

A loose proof of the Poisson probability formula

jh2946

March 2024

Introduction

This paper will provide strong reasoning (but not a formal proof) for the Poisson probability formula:

$$P(X = n) = \frac{1}{n!} \lambda^n e^{-\lambda}$$

Imagine an experiment on a rainy day where we have a stopwatch and a bucket in open air. In 1 second, the bucket receives 1 drop of rain on average.

Start the stopwatch at $t = 0$. Then when the stopwatch reads t seconds, the number of raindrops that have hit the bucket follows the distribution $Po(t)$. We'll assign a random variable to the number of drops that have fallen after t seconds: $X(t) \sim Po(t)$.

Define $y_n(t) = P(X(t) = n)$. Our hypothesis is hence

$$y_n(t) = \frac{1}{n!} t^n e^{-t} \tag{1}$$

Base case

Firstly let's express our system in terms of "pixel time". We define a constant duration Δt , and section all of time into disjoint time "pixels", each Δt seconds long. In each pixel, an event can only either happen 0 or 1 times. We do not care about the exact time the event occurs, we only look at which pixel the event occurs in. Under these conditions, it's impossible to define a Poisson distribution, since in a real Poisson distribution you could always have 5 raindrops falling in the same millisecond – possible, but quite unlikely. But we can still define the behaviour of the raindrops such that the expected number of raindrops is 1 per second (strictly speaking for our argument, each second would have to have an integer number of pixels, but in reality this isn't necessary). If we set the probability of a raindrop falling in each pixel proportional to the length of the pixel, then shrink the pixels arbitrarily small, we get a system that's closer and closer to how a Poisson distribution behaves in reality.

To illustrate with an example, let's set $\Delta t = \frac{1}{2}$, meaning an event occurs once or none at all in each half-second pixel. To get an expected raindrop count of 1 per second, we set the probability of each pixel producing a raindrop to be $\frac{1}{2}$ as well. If we set $\Delta t = 0.001$, then the probability of each pixel producing a raindrop is also 0.001. Each second has $\frac{1}{\Delta t}$ trials (assuming $\frac{1}{\Delta t}$ is an integer, but to some extent this applies to all Δt), hence if we set the probability of a raindrop per pixel to be Δt as well, we get an expected raindrop count of $\frac{\Delta t}{\Delta t} = 1$ per second.

Define $\Delta X = X(t + \Delta t) - X(t)$, which is to say, ΔX is the number of raindrops that fall during the new pixel, which can be 0 or 1.

To start off, what's the probability that we'll see 0 raindrops after waiting for t seconds? Not sure, but we can say that if we want a total of 0 raindrops from the start of the measurement to the next

pixel ($X(t + \Delta t) = 0$), we have to see 0 total raindrops from the start of the experiment to the current pixel ($X(t) = 0$), then have 0 more raindrops fall in the next pixel ($\Delta X = 0$). The last two events are independent of each other, so we can express the first probability as the product of the latter two probabilities:

$$P(X(t + \Delta t) = 0) = P(X(t) = 0) \cdot P(\Delta X = 0)$$

Since the probability of a raindrop falling in one pixel is Δt , the probability of none falling is $1 - \Delta t$, hence yielding

$$P(X(t + \Delta t) = 0) = (1 - \Delta t)P(X(t) = 0)$$

Converting to our y definition we get:

$$y_0(t + \Delta t) = (1 - \Delta t)y_0(t)$$

Then rearrange to get:

$$\frac{y_0(t + \Delta t) - y_0(t)}{\Delta t} = -y_0(t)$$

If we shrink the pixel size arbitrarily small and we take the limit as $\Delta t \rightarrow 0^+$, then Δt turns into dt , yielding

$$\begin{aligned} \frac{y_0(t + dt) - y_0(t)}{dt} &= -y_0(t) \\ \frac{d}{dt}y_0(t) &= -y_0(t) \end{aligned}$$

Approximating a system using somewhat ill-defined limits isn't very rigorous, but I find it quite sufficient as a mathematical model of a complex physical process.

The initial condition is $y_0(t) = 1$, since at the instant you start measuring at $t = 0$ you're guaranteed to have 0 total raindrops. Solving the above ODE with this initial condition yields

$$y_0(t) = e^{-t}$$

According to (1), when $n = 0$,

$$\begin{aligned} y_0(t) &= \frac{1}{0!}t^0e^{-t} \\ &= e^{-t} \end{aligned}$$

hence the hypothesis matches our result for $n = 0$.

Inductive case

For cases where we see 1 or more total raindrops after $t + \Delta t$ seconds, there are two possibilities: one where $X(t + \Delta t) = X(t)$ (same as our 0 case) and one where $X(t + \Delta t) = X(t) + 1$ (since it might be possible a raindrop caused our count to increase to $X(t + \Delta t)$). Then our probability equation has two terms on the RHS for $n \geq 1$:

$$P(X(t + \Delta t) = n) = [P(X(t) = n) \cdot P(\Delta X = 0)] + [P(X(t) = n - 1) \cdot P(\Delta X = 1)]$$

Recall that $P(\Delta X = 1) = \Delta t$, which yields

$$\begin{aligned} P(X(t + \Delta t) = n) &= (1 - \Delta t)P(X(t) = n) + P(X(t) = n - 1)\Delta t \\ y_n(t + \Delta t) &= (1 - \Delta t)y_n(t) + y_{n-1}(t)\Delta t \\ \frac{y_n(t + \Delta t) - y_n(t)}{\Delta t} &= y_{n-1}(t) - y_n(t) \end{aligned}$$

As $\Delta t \rightarrow 0^+$,

$$\begin{aligned}\frac{y_n(t + dt) - y_n(t)}{dt} &= y_{n-1}(t) - y_n(t) \\ \frac{d}{dt}y_n(t) &= y_{n-1}(t) - y_n(t)\end{aligned}\tag{2}$$

This is true for $n \geq 1$.

Substituting (1) into the RHS of (2) we get

$$\frac{1}{(n-1)!}t^{n-1}e^{-t} - \frac{1}{n!}t^n e^{-t}$$

Differentiating (1) yields the LHS of (2):

$$\begin{aligned}\frac{d}{dt} \frac{1}{n!}t^n e^{-t} &= \frac{1}{n!}(nt^{n-1}e^{-t} - t^n e^{-t}) \\ &= \frac{1}{n!}nt^{n-1}e^{-t} - \frac{1}{n!}t^n e^{-t} \\ &= \frac{1}{(n-1)!}t^{n-1}e^{-t} - \frac{1}{n!}t^n e^{-t}\end{aligned}$$

Hence, assuming that (1) is correct for the $n-1$ case, the first-order ODE is fulfilled for the n case.

The initial condition is $y_n(0) = 0$ for all $n \geq 1$, since you can't have a non-zero number of raindrops falling in 0 seconds.

Substituting $t = 0$ into (2) yields:

$$\begin{aligned}y_n(0) &= \frac{1}{n!}0^n e^{-0} \\ &= 0\end{aligned}$$

The initial condition of (1) also matches our understanding, hence given that (1) is correct for the $n-1$ case, by the existence and uniqueness theorem for first-order ODEs, (1) is also correct for the n case. Hence completes the proof by induction for

$$y_n = \frac{1}{n!}t^n e^{-t}$$

■.