Derivation of the Poisson Distribution

stats methodologists meeting

9 February 2017

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

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

www.pp.rhul.ac.uk/~cowan/stat/notes/PoissonNote.pdf

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

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

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.

Poisson processes are (totally) random

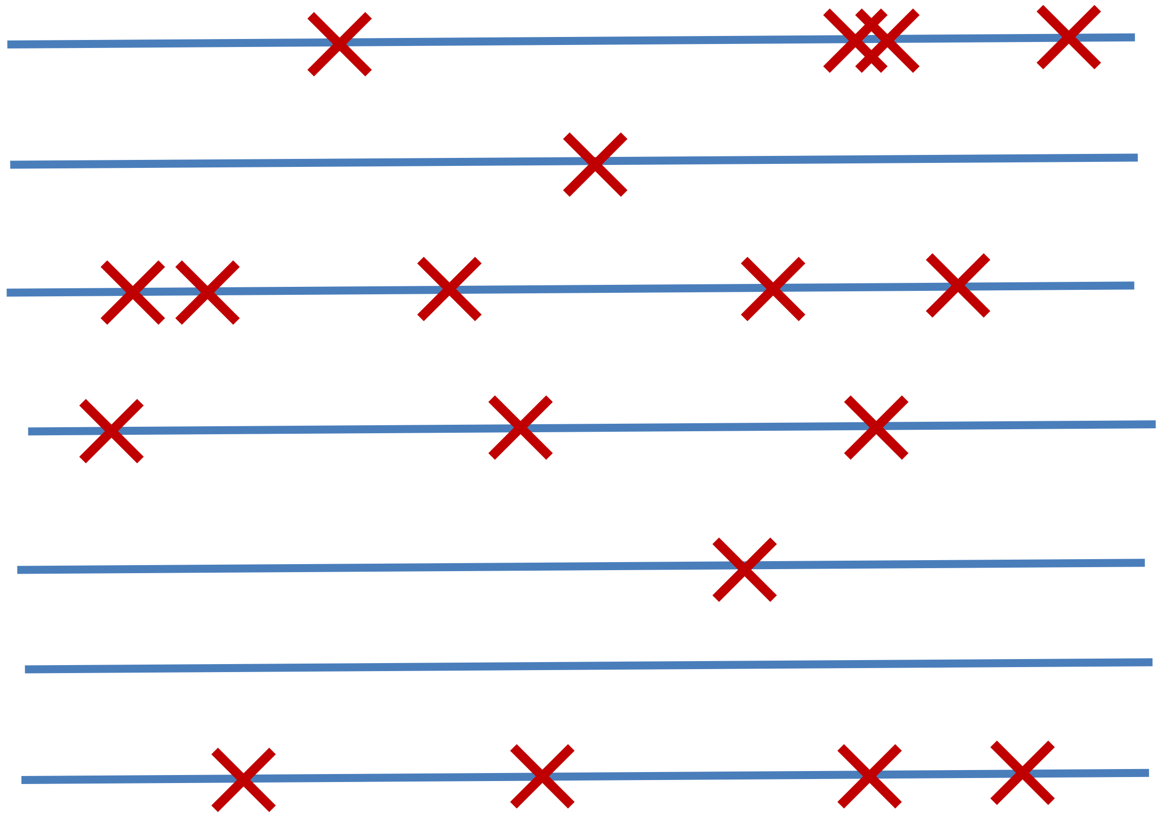
- we watch for a period of time…

- …events occur unpredictably…

- …but at a constant average rate

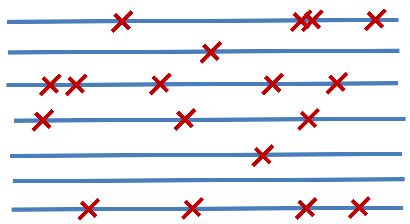
look at 'Distribution' this generates…

- collect ***lots*** of representative samples ('sticks')…



look at 'Distribution' this generates...

- collect ***lots*** of sample 'sticks'...



- ...and group them together according to the number of events

No. of events	0	1	2	k
No. of 'sticks'	n_0	n_1	n_2	n_k

The Poisson Distribution

Group (#events)	0	1	2	k
No. of 'sticks'	n_0	n_1	n_2	n_k

- $n_0 + n_1 + n_2 \ldots = N$ (total number of sticks)
- $0.n_0 + 1.n_1 + 2.n_2 \ldots = L$ (total number of events)
- average number of events = $L/N = \mu$ (say)
- Prob[k events] = $n_k/N = P_k$ (say)
- As if L cards were dealt to N players totally at random...

The Poisson Distribution...

Group (#events)	0	1	2	k
No. of 'sticks'	n_0	n_1	n_2	n_k

- ...Generated by dealing L cards ***randomly*** to N players...
- So we won't change the distribution by re-allocating one card at random
- *Therefore* the chance of taking it off a player in Group k must equal the chance of giving it to a player in Group (k-1)

Poisson Distribution: re-deal a card

Group (#events)	0	1	2	k
No. of 'sticks'	n_0	n_1	n_2	n_k

- chance of taking it off a player in Group k = total cards held by Group k players / total cards

  $= k.n_k/L = k(n_k/N)/(L/N) = k.\text{Prob}[k \text{ events}] /\mu = k.P_k/\mu$

- the chance of giving it to a player in Group (k-1) = #Group (k-1) players / total players

  $= \text{Prob}[(k-1) \text{ events}] = P_{k-1}$

Poisson Distribution: re-deal a card

Group (#events)	0	1	2	k
No. of 'sticks'	n_0	n_1	n_2	n_k

- k.Prob[k events] /μ = Prob[(k-1) events]
- $P_k = (\mu/k).P_{k-1}$
- $P_k = (\mu^k/k!).P_0$ Note: $Sum(P_k)=1$
- $P_0 = e^{-\mu}$
- $P_k = e^{-\mu}\,\mu^k / k!$