Quick facts #3: Poisson noise

1 Introduction

The idea of *noise* appears whenever we consider the accuracy (or sensitivity) of a measurement. We find quite generally that repeated measurements of an assumed property of a system (such as its length or mass) tend to give slightly different results. This is always true at some level. For example, if we were to measure

the angular size of the Sun from the Earth our measurement would be affected by atmospheric shimmer. If we were to go above the atmosphere, we would see that the solar surface is not smooth, and that the angular diameter changes in time.

Any unpredictable variation in the measured value of a quantity is called *noise*. We can usually improve the accuracy of our measurement if we know something about the statistics of the noise. For example, if the noise is additive and has a mean of zero, averaging successive measurements will help reduce it.

2 Signal-to-noise ratio

It is useful to define a number that represents the quality of a measurement. Sometimes we quote measurement with an associated error, such as $S = 14.8 \pm 0.4$ Jy, or we could quote the error as a percentage of the signal ($S = 14.8\,\text{Jy} \pm 2.7\%$). It's common to define a *signal-to-noise ratio* (SNR) for the measurement, as

$$\text{SNR} = \frac{\text{expected signal strength}}{\text{expected noise strength}}.$$

In our example, the SNR would be $14.8/0.4 = 37$. Clearly, a high value of SNR is good. Measurements with SNRs less than 2 contain very little information on what we are trying to measure. This system is used widely. For example, cassette tapes can reproduce sound with a SNR up to about 10^7, whereas CDs give a SNR greater than 10^9.

The *dynamic range* of an instrument is its ability to measure both weak and strong signals, and is closely related to the maximum signal-to-noise ratio it can deliver.

3 Poisson statistics

In astronomy, we are often interested in estimating flux density by measuring the average arrival rate of photons from a source. Although these photons may be expected to arrive at a constant *average* rate of r photons per unit time interval, they cannot be expected to arrive at regular intervals. In any particular interval of time, τ, the number of photons *expected* is $\langle N \rangle = r\tau$, but the number actually received will probably not equal exactly this. Photons are not aware of each other – they each arrive with some fixed probability per unit time without reference to the number that have already arrived. Such arrivals are said to be *uncorrelated*, and their statistics are particularly straightforward.

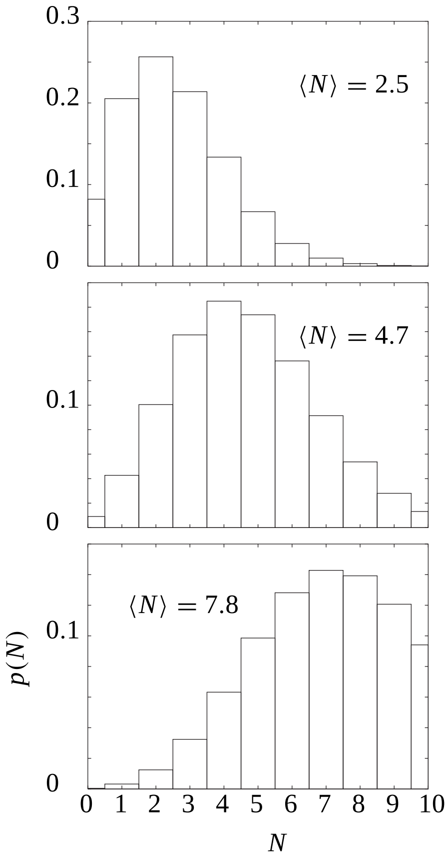

Figure 1: Poisson distributions for three mean arrival rates.

It turns out that if we expect to detect $\langle N \rangle$ photons on average over many observing runs, we can expect to receive about $N = \langle N \rangle \pm \sqrt{\langle N \rangle}$ during any particular run. The probability of detecting exactly N photons is in fact

$$p(N) = \exp(-\langle N \rangle) \frac{\langle N \rangle^N}{N}.$$

Of course we can rewrite this in terms of the arrival rate, r and the observing time, τ:

$$p(N) = \exp(-r\tau) \frac{(r\tau)^N}{N}.$$

This is called the *Poisson distribution* (see Figure 1), and can be derived from fundamental probability theory (see next section). Its most important property is that the standard deviation of the distribution[1] equals the square root of the

[1] 'Standard deviation', σ, is a measure of the statistical spread of a distribution, and characterises the variation between measurements. It is defined as $\sigma = (\langle N^2 \rangle - \langle N \rangle^2)^{1/2}$. σ^2 is called the *variance* of the distribution.

mean of the distribution, i.e.,

$$\sigma = \sqrt{\langle N \rangle} = \sqrt{r\tau}.$$

If we just make one measurement, and receive N photons, then our best estimate of $\langle N \rangle$ is N. If we associate σ with the noise in the measurement, then our best estimate of our error is $\sqrt{N}$.

This kind of noise is variously called 'Poisson noise', 'shot noise', 'self noise' or 'photon noise'. It is a direct consequence of the statistics of photon-counting. The fundamental result is that, if we base a measurement on the arrival of N photons, our SNR cannot be greater than

$$\mathrm{SNR} = \frac{N}{\sqrt{N}} = \sqrt{N}.$$

Other noise sources may be present, and they can only reduce this SNR further. The rule is that we add the squares of the standard deviations, called the *variances*, of the various noise processes to get the overall variance, and therefore the overall standard deviation.

In the limit of very large N, the Poisson distribution approaches the (perhaps more familiar) Gaussian distribution. This is a less awkward expression for large N, but the same rules apply. Quite generally, the SNR of a measurement increases as $\sqrt{N} = \sqrt{r\tau}$, i.e., as the square root of the observing time. Even in radio astronomy, where we are no longer counting photons, the SNR $\propto \sqrt{\Delta\nu\tau}$ where $\Delta\nu$ is the observing bandwidth.[2]

4 Deriving the Poisson distribution

This section is not examinable, but demonstrates how the Poisson distribution arises.

We expect, on average, $\langle N \rangle = r\tau$ photons to be detected during our observing period τ. Photons will arrive at random times during this period, but we can imagine dividing τ into a very large number of sub-intervals (M of them) which are sufficiently short that they either contain one photon or no photons. The probability, p_1, that any one sub-interval contains a photon is the expected number of photons divided by the number of sub-intervals, i.e.,

$$p_1 = \frac{r\tau}{M}.$$

The probability, p_0, that the sub-interval does not contain a photon is simply

$$p_0 = 1 - p_1,$$

because one or other outcome is certain[3] ($p_0 + p_1 = 1$).

[2] You can think of $\Delta\nu\tau$ as the number of independent measurements of the electric field than can be made in the observing time – the equivalent of N when we are photon counting.

[3] We have assumed that M is so big that the chances of two photons arriving in the same sub-interval is zero.

The probability of, say, the first N sub-intervals containing a photon and the remaining $M - N$ being empty is $p_1^N p_0^{M-N}$. This is just one way in which N photons can be distributed in the interval. There are many other ways. In fact there are M ways to insert the first photon, $M - 1$ ways to insert the second (avoiding the first), $M - 2$ ways to insert the third and so on. There are therefore $M/(M - N)$ ways to insert all the photons. But we have over-counted. We have included identical configurations that have been arrived at merely by inserting photons in different orders. There are N ways to do this (the number of rearrangements of N objects), so the final number of arrangements of N photons in M sub-intervals is

$$_M C_N = \frac{M}{(M - N)\, N}.$$

Each of these configurations has an equal probability of happening ($= p_1^N p_0^{M-N}$), so the overall *probability* of receiving N photons is

$$p(N) = \frac{M}{(M - N)\, N} p_1^N p_0^{M-N}.$$

This is known as the *binomial* distribution, and is valid for all values of M. However, we have assumed that M is very large, so we can consider its limiting form as $M \to \infty$ (and therefore as $p_1 \to 0$ with $Mp_1 = r\tau$). Without approximation, we can rewrite the above equation as

$$p(N) = \frac{(r\tau)^N}{N} p_0^{M-N} f(M),$$

where

$$f(M) = \frac{M(M-1)(M-2)\dots(M-N+1)}{M^N}.$$

As $M \to \infty$, $f(M)$ approaches unity, as N becomes trivially small compared with M. We can also say that

$$\begin{aligned} p_0^{M-N} &= (1 - p_1)^{M-N} \\ &= \left(1 - \frac{r\tau}{M}\right)^{M-N} \\ &= \frac{(1 - r\tau/M)^M}{(1 - r\tau/M)^N}. \end{aligned}$$

As $M \to \infty$ the numerator in the above expression approaches $\exp(-r\tau)$ and the denominator approaches unity. The final expression for $p(N)$ in the limit we require of large M is therefore

$$p(N) = \exp(-r\tau) \frac{(r\tau)^N}{N},$$

which is our expression for the Poisson distribution.

G.W.
2009