts are on average seen in a time of 100 s, the rate $\lambda = 0.5\,\text{s}^{-1}$, and the average interval between counts is $\tau = 2\,\text{s}$.

Next question: by how much can the time between counts, or the total number of counts, be expected to vary, just by random statistical fluctuations? Figure 1 shows what a typical set of randomly arriving counts looks like. They were obtained using a counter connected to a fast data-logger. You may notice that quite closely spaced counts occur, but that there are also some quite large intervals.

these data are the first 0.2 s of a 0.8 s interval containing 263 random counts

count rate $\lambda = \dfrac{263}{0.8} = 329$ Bq

mean waiting time $\langle\,\tau\,\rangle = \dfrac{1}{\lambda} = \dfrac{1}{329} = 0.00304 = 3.0$ ms
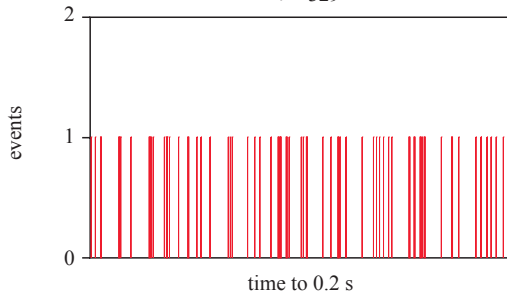

Figure 1. Counts from a radioactive source arriving randomly.

Figure 1 is in stark contrast to the appearance of a deterministic time variation, such as a harmonic oscillation[2]. For this reason, random counts per second are given the unit Bq, whereas frequencies are given the unit Hz, even though both have the dimension time^{-1}. The choice of names is very appropriate: Becquerel discovered radioactivity, and Hertz showed that electromagnetic oscillations could propagate as waves.

How long might you wait between counts?

Let's ask how long you may have to wait between counts. The answer is quite surprising and is easily testable experimentally.

You are waiting for a count. In the coming short interval Δt the probability of a count is $p = \lambda \Delta t$, so the probability of not getting a count is $(1 - \lambda \Delta t)$.

Suppose that $\lambda = 0.01\,\text{s}^{-1}$, and that you wait 100 s, so that $\lambda t = 1$. The probability of no count in each second is $0.99 = 1 - 0.01$. It's the same for the next time-slice, and the next, and so on. Probabilities of independent events that all happen jointly must be multiplied. Thus the probability of getting no count in each and every second out of 100 s is:

$$0.99 \times 0.99 \times 0.99 \times 0.99 \times 0.99 \times \ldots \ (100\,\text{times}).$$

The 'no memory' condition is present as the fact that the multiplying factor is constant, the same for every time interval. If you calculate the value of the product (0.99^{100}) you get 0.37. You may recognize[3] this as $1/e$.

In fact, the probability of getting no count in time t is just

$$p_0(t) = \exp(-\lambda t).$$

The probability decays exponentially with time simply because at each successive interval it is multiplied by a constant factor. That's how exponential change arises.

[2] The difference runs very deep. For regular deterministic variations, the whole panoply of Fourier analysis can be deployed. But this is useless for random variation, for which quite different (statistical) tools are needed. Indeed, this basic difference is one of our reasons for writing this pair of articles.
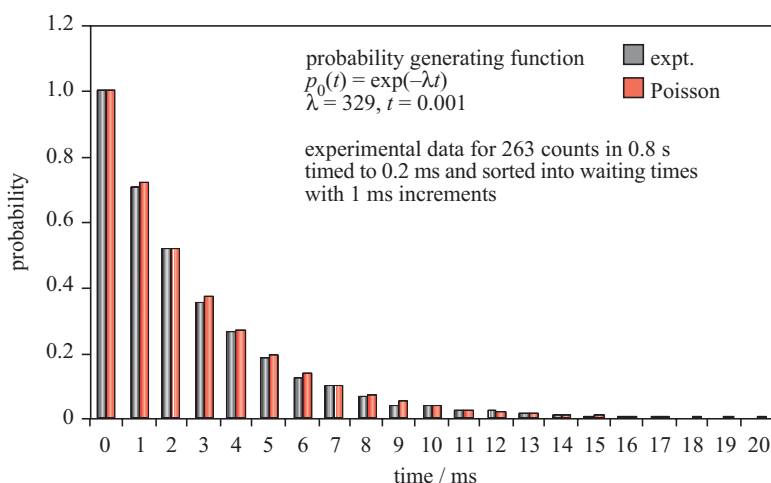[3] The limit of $(1 - 1/N)^N$ as N tends to infinity is e^{-1}.

Figure 2. Fraction of time intervals greater than t varies exponentially with t.
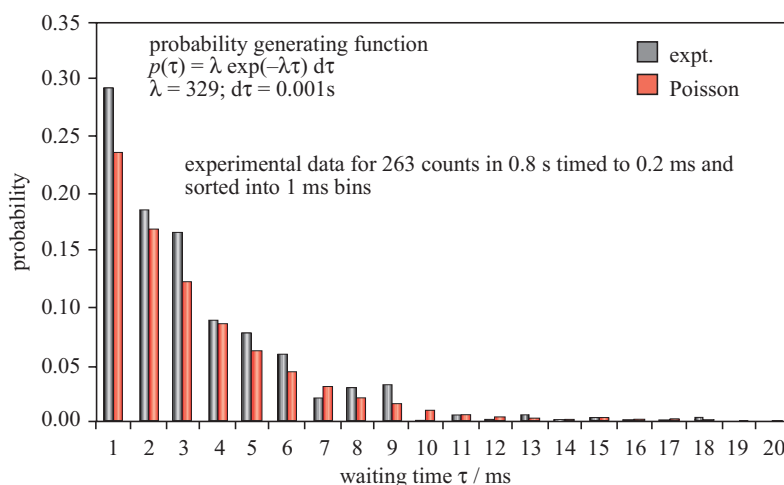

Figure 3. Time intervals between counts are distributed exponentially, too.

Testing waiting time statistics

This result is quite easy to test experimentally. You use a data-logger to record counts over a long period, and extract the time intervals τ between counts. If no event has occurred in time t then the time interval τ between counts is certainly greater than t. Thus the probability that the interval τ between counts is greater than t is the same as the probability that no count has occurred in time t:

$$p_{\tau > t}(t) = p_0(t) = \exp(-\lambda t).$$

The test is then to find the fraction of intervals τ between counts that are greater than a succession of times t. All intervals are greater than $t = 0$,

so the graph starts at unity, as the equation says it should. You can find the rate λ from the total number of counts in a long time, and test the fit of the equation. Figure 2 shows some sample results.

You will also find experimentally that the waiting times themselves are exponentially distributed, with mean time $1/\lambda$ as discussed above. The probability that the waiting time τ will lie in the range τ to $\tau + d\tau$ is given by

$$p(\tau) = \lambda \exp(-\lambda \tau)\, d\tau.$$

This expression is easy to understand. It is just the probability $\lambda\, d\tau$ that there *is* a count in the short interval $d\tau$, multiplied by the probability

$\exp(-\lambda\tau)$ that it did *not* happen in the previous time interval τ. Figure 3 shows some results from an experimental test.

The equation $p(\tau) = \lambda\exp(-\lambda\tau)\,\mathrm{d}\tau$ can be regarded as the fundamental equation describing Poisson sequences in one dimension. Indeed the Poisson distribution can be derived from it (see below). It is not very difficult to show that the equation correctly sums to 1 over all times from zero to infinity, and has mean $1/\lambda$.

It could well be that people arrive randomly at supermarket checkout queues. If so, these waiting time equations give you some handle on working out whether queues will grow or not, and on how their length may vary. Another application is the occurrence in time of faults in communication systems. Perhaps the real surprise is that the distribution of waiting times is exponential, with no peak. Shorter times are always more frequent than longer times.

Proving that $p_0(t) = \exp(-\lambda t)$

Above, we gave only a faint sketch of an argument. A proof is quite simple, and unexpectedly revises a piece of mathematics that is very familiar in A-level physics. So here goes.

Suppose that the probability that no events have arrived after time t is $p_0(t)$. What is the probability $p_0(t + \mathrm{d}t)$ that there will still be no event after waiting a further time $\mathrm{d}t$? It is the probability of two things happening together: no event up to time t and no event in the further time $\mathrm{d}t$. The probability of the joint occurrence is the product of the two probabilities. Thus

$p_0(t + \mathrm{d}t)$

$\quad = p_0(t) \times$ probability of no arrival in time $\mathrm{d}t$.

The probability of no arrival in time $\mathrm{d}t$ follows from the basic assumptions we made previously about a Poisson sequence. It is $(1 - \lambda\,\mathrm{d}t)$. So now we can write down how the probability of no arrival changes with time:

$$p_0(t + \mathrm{d}t) = p_0(t) \times (1 - \lambda\,\mathrm{d}t).$$

You can think of this as the equation of suspense: 'Nothing yet; how long can this go on?' It gives a simple and rather familiar equation for the rate of change of the probability of having no event. Multiplying out the right-hand side, and

subtracting $p_0(t)$ from both sides gives the change of the probability in time $\mathrm{d}t$:

$$\mathrm{d}p_0(t) = p_0(t + \mathrm{d}t) - p_0(t) = -\lambda p_0(t)\,\mathrm{d}t$$

or

$$\frac{\mathrm{d}p_0(t)}{\mathrm{d}t} = -\lambda p_0(t).$$

This is the well-known equation for exponential decay. The probability of having had no arrivals obviously starts at 1 (certainty) when the clock starts ($t = 0$). Then it decreases exponentially with time, with rate λ. That is,

$$p_0(t) = \exp(-\lambda t).$$

The calculation is familiar because it is very like that for the decay of numbers of particles of a radioactive species. But of course it is also very different. It is the change in time of a probability, not the change in time of a number of particles.

Probability of different numbers of counts in time t: Poisson distribution

If you count radioactive emissions for a time t the average number will be near $N = \lambda t$. But how may the number vary? How often will you see 0, 1, 2, 3, etc counts? The answer to this question is the Poisson distribution.

Probability of just 1 count in time t

We already know the probability of getting zero counts. It is

$$p_0(t) = \exp(-\lambda t).$$

From this, we can obtain the probability of getting just one count in time t. Imagine the time line divided into a very large number of tiny intervals $\mathrm{d}t$, as in figure 4. In every interval there is the
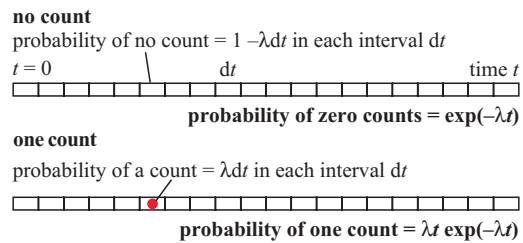
no count

probability of no count = $1 - \lambda\mathrm{d}t$ in each interval $\mathrm{d}t$

$t = 0$ $\mathrm{d}t$ time t

probability of zero counts = $\exp(-\lambda t)$

one count

probability of a count = $\lambda\mathrm{d}t$ in each interval $\mathrm{d}t$

probability of one count = $\lambda t\exp(-\lambda t)$

Figure 4. One count arrives at a random moment during time t.

small probability $\lambda \, \mathrm{d}t$ that one count will occur. So the probability of getting one count in (say) the first interval $\mathrm{d}t$ during time t is equal to:

probability of no count in rest of the time $\times \lambda \, \mathrm{d}t$.

The first part of this expression is very nearly equal to $\exp(-\lambda t)$, since nothing arrives in all but one interval $\mathrm{d}t$. But the single count can come in *any* of the intervals $\mathrm{d}t$. So you must add these probabilities over the total time t. Thus the probability of a single count in just one of the intervals $\mathrm{d}t$ is

$$p_1(t) = \lambda t \exp(-\lambda t).$$

Notice the simple relation $p_1(t) = \lambda t p_0(t)$. We will show below that there is a similar simple general relationship between the probabilities of adjacent numbers of occurrences, which describes the Poisson distribution in a very clear way.

Probability of just 2 counts in time t

The argument for getting to the probability for two counts from the probability for one count is almost the same as above. Just one further count in any one of the intervals $\mathrm{d}t$ is needed. Adding up the probabilities for individual time slices gives, as above, another multiplying factor of λt (we again neglect the small reduction in available intervals). But now there is a subtlety. We have double counted[4] in figure 5. The 'extra' count could have been either the first or the second. So the probability for just two counts in time t is equal to the probability for just one count, multiplied by $\lambda t/2$. Thus

$$p_2(t) = \frac{\lambda t}{2} p_1(t)$$

---

[4] We here gloss over some subtle combinatorial arguments. $n-1$ counts can be permuted in $(n-1)!$ ways without changing the pattern. Adding one extra count multiplies this by n.
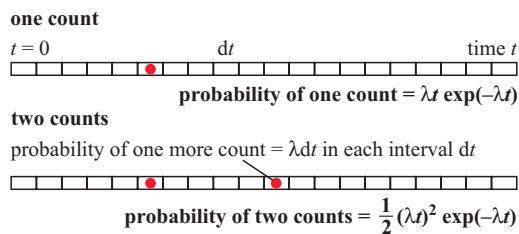
one count

$t = 0$ $\mathrm{d}t$ time t

probability of one count $= \lambda t \exp(-\lambda t)$

two counts

probability of one more count $= \lambda \mathrm{d}t$ in each interval $\mathrm{d}t$

probability of two counts $= \frac{1}{2}(\lambda t)^2 \exp(-\lambda t)$

Figure 5. Two counts arrive at random moments during time t.

and

$$p_2(t) = \frac{(\lambda t)^2}{2} p_0(t) = \frac{(\lambda t)^2}{2} \exp(-\lambda t).$$

Probability of n counts in time t

The argument generalizes straightforwardly. To get a third count multiplies the probability for two counts by $\lambda t/3$ (the factor 3 coming from triple counting). So

$$p_3(t) = \frac{\lambda t}{3} p_2(t)$$

and

$$p_3(t) = \frac{(\lambda t)^3}{3 \times 2} \exp(-\lambda t).$$

In general, to go from the probability for $n - 1$ counts to that for n counts you multiply by the simple factor $\lambda t/n$. Thus the general relationships are

$$p_n(t) = \frac{\lambda t}{n} p_{n-1}(t)$$

and

$$p_n(t) = \frac{(\lambda t)^n}{n!} \exp(-\lambda t)$$

where $n! = n(n-1)(n-2) \ldots 1$ is called 'factorial n'.

This is the *Poisson distribution*. One of the earliest tests of it was nothing to do with physics: it was the distribution of the number of deaths per year by the rare event of a lethal kick from a horse, in cavalry regiments of the 19th century Prussian army (see Appendix).

It is easy to construct a theoretical Poisson distribution on a spreadsheet. Choose a constant rate λ and a time t. Calculate the product λt. Calculate the probability for getting zero counts, $p_0(t) = \exp(-\lambda t)$. Then find the probabilities for, successively, one, two, three etc counts by multiplying the previous value by $\lambda t/n$.

You can see where the 'hump' comes from. For small n, $\lambda t > n$ and the probabilities increase. But once past the peak, where $\lambda t = n$, the probabilities start to decrease. Figure 6 shows some samples, illustrating how the distribution shifts as time t increases. The peak moves to higher values as t increases, and the spread broadens. Notice that the spread is such that the height of the distribution roughly halves in height at distances $\pm\sqrt{N}$ from the peak.
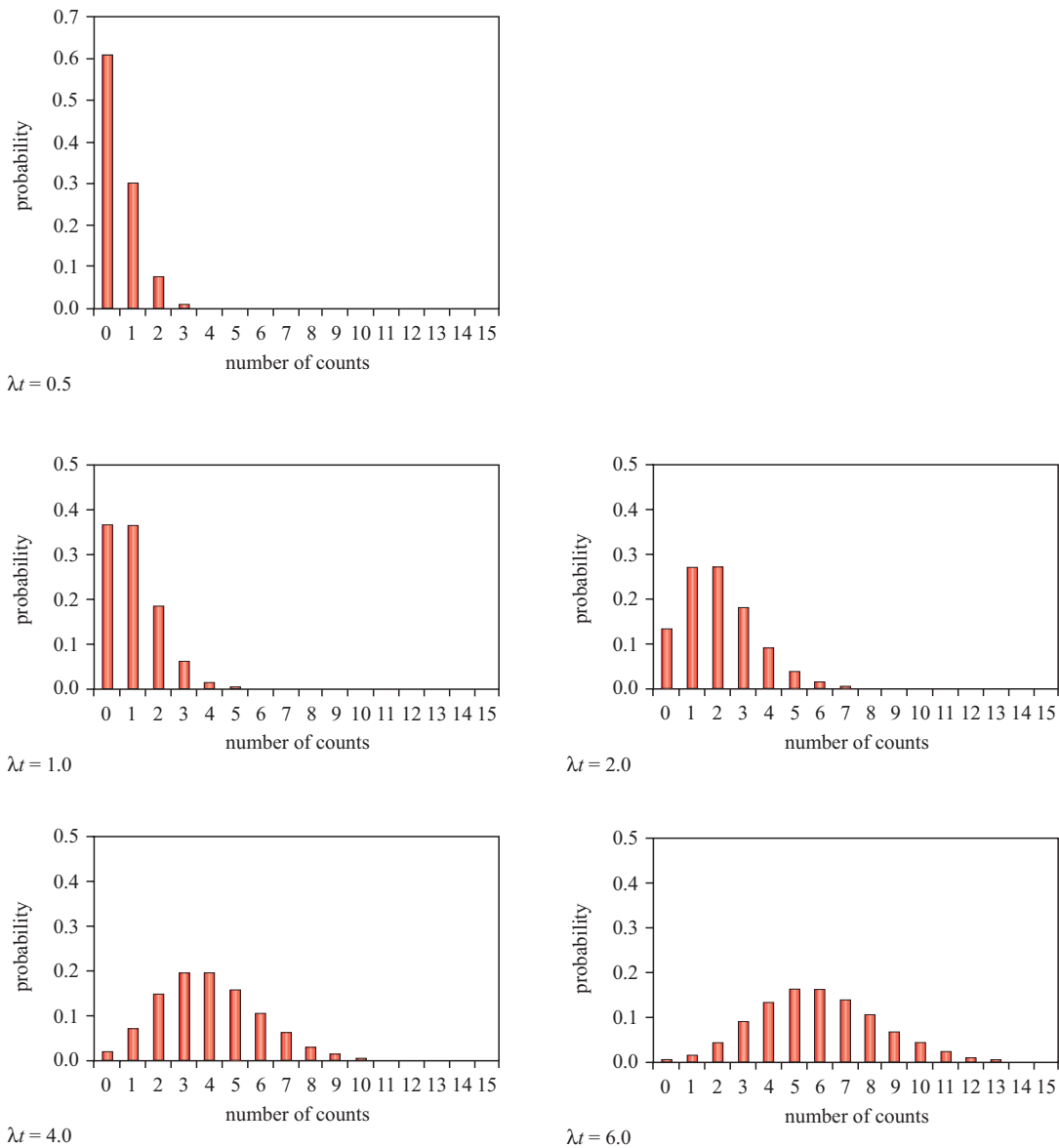
Figure 6. Sample Poisson distributions with λt from 0.5 to 6.

Experimental check on Poisson distribution

It is simple to test whether the numbers of counts in a fixed time t from repeated observations of (say) a radioactive source fit the Poisson distribution. Estimate the rate λ from the total number of counts in all observations, and calculate the expected distribution as above. Superpose experimental and theoretical plots. Figures 7(a) and 7(b) show data taken at two different count rates.

The width of the distribution, showing how much the number of counts in time t typically varies, is of interest. You can approach this experimentally or by numerical simulation. You will find that when the distribution centres on a mean number of counts $N = \lambda t$, the distribution gives appreciable numbers of counts either side of the mean, out to around $N \pm \sqrt{N}$. Proving this is rather heavy work, so we will not attempt it
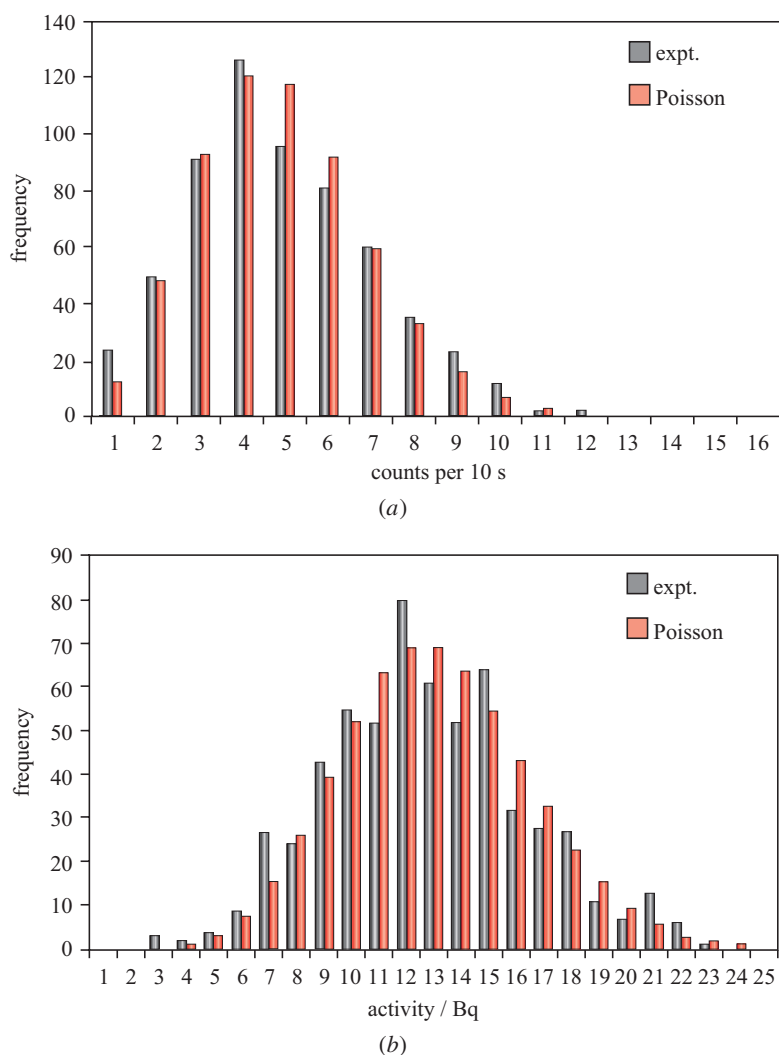
Figure 7. (*a*) Distribution of varying numbers of counts, compared with a Poisson distribution (600 repeated 10 s readings of background count). (*b*) Distribution of varying numbers of counts, compared with a Poisson distribution (600 repeated 1 s readings of counts from a source).

here[5]. Instead, in the preceding article we offered a simpler random walk argument.

Conclusion

We have shown how the Poisson distribution, the most fundamental distribution of random events in continuous time or space, can be approached rather simply through waiting times, with almost no use of difficult combinatorial arguments. This opens up the possibility of testing whether counts from a radioactive source are really random, and

[5] A proof can be obtained by e-mail from j.ogborn@physics.org

so of getting to grips experimentally with the fundamentals of quantum behaviour.

Appendix. Two historic examples of Poisson sequences

Deaths by horse kicks in the Prussian cavalry

Von Bortkiewicz reported the rare event of death by horse kicks in the Prussian army in the late 19th century. Table A1 shows the number of cavalry corps-years in which different numbers of such deaths occurred.

Figure A1 shows that the numbers of cases fits well with the number expected from the Poisson

Table A1. The number of cavalry corps-years in which different numbers of deaths by horse kicks occurred.

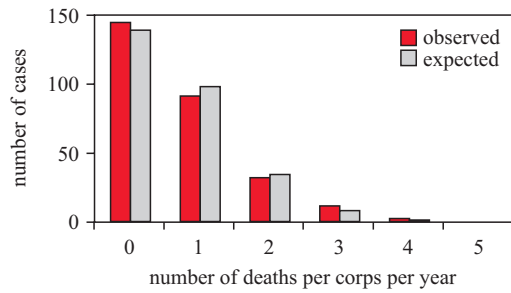| Number of deaths	Number of cases
0	144
1	91
2	32
3	11
4	2
5 or more	0

Figure A1. Deaths by horse kick.

distribution, at a rate of 0.7 deaths per corps per year.

Rutherford tests whether alpha emission is random

In 1910, Rutherford published results of a study of the randomness of the emission of alpha particles. The numbers of particles emitted by a source in a sequence of 2608 time intervals each 1/8 minute long were counted. Table A2 shows the number of intervals for each different number of particles emitted.

Figure A2 shows a good fit between the observed values and the numbers expected from a Poisson distribution with a rate of 3.87 emissions per interval.

Table A2. The numbers of intervals for different number of particles emitted.

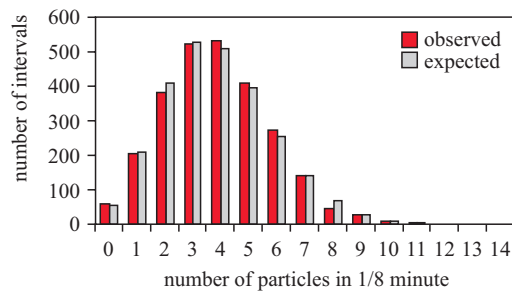| Number of alpha particles emitted	Number of 1/8 minute intervals
0	57
1	203
2	383
3	525
4	532
5	408
6	273
7	139
8	45
9	27
10	10
11	4
12	0
13	1
14	1

Figure A2. Random emission of alpha particles. Reference: Rutherford, Geiger and Bateman 1910 *Phil. Mag.* **20** 698–707.

Jon Ogborn directed the Institute of Physics' *Advancing Physics* project, and is Emeritus Professor of Science Education, Institute of Education, University of London. **Mick Brown** is a Fellow of the Royal Society, Emeritus Professor of Physics, University of Cambridge, and has a long-standing interest in improving physics education. **Simon Collins** teaches physics at Stowe School, and is a Principal Examiner for *Advancing Physics*.