The Poisson Process

George Theakston, Gregory Pollard, Tara McClaughlin, Lorenzo Colpani, Lokti Spryou-Kirby and Ben Walton

Supervisor: Professor Kostantinos Fokianos

November 29, 2020

Contents

1 Introduction

Suppose we wish to know the probability that n cars pass a point in the road in a given time frame or that we wish to know the chances of exactly 10 customers entering a store in an hour. These situations can be modeled by a counting process known as the Poisson process.

The Poisson process is a special counting process in which the events being counted in each time interval follow a Poisson distribution, first formally used by Poisson in [5], in which the events have a mean number of occurrences, λ, per interval of time or space

In this report we will discuss the Poisson process as it relates to examples that consider only time and two dimensional space and we will cover the differences between the properties both of the homogeneous Poisson process (which may henceforth be referred to as HPP) and the non-homogeneous Poisson process (NHPP), providing examples of situations in which these may be applied, and simulating both the homogeneous and non-homogeneous processes.

2 Properties

2.1 Homogeneous Poisson Process

The following definitions are from Ross (1996), we will begin by defining what is a counting process:

Definition 2.1. A random process { N(t), $t \in [0, \infty)$ } is said to be a counting process if N(t) satisfy:

I $N(t) \neq 0$.

II $N(t)$ is integer valued.

III If $s < t$, then $N(s) \leq N(t)$.

IV For $s < t$, $N(t) - N(s)$ equals to number of events that have occurred in the interval (s,t].

Now we will give a very small definition, but which is critical, a function o(h):

Definition 2.2. A function f is said to be o(h) if
$\lim_{h \to 0} \frac{f(h)}{h} = 0$.

Lastly but not least, The Poison process, which obviously is a critical definition for this topic.

Definition 2.3. The counting process {N(t), $t \geq 0$} is said to be a Poison process with rate $\lambda, \lambda > 0$ if:

I $N(0) = 0$.

II The process has stationary and independent increments.

III $P\{N(h) = 1\} = \lambda h + o(h)$.

IV $P\{N(h) \geq 2\} = o(h)$.

The condition of a Poisson process arise from the definition of a Poisson Distribution and counting processes shown in the following proof:

Proof. let t be a set time and define a probability Pn(h) such that,

$$P_n(t) = P\{N(t) = n\}.$$

therefore we can define $P_0(t+h)$ as:

$$P_0(t+h) = P\{N(t+h) = 0\}.$$

Because The process has stationary and independent increments we can split into multiple probabilities which will arise from splitting P(No events in t) and P(no events in t to t + h) also using the 3^{rd} assumption of Definition 2.3, we have that $P_0(t+h)$ equals to:

$$= P\{N(t) = 0, N(t+h) - N(t) = 0\}.$$

$$= P\{N(t) = 0\}P\{N(t+h) - N(t) = 0\}.$$

$$P\{N(h) = 0\} = 1 - \lambda h + o(h).$$

. utilizing the definition of differentiation we can utilize the previous result to create the differential of P_0 such that it equals to:

$$\frac{P_0(t+h) - P_0(t)}{h} = -\lambda P_o(t) + \frac{o(h)}{h}$$

.
Therefore by definition 2.2 the function o(h)/h will tend to 0 as h tends to 0 and the right hand side its the differential of P_0 as h tends to 0:

$$P_0'(t) = -\lambda P_0(t).$$

rearranging it so it is easier to integrate and integrating after we will get as follow:

$$log(P_0(t)) = -\lambda t + c.$$

raising it to the power of e, we have:

$$P_0(t) = Ke^{-\lambda t}, \text{where } K = e^c.$$

Since $P_0(0) = P\{N(0) = 0\} = 1$, results to K = 1 and so:

$$P_0(t) = e^{-\lambda t}.$$

Now we will prove for $n >= 1$ very similarly as with $n = 0$, therefore let $P_n(t+h)$ equal to:

$$P_n(t+h) = P\{N(t+h) = n\}.$$

if we split it with the three possible cases, we have:

$$= P\{N(t) = n, N(t+h) - N(t) = 0\} + P\{N(t) = n-1, N(t+h) - N(t) = 1\}$$

$$+P\{N(t+h) = n, N(t+h) - N(t) >= 2\}.$$

since definition 2.3 4^{th} assumption, the last term is o(h). And using as well the second assumption, which was about the process has stationary and independent increment. Lastly using all the values that we already know or have calculated previously. We can derive that $P_n(t+h)$ is:

$$P_n(t+h) = P_n(t)P_0(h) + P_{n-1}(t)P_1(h) + o(h) = (1 - \lambda h)P_n(t) + \lambda h P_{n-1}(t) + o(h).$$

Hence if we want to find its derivative we do as with the case were n = 0, and it will result to:

$$\frac{P_n(t+h) - P_n(t)}{h} = -\lambda P_n(t) + \lambda h P_{n-1}(t) + \frac{o(h)}{h}$$

.

As $h \to 0$, the left hand side becomes the definition of differentiation and using definitions 2.2 about it's stationary and independent increments . We get:

$$P_n'(t) = -\lambda P_n(t) + \lambda P_{n-1}(t).$$

Reorganizing such that the integration will be easily done, multiplying both sides by $e^{\lambda t}$ such that the right hand side is a product rule and integrating thereafter:

$$\frac{d}{dt} e^{\lambda t} P_n(t) = \lambda e^{\lambda t} P_{n-1}(t)$$

Now by using induction we shall proof that the Poison distribution arise from Poison process and counting process:

Since the Poisson distribution is $\frac{e^{-\lambda t}(\lambda t)^n}{n!}$ therefore for n = 1, we have:

$$\frac{d}{dt}(e^{\lambda t} P_1(t)) = \lambda e^{\lambda t} * e^{-\lambda t}$$

Therefore if we integrate, $P_1(t) = (\lambda t + c) * e^{-\lambda t}$

since P(0) = 0, yields:

$$P_1(t) = \lambda t * e^{-\lambda t}$$

assuming that for n - 1 is true. we have: $\frac{d}{dt}(e^{\lambda t}P_n(t)) = \frac{\lambda(\lambda t)^{n-1}}{(n-1)!}$.
Integrating both sides: $e^{\lambda t}P_n(t) = \frac{\lambda(\lambda t)^n}{(n)!} + c$, since $p_n(0) = P\{N(0) = n\} = 0, c = 0$ therefore,

$$P_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}.$$

Since for n = 1 is true, and assuming that n - 1 is true. we get that n is also true therefore. n = 1,2,3... are also true. □

2.2 Inter-arrival time

Now we will considered about Inter-arrival time in a Poisson distribution. So firstly let's define a sequence of inter-arrivals times as follow:

Definition 2.4. Considering a Poisson process, and that X_1, denotes the time of the first event. and for $n \geq 1$ let X_n be the difference of the n^{th} term and $(n-1)^{th}$ term. This sequence $\{X_n, n \geq 1\}$ is called a sequence of inter-arrivals times.

Proposition 2.5. *The sequence X_n, for $n = 1, 2, 3...$ are distributed identically, independently with exponential random value and mean $\frac{1}{\lambda}$*

Proof. Now we shall determine the distribution that this sequence X_n follows. Starting we note that for the event $\{X_1 > t\}$ to exist happens if and only if there are no events of the Poisson process in the interval $[0, t]$. This means:

$$P\{X_1 > t\} = P\{N(t) = o\} = e^{-\lambda t},$$

This means that X_1 is distributed as an exponential distribution with mean $\frac{1}{-\lambda}$
Now lets try with the next event so X_2,

$$P\{X_2 > t \| X_1 = s\}$$

This means that there are no events in the interval $(s, s+t]$. In a mathematical notation it means:

$$P\{0 \text{ events in}(s, s+t] \| X_1 = s\},$$

but because of the propriety of the independent increments, we can ignore the that $X_1 = s$ and applying stationary increments, we end up with the following:

$$P\{0 \text{ events in}(s, s+t]\} = e^{-\lambda t}.$$

therefore it is independent of x_1 and is distributed the same. and if we keep doing this process over and over we will see that they are all independent and distributed with an exponential distribution with mean $\frac{-1}{\lambda}$. □

This proposition obviously follows since its proprieties. All the points are independent from what happened before because of independent increments, and it will have the same distribution has the first one because of stationary increments. This means that this process has no memory and therefore proposition 2.5 follows.
Let's give another interesting definition about this subject:

Definition 2.6. Let's define a sequence S_n has the arrival time of the n^{th} term, where

$$S_n = \sum_{i=1}^{n} X_i, \text{for } n > 1.$$

Now using proposition 2.5 we can prove that the sequence S_n has a gamma distribution, we could have proofed in at least two other ways but this is the easiest of the all.

Proof. Since proposition 2.5 implies that S_n is distributed by a Gamma distribution with parameters n and λ, its probability density is:

$$f(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}, t \geq 0$$

□

Because of this result we have another way of defining a Poisson process. Supposing that we have a sequence like in proposition 2.5 and defining the counting process that n^{th} occurs at time S_n has defined above $S_n = X_1 + X_2 + \ldots + X_n$
This means that the resulting counting process $\{N(t), t \geq 0\}$ is a Poisson with rate λ

3 Non-homogeneous Poisson Process

As well as the HPP defined in the previous section, there is also a non-homogeneous Poisson process. From [6], a NHPP can be defined as;

Definition 3.1 (Non-homogeneous Poisson Process)**.** The counting process $\{N(t), t \geq 0\}$ is said to be a non-stationary or non-homogeneous Poisson process with intensity function $\lambda(t), t \geq 0$ if:

I $N(0) = 0$

II $\{N(t), t \geq 0\}$ has independent increments

III $P\{N(t+h) - N(t) \geq 2\} = o(h)$

IV $P\{N(t+h) - N(t) = 1\} = \lambda(t)h + o(h)$

To elaborate on this definition in simpler terms; a NHPP is one where λ does not take a constant value for the entire period. Instead, it is defined by $\lambda(t)$. Thus, the counting process for $N(t+h) - N(t)$ is distributed as $Poisson(\int_t^{t+h} \lambda(\alpha)d\alpha)$. Note that because of this, the intensity function $\lambda(t)$ must be integrable - meaning that the function will be both bounded and continuous for the entire time period.

Much like the HPP, the conditions for NHPP arise from the definition of the Poisson distribution and a counting process. (i) $N(0) = 0$ follows from the fact that if no time has passed, an event cannot have occurred. Independence of time periods is obvious. The proof for (iii) and (iv) also follow like the proofs for HPP (iii) and (iv), where λ is replaced with $\lambda(t)$.

Proposition 3.2. *$N(t+h) - N(t)$ is Poisson distributed with mean* $\int_t^{t+h} \lambda(\alpha)d\alpha$

Proof. This proof will be similar in part to our proof of Definition 2.3, however due to the nature of a NHPP, it will be changed slightly. First, let t be a set time and define a probability $P_n(h)$ such that,

$$P_n(h) = P(N(t+h) - N(t) = n)$$

Therefore, $P_0(h+r)$ will be equal to,

$$P_0(h+r) = P(N(h+t+r) - N(t) = 0)$$

This can be split into P(No events in period (t) to $(t+h)$) and P(No events in period $(t+h)$ to $(t+h+r)$). This follows from the independence of increments of time (Definition 3.1(II)). Therefore, using (III) and (IV) - which imply the definition of $P_0(t)$ - can be expressed in the form,

$$P_0(h+r) = P_0(h) \cdot [1 - \lambda(t+r)h + o(h)]$$

Hence, this can be rearranged to give

$$\frac{P_0(h+r) - P_0(h)}{h} = -\lambda(t+r)P_0(h) + \frac{o(h)}{h}$$

If we take a limit of $h \to 0$, this will yield a differential that we can integrate to give

$$\log P_0(r) = -\int_0^r \lambda(t+u)du$$

The obvious result being that this can be set as a power of the exponential and thus $P_0(r)$ (the probability of zero events in a period r) is Poisson distributed with mean $\int_0^r \lambda(t+u)du$. Therefore, our original time period, $P_0(t+h)$, is Poisson distributed with mean $\int_t^{t+h} \lambda(t+u)du$. This can be proved similarly for other values of n.

□

4 Applications

The Poisson distribution can be applied to many variables in the real world and these can be sub-categorised tidily as 'time orientated', and 'space orientated' variables, [3]. In general, time orientated variables are more popular and prolific in example numbers than space orientated variables.

4.1 Time Orientated

When using the Poisson distribution to model situations, or in this case, the number of events that occur within a certain time frame, the underlying assumptions are:

- the mean (μ: the number of events that are expected to occur within the time frame), or rate is constant
- all the events are independent from one another.

Further, the Poisson process, in addition to being derived from by letting $n \to \infty$ in the Binomial distribution law, can be obtained from "a series of exponentially distributed random variables, $S_n = X_1 + X_2 + ... + X_n$" [2], which yields the random variable $N(t)$, with each $X \in S_n$ to have equal Exponential distribution and such that $N(t) = max\{n : S_n \leq t\}$, to have a Poisson distribution, i.e. μ relying on t, which allows the variable to be varied over differing lengths of time period. This allows this distribution to be a versatile tool usable in many circumstances in actuality.

Some time orientated examples of the Poisson distribution application (obeying the aforementioned underlying assumptions) may be seen as, the number of drunk people entering a MacDonald's over a night period, the number of calls to a radio competition phone line within the program length, or the number of leaves to fall off a tree within an hour. Unlike with statistical computation of data, time orientated real world applications of the Poisson distribution may be subject to seasonality, therefore the expected rate of events might only remain constant for a certain amount of time. This might result in irregular results if collection of such data is continued over such 'seasons' as pattern in human behaviour may change with the hours of available daylight for example.

4.2 Space Orientated

For the use of Poisson distribution, these variables are governed by the same underlying assumptions as detailed above for the time orientated variables, however instead of events being counted over a set time period, they're recorded from a place in space, i.e. on an object or place of particular size. Examples of such variables include:

- number of mistakes found in a specified length of code.
- number of sesame seeds found on a single bagel.
- number of hits by V-1 buzz bombs in WWII London [7].

This final example highlights how such applications of the Poisson distribution can help plan for or model the fall out of catastrophic events (for example, air raids during a world war). Calculations and method completed in [2] show that the actual number of V-1 buzz bombs that hit London in one raid, is incredibly close to the result that is yielded as the expected number by the Poisson distribution.

The Waiting Time Paradox

For the average person in a metropolitan area there are many benefits: great food; diverse cultures and a wide array of events and parties. However, these positives are hampered at times by public transport and, in particular, buses. Companies like Stagecoach make bold claims like there is a bus 'every ten minutes'. Yet, for most it would seem that you can wait for hours at the bus station. As a mathematician, I'm sure you're already considering how we can understand the deeper underlying pattern.

The waiting time paradox helps us to understand what is going on here. If the buses you are using arrive exactly on time every given interval, then your average waiting time will be half that interval. However, due to the small joys of life, such as roadworks and breakdowns, we can confidently say that public transport rarely runs on time. However, if we set up some assumptions and begin to look into this further, VanderPlas says in [8] that; “When waiting for a bus that comes on average every ten minutes, your average waiting time will be ten minutes.”

Explanation

Let us first set the scene of what is happening:

- Buses pass a given point on the road following a Poisson process.
- Average time between buses is a constant.
- Passengers arrive at the same point of road at a random instant of time.

If we were to split up a day into varied inter-arrival intervals for the bus, then it is clear that the probability of a passenger arriving at the bus stop during a longer interval is greater than that of a passenger arriving during a short interval. We must assume that, within each inter-arrival interval, the arrival of the passenger to the fixed point on the road can be modeled with a uniform distribution. This implies that the expected waiting time should be half of the duration of the interval. However, given that the long intervals occur more frequently than the short ones, the overall average waiting time will be larger. This is represented visually in Figure 1 taken from [4].

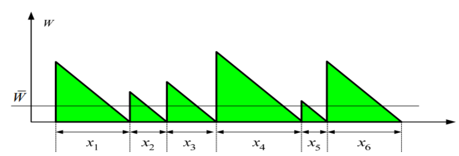

Figure 1: Sawtooth Graph of Waiting Times

However, we must ask ourselves, how accurate is the waiting paradox? Whilst it provides us with some explanation to why our bus always seems to be late, we can not necessarily put too much weight on the result. For example, arrival time might not be random, nor is it always uniformly distributed. For example, more people will turn up during rush hour than during the day. Also, buses do not run randomly. They aim to run on a vigorous schedule. It is doubts like these that lessen the credibility of the waiting time paradox.

When it comes to catching a bus, you’re better off arriving early and not basing your schedule on this paradox.

5 Simulations

Now that we’ve defined the Poisson process, observed some real applications of it, and defined the non-homogeneous variant; we can now simulate how this process behaves in practice. In this section we’ll give examples of both homogeneous and non-homogeneous situations by making simulations in R. These simulations will consider values on the positive half line such as time, (so we will be using the counting process, denoted $\{N(t), t \geq 0\}$).

NOTE: All of the simulations in this section were made using the “Poisson” package found here: [1]. This package uses the ’thinning’ technique to simulate the Poisson point process.

5.1 Homogeneous Simulations

First, recall that the homogeneous Poisson process has a constant value for it's intensity function, this in-turn means that when an event occurs between two points in time it is uniformly distributed. Our intensity function is simply: $\lambda(t) = 1$, and so the mean rate of events at time T is then:

$$E[N(T)] = \lambda \int_0^T \lambda(t)dt = \lambda \tag{5.1}$$

Now we'll use an example with actual values to showcase the HPP; say that we wanted to model how many cars n pass a point on a road in exactly $T = 15$ minutes. If we know that on average, 5 cars pass by in 1 minute, then we take $\lambda = 5$ where 1 minute is a single 'unit' in time t. Before we do any simulations, we know that the expected outcome is 75 minutes since, using (5.1), $E[N(15)] = 5*15 = 75$. However, the inter-arrival times follow an exponential distribution so there's of course going to be deviations from this expected value. Figure 2 shows 100 simulations of this scenario[1], where a single simulation would be the act of recording the amount of cars that pass by until 15 minutes have passed. From the plot, we can see that our expected outcome $E[N(t)] = 75$ becomes less accurate over time (however all the points still gravitate around our $n = \lambda t$ line). The cone-like shape comes from the variability of the events in each minute building up, (which is 5 as they follow a Poisson distribution) this means the simulations fan outward as time passes and more events occur.

5.2 Non-homogeneous Simulations

Simulating the NHPP is similar, however the time at which events occur is not uniformly distributed, they're unevenly distributed according to a given intensity function. In the context of the counting process, instead of counting from $t = 0$ up to a specified time $t = T$, we can start counting at any point in time and then stop at a specified time. This form of the NHPP can be used to model scenarios where the intensity changes values as time goes on by simply applying multiple processes together, the following examples shows this in practice as the intensity function changes at a particular point in time.

Let's say we want to model the amount of cars n that pass a point on the road again. However this time around, the rate at which cars pass by is slow while t is small and larger as t increases over the course of 80 minutes to simulate rush hour traffic building up. We'll assume that we know the following intensity function for the rate of cars passing in the next hour, this is $\lambda(t) = min(t/60, 1)$ with $\lambda = 5$, so, again using (5.1), $E(N(T)) = \lambda \int_0^T \lambda(t)dt = \frac{5T^2}{120}$ for the first 60 minutes, and then becomes $E(N(T)) = 5T - 150$ after that. So as time passes, we expect to see the amount of cars passing by the to increase for each minute that we wait up to 60 minutes, after that the rate will stay constant. Figure 3 shows this intensity function as a plot, and Figure 2 is a plot of 100 simulations for this scenario. Figure 2 allows us to see clearly how the change of intensity affects n as time passes, for example we can indeed see that the simulations have fewer events when t is small and more as t increases. As a comparison to the HPP, a vertical line at $T = 15$ is on the plot to see the difference in rates from one type of process to another. Here, n is expected to be roughly $5\int_0^{15} \frac{t}{60}dt = \frac{5*15^2}{120} = 9.375$ cars, which is drastically smaller than our previous example, as to be expected considering the intensity is lower at the 15 minute mark in this scenario.

[1] 100 simulations is a reasonable amount to perform given our small λ and T.

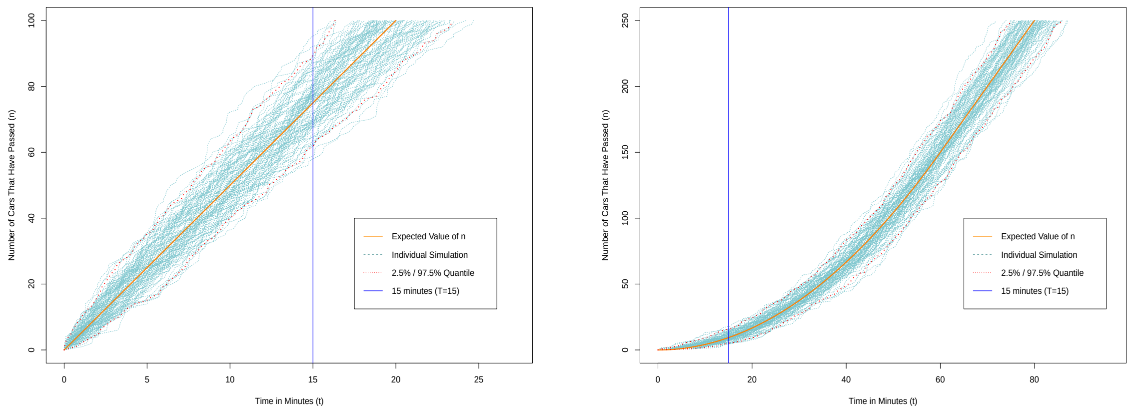

Figure 2: Left Plot: 100 homogeneous realisations obtained by the model described in 5.1 where the intensity function is $\lambda(t) = 1$ where $\lambda = 5$. Right Plot: 100 non-homogeneous realisations obtained by the model described in 5.2 where the intensity function is $\lambda(t) = min(t/60, 1)$ where $\lambda = 5$. In both plots we consider a single unit of time to be 1 minute.

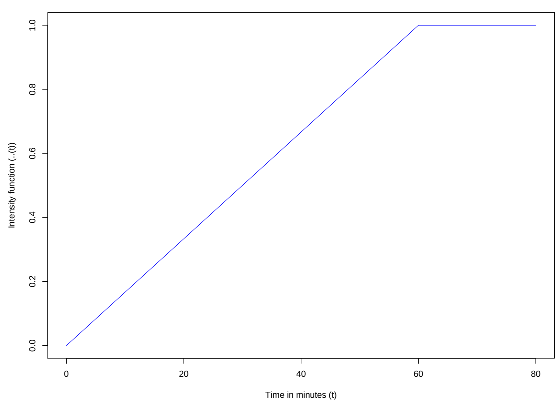

Figure 3: Simple plot showing the intensity function for our NHPP scenario where $\lambda(t) = min(t/60, 1)$, after 60 minutes the rate stops increasing and remains constant.

6 Conclusion

We have seen that the Poisson process can be applied to many real life scenarios such as the number of busses to arrive in a given time or the number of mistakes in a specified length of code. In addition, the same process can be applied to most situations in which events follow a Poisson distribution with independent incriments of space or time and either a fixed rate parameter λ as in the HPP or with a rate parameter that varies as a function of time $\lambda(t)$ as in the NHPP. From Section 5 we saw that it is possible to model both the HPP and NHPP using computer simulations with realisations having mean and variance $\lambda(t)T$, resulting in both higher mean and more varied outcomes as more time passes or, in the case of the NHPP, mean and variance change in proportion to λ as it varies with time.

References

[1] K Brock. Package "poisson". https://cran.r-project.org/web/packages/poisson/index.html, 2015. 9

[2] W Feller. *An Introduction to Probability Theory and Its Applications, Vol. 2.* Boston, MA: Addison Wesley, Pearson Education, Inc., 2007. 8

[3] J Letkowski. Applications of the poisson probability distribution, n.d. 7

[4] Ben-Gurion University of the Negev. Poisson process. http://www.cse.bgu.ac.il/common/download.asp?FileName=PoissonProcesss.pdf&AppID=2&MainID=497&SecID=4329&MinID=3, 2011. 9

[5] S. D Poisson. *Probabilité des jugements en matière criminelle et en matière civile,précédées des règles générales du calcul des probabilitiés.* Rue de Jardinet, 1837. 2

[6] S M Ross. *Stochastic Processes.* John Wiley & Sons, 2nd edition, 1996. 6

[7] M. F Triola. *Elementary Statistics Using Excel®.* NY: John Wiley & Sons Inc., 1966. 8

[8] J VanderPlas. The waiting time paradox, or, why is my bus always late? https://jakevdp.github.io/blog/2018/09/13/waiting-time-paradox, 2018. 9