Glen Cowan
RHUL Physics
1 December, 2009

Derivation of the Poisson distribution

I this note we derive the functional form of the Poisson distribution and investigate some of its properties. Consider a time t in which some number n of events may occur. Examples are the number of photons collected by a telescope or the number of decays of a large sample of radioactive nuclei. Suppose that the events are *independent*, i.e., the occurrence of one event has no influence on the probability for the occurrence of another. Furthermore, suppose that the probability of a single event in any short time interval δt is

$$P(1; \delta t) = \lambda \, \delta t \, , \tag{1}$$

where λ is a constant. In Section 1 we will show that the probability for n events in the time t is given by

$$P(n; \nu) = \frac{\nu^n}{n!} e^{-\nu} \, , \tag{2}$$

where the parameter ν is related to λ by

$$\nu = \lambda t \, . \tag{3}$$

We will follow the convention that arguments in a probability distribution to the left of the semi-colon are random variables, that is, outcomes of a repeatable experiment, such as the number of events n. Arguments to the right of the semi-colon are parameters, i.e., constants.

The Poisson distribution is shown in Fig. 1 for several values of the parameter ν. In Section 2 we will show that the mean value $\langle n \rangle$ of the Poisson distribution is given by

$$\langle n \rangle = \nu \, , \tag{4}$$

and that the standard deviation σ is

$$\sigma = \sqrt{\nu} \, . \tag{5}$$

The mean ν roughly indicates the central region of the distribution, but this is not the same as the most probable value of n. Indeed n is an integer but ν in general is not. The standard deviation is a measure of the width of the distribution.

1 Derivation of the Poisson distribution

Consider the time interval t broken into small subintervals of length δt. If δt is sufficiently short then we can neglect the probability that two events will occur in it. We will find one event with probability

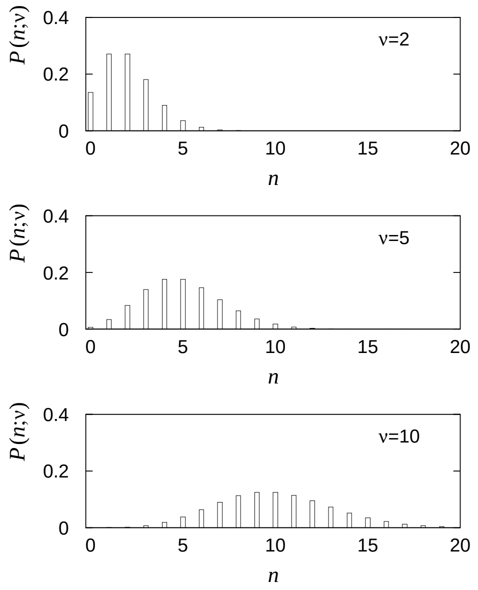

Figure 1: The Poisson distribution $P(n;\nu)$ for several values of the mean ν.

$$P(1;\delta t) = \lambda\,\delta t \tag{6}$$

and no events with probability

$$P(0;\delta t) = 1 - \lambda\,\delta t\ . \tag{7}$$

What we want to find is the probability to find n events in t. We can start by finding the probability to find zero events in t, $P(0;t)$ and then generalize this result by induction.

Suppose we knew $P(0;t)$. We could then ask what is the probability to find no events in the time $t+\delta t$. Since the events are independent, the probability for no events in both intervals, first none in t and then none in δt, is given by the product of the two individual probabilities. That is,

$$P(0;t+\delta t) = P(0;t)(1-\lambda\,\delta t)\ . \tag{8}$$

This can be rewritten as

$$\frac{P(0;t+\delta t) - P(0;t)}{\delta t} = -\lambda P(0;t)\ , \tag{9}$$

which in the limit of small δt becomes a differential equation,

$$\frac{dP(0;t)}{dt} = -\lambda P(0;t)\ . \tag{10}$$

Integrating to find the solution gives

$$P(0;t) = Ce^{-\lambda t}\ . \tag{11}$$

For a length of time $t=0$ we must have zero events, i.e., we require the boundary condition $P(0;0)=1$. The constant C must therefore be 1 and we obtain

$$P(0;t) = e^{-\lambda t} \; . \tag{12}$$

Now consider the case where the number of events n is not zero. The probability of finding n events in a time $t + \delta t$ is given by the sum of two terms:

$$P(n; t + \delta t) = P(n;t)(1 - \lambda \, \delta t) + P(n-1;t)\lambda \, \delta t \; . \tag{13}$$

The first term gives the probability to have all n events in the first subinterval of time t and then no events in the final δt. The second term corresponds to having $n - 1$ events in t followed by one event in the last δt. In the limit of small δt this gives a differential equation for $P(n;t)$:

$$\frac{dP(n;t)}{dt} + \lambda P(n;t) = \lambda P(n-1;t) \; . \tag{14}$$

We can solve equation (14) by finding an integrating factor $\mu(t)$, i.e., a function which when multiplied by the left-hand side of the equation results in a total derivative with respect to t. That is, we want a function $\mu(t)$ such that

$$\mu(t) \left[\frac{dP(n;t)}{dt} + \lambda P(n;t) \right] = \frac{d}{dt} \left[\mu(t) P(n;t) \right] \; . \tag{15}$$

We can easily show that the function

$$\mu(t) = e^{\lambda t} \tag{16}$$

has the desired property and therefore we find

$$\frac{d}{dt} \left[e^{\lambda t} P(n;t) \right] = e^{\lambda t} \lambda P(n-1;t) \; . \tag{17}$$

We can use this result, for example, with $n = 1$ to find

$$\frac{d}{dt} \left[e^{\lambda t} P(1;t) \right] = \lambda e^{\lambda t} P(0;t) = \lambda e^{\lambda t} e^{-\lambda t} = \lambda \; , \tag{18}$$

where we substituted our previous result (12) for $P(0;t)$. Integrating equation (18) gives

$$e^{\lambda t} P(1;t) = \int \lambda \, dt = \lambda t + C \; . \tag{19}$$

Now the probability to find one event in zero time must be zero, i.e., $P(1;0) = 0$ and therefore $C = 0$, so we find

$$P(1;t) = \lambda t e^{-\lambda t} \; . \tag{20}$$

We can generalize this result to arbitrary n by induction. We assert that the probability to find n events in a time t is

$$P(n;t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \; . \tag{21}$$

We have already shown that this is true for $n = 0$ as well as for $n = 1$. Using the differential equation (17) with $n + 1$ on the left-hand side and substituting (21) on the right, we find

$$\frac{d}{dt}\left[e^{\lambda t} P(n+1;t)\right] = e^{\lambda t}\lambda P(n;t) = e^{\lambda t}\lambda \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \lambda \frac{(\lambda t)^n}{n!} \,. \tag{22}$$

Integrating equation (22) gives

$$e^{\lambda t} P(n+1;t) = \int \lambda \frac{(\lambda t)^n}{n!}\, dt = \frac{(\lambda t)^{n+1}}{(n+1)!} + C \,. \tag{23}$$

Imposing the boundary condition $P(n + 1; 0) = 0$ implies $C = 0$ and therefore

$$P(n+1;t) = \frac{(\lambda t)^{n+1}}{(n+1)!} e^{-\lambda t} \,. \tag{24}$$

Thus the assertion (21) for n also holds for $n + 1$ and the result is proved by induction.

2 Mean and standard deviation of the Poisson distribution

First we can verify that the sum of the probabilities for all n is equal to unity. Using now $\nu = \lambda t$, we find

$$\begin{aligned} \sum_{n=0}^{\infty} P(n;\nu) &= \sum_{n=0}^{\infty} \frac{\nu^n}{n!} e^{-\nu} \\ &= e^{-\nu} \sum_{n=0}^{\infty} \frac{\nu^n}{n!} \\ &= e^{-\nu} e^{\nu} \\ &= 1 \,, \end{aligned} \tag{25}$$

where we have identified the final sum with the Taylor expansion of e^{ν}.

The *mean value* (or *expectation value*) of a discrete random variable n is defined as

$$\langle n \rangle = \sum_n n P(n) \,, \tag{26}$$

where $P(n)$ is the probability to observe n and the sum extends over all possible outcomes. In the case of the Poisson distribution this is

$$\langle n \rangle = \sum_{n=0}^{\infty} n P(n;\nu) = \sum_{n=0}^{\infty} n \frac{\nu^n}{n!} e^{-\nu} \,. \tag{27}$$

To carry out the sum note first that the $n = 0$ term is zero and therefore

$$\begin{aligned}
\langle n \rangle &= e^{-\nu} \sum_{n=1}^{\infty} n \frac{\nu^n}{n!} \\
&= \nu e^{-\nu} \sum_{n=1}^{\infty} \frac{\nu^{n-1}}{(n-1)!} \\
&= \nu e^{-\nu} \sum_{m=0}^{\infty} \frac{\nu^m}{m!} \\
&= \nu e^{-\nu} e^{\nu} \\
&= \nu \; .
\end{aligned} \tag{28}$$

Here in the third line we simply relabelled the index with the replacement $m = n - 1$ and then we again identified the Taylor expansion of e^{ν}.

To find the standard deviation σ of n we use the defining relation

$$\sigma^2 = \langle n^2 \rangle - \langle n \rangle^2 \; . \tag{29}$$

We already have $\langle n \rangle$, and we can find $\langle n^2 \rangle$ using the following trick:

$$\langle n^2 \rangle = \langle n(n-1) \rangle + \langle n \rangle \; . \tag{30}$$

We can find $\langle n(n-1) \rangle$ in a manner similar that used to find $\langle n \rangle$, namely,

$$\begin{aligned}
\langle n(n-1) \rangle &= \sum_{n=0}^{\infty} n(n-1) \frac{\nu^n}{n!} e^{-\nu} \\
&= \nu^2 e^{-\nu} \sum_{n=2}^{\infty} \frac{\nu^{n-2}}{(n-2)!} \\
&= \nu^2 e^{-\nu} \sum_{m=0}^{\infty} \frac{\nu^m}{m!} \\
&= \nu^2 e^{-\nu} e^{\nu} \\
&= \nu^2 \; ,
\end{aligned} \tag{31}$$

where we used the fact that the $n = 0$ and $n = 1$ terms are zero. In the third line we relabelled the index using $m = n - 2$ and identified the resulting series with e^{ν}. Putting this into equation (29) for σ^2 gives $\sigma^2 = \nu^2 + \nu - \nu^2 = \nu$ or

$$\sigma = \sqrt{\nu} \; . \tag{32}$$

This is the important result that the standard deviation of a Poisson distribution is equal to the square root of its mean.