Project

Gifty Osei, Carol, Patrick

11/9/2021

1 POISSON AND EXPONENTIAL DISTRIBUTION (WAITING TIME)

1.1 INTRODUCTION

how so?

Probability is a measure of uncertainty, like a logical argument. Probability distributions are common statistical distributions used in modelling real life incidents. For every probability distribution, we have the mean, variance and other useful descriptive measures that aid in understanding interesting relationships. The Poisson distribution is a widely used discrete distribution in modeling count data. For example, modeling spatial distributions and the number of occurrences in a given time interval. In modeling lifetime data analysis problems the most widely used distribution is the exponential distribution and we see that ~~with~~ in Cooper (2005).

1.2 CONCEPT

A continuous random variable is a random variable which can take any value in some interval and be characterized by ~~the~~ a probability density function, a graph which has a total area of 1 , more to def. beneath it (The probability at a point is essentially 0). A Poisson distribution deals with the number of occurrences in a fixed period of time and an exponential distribution deals with the time between occurrences of successive events as time flows by continuously.

The Poisson distribution has a probability mass function

$$f_X(x) = \frac{e^{-\mu}\mu^x}{x!} \tag{1}$$

where, x = 0,1,2,...

The exponential distribution has a probability density function;

$$f_T(t) = \lambda e^{-\lambda t} \tag{2}$$

where $t \leq 0$ and $\lambda > 0$

1.3 APPLICATION OF ARRIVAL AND INTERARRIVAL TIMES USING POISSON AND EXPONENTIAL DISTRIBUTIONS

To look at the relationship between the exponential distribution and the Poisson distribution, we must first understand a Poisson process. As we will see, the intervals of a Poisson process can be modeled with an exponential distribution. So, let us start at step one and understand what a Poisson process is. A Poisson process is used to model events that happen randomly, but at a certain rate. An example of phenomena that Poisson distributions can model are earthquakes and spatial distributions. While earthquakes appear to occur randomly, there is historical data that gives us a rate of occurrence within a certain time period. To be considered a Poisson process, there are three requirements. Firstly, at time zero there cannot have been any events yet. Secondly, all time intervals must be independent of each other. Thirdly, the Poisson process must follow a Poisson distribution, but with an extra rate parameter.

1.3.1 POISSON PROCESS

To understand the origins of a Poisson process, it can be quite helpful to begin by approximating it with a binomial distribution. Suppose a researcher knows that an average of 6 earthquakes in a year of magnitude ≥ 8 (hereafter considered high-magnitude) and has reason to believe the odds of an earthquake occurring are stable from moment to moment. They might approximate the number of earthquakes in a year by treating each month as an independent Bernoulli trial in which an earthquake might occur with 50% odds. Then if K is the (approximate) number of earthquakes that occurred in a year, $K \sim Binom(12, \frac{6}{12})$. There's a problem with this approximation, though. There's nothing really bounding the number of high-magnitude earthquakes to 12, nor is there any reason more than 1 high-magnitude earthquake couldn't occur in a single month. The researcher, therefore, decides to go deeper. Instead of modeling months, she models by day. Following from our previous work, $K_{day} \sim (365, \frac{6}{365})$. Again, however, we run into the problem of fixing the maximum number of possible earthquakes to 365, and the number of earthquakes that can occur in a day to 1. The researcher, understanding that this is a race to the bottom, now considers the outcome as the number of trials approaches infinity and the length of time between trials approaches 0. The general PMF for a binomial distribution is $\binom{n}{k}p^k q^{n-k}$. For the earthquake model, p is always $\frac{6}{n}$ (this is important because the mean of a binomial distribution is np, so the n's will cancel and the mean will always be 6). Let's generalize this by letting $6 = \lambda$. Then, the researcher is trying to find $\lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$. Simplifying the expression, she sees that

haven't abbreviated anything until here

$$= \lim_{n\to\infty} \frac{n!}{(n-k)!k!} \frac{\lambda^k}{n^k} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k}$$

$$= \lim_{n\to\infty} \frac{\prod_{i=0}^{k-1}(n-i)}{k!} \frac{\lambda^k}{n^k} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k} \qquad \frac{n!}{(n-k)!} = n\cdot(n-1)...(n-(k-1)) = \prod_{i=0}^{k-1}(n-i)$$

$$= \lim_{n\to\infty} \frac{\prod_{i=0}^{k-1}(n-i)}{n^k} \frac{\lambda^k}{k!} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k}$$

Largest term in $\prod_{i=0}^{k-1}(n-i)$ is going to be n^k ?

combine on 1 line

$$= 1\cdot \lim_{n\to\infty} \frac{\lambda^k}{k!} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k}$$

As n approaches ∞, first term approaches $\frac{n^k}{n^k} = 1$

$$= e^{-\lambda}\frac{\lambda^k}{k!} \cdot \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{-k} \qquad \lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = e^a$$

$$= e^{-\lambda}\frac{\lambda^k}{k!}(1-0)^{-k}$$

$$= e^{-\lambda}\frac{\lambda^k}{k!}$$

hadn't abbreviated before now

This is the general PDF of a Poisson random variable and the rationale behind it. So, if there ~~are~~ is an average of 6 earthquakes in a year and every infinitesimal moment has an equal probability of producing an earthquake, then the probability of k earthquakes in a year $\forall k \in \mathbb{Z}, k \geq 0$ is $e^{-6}\frac{6^k}{k!}$.

Let X represent the time until a new event occurs and K_x represent the number of events in time period x . Then $P(X > x) = P(K_x = 0)$. By the complement rule we also see that $P(X \leq x) = 1 - P(K_x = 0)$. Noting that $K_x \sim Pois(\lambda x)$, why? $P(K_x) = \frac{(\lambda x)^0}{0!}e^{-\lambda x} = e^{-\lambda x}$. So $P(X \leq x) = 1 - e^{-\lambda x}$. This is the CDF of an exponential distribution with *rate* parameter λ.[1]

define $X \sim Exp$

So, if the Poisson process is modeling the time at which an event occurs, then the exponential is modeling the time between arrivals. I think this is easier to understand visually.

we (group paper)

[1] $X \sim Exp(\frac{1}{\lambda})$ in our ~~class~~ case

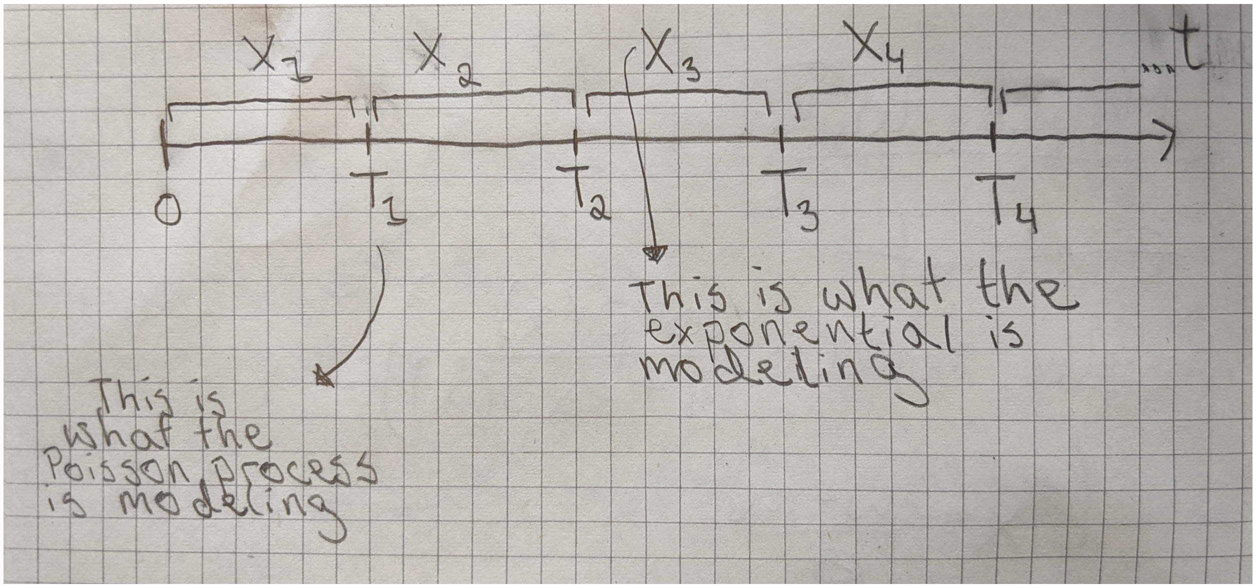

Figure 1: How to Visualize Exponential (Pishro Nik (n.d.))

Is this drawing, the T_i's are the times that events occur along the timeline. The X_i's are the inter-arrival times. Importantly, if the Poisson process has rate λ then the exponential will also have rate λ.

1.4 ARRIVAL AND INTERARRIVAL TIMES FOR SCHEDULING OF APPOINTMENTS (FIRST-COME-FIRST-SERVE)

For every system or company in the customer services field, it is very important to have a system time scheduler for appointments such that employees are not dormant nor are customers spend ing more time than the have to. Let T_x be the inter-arrival time between customers x and no hyphen $x + 1$ and W_x be the arrival time / service time of customer x. Assuming that the service time

$$S_{x+1} = P_x A_x + B_x \tag{3}$$

where, A_x, P_x and B_x are 3 sequences of independent and identically distributed (i.i.d) non-negative random variables and are independent of each other. These ~~This~~ inter-arrival and arrival (service) times have been found in communication systems, transportation systems, producwhy?

tion processes and mainly in the customer service industries like hospitals and banks. P_x is constant, B_x is zero and A_x is exponentially distributed ($A_x \sim exp(\lambda)$). Dai and Hu (2021) explored more on how ~~this~~ these two things correlate.

include in bulleted list

Other applications for inter-arrival times and arrival times are;

- Extensions and modification of the queue for customer collection on communication systems.
- Performance measures
- Traffic systems using the Markovian Arrival Processes (MAP's) that captures statistics of any order of inter-arrival times Casale, Zhang, and Smirni (2008).
- The Inter-arrival times using the weibull distribution and other count modelling McShane et al. (2008)

just said $A_x \sim Exp$??

Other studies also show that the S_{x+1} and A_x have a bivariate exponential distribution, (Downton's) that is the inter-arrival time and the arrival time is given ~~be~~ as a mixture of joint density, distributions.

REFERENCES

Casale, Giuliano, Eddy Z. Zhang, and Evgenia Smirni. 2008. "Interarrival Times Characterization and Fitting for Markovian Traffic Analysis." In *Numerical Methods for Structured Markov Chains*, edited by Dario Bini, Beatrice Meini, Vaidyanathan Ramaswami, Marie-Ange Remiche, and Peter Taylor. Dagstuhl Seminar Proceedings 07461. Dagstuhl, Germany: Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany. http://drops.dagstuhl.de/opus/volltexte/2008/1390.

Cooper, John CB. 2005. "The Poisson and Exponential Distributions." *Mathematical Spectrum* 37 (3): 123–25.

Dai, Weimin, and Jian-Qiang Hu. 2021. "Correlated Queues with Service Times Depending on Inter-Arrival Times." *Queueing Systems*, 1–20.

McShane, Blake, Moshe Adrian, Eric T Bradlow, and Peter S Fader. 2008. "Count Models Based on Weibull Interarrival Times." *Journal of Business & Economic Statistics* 26 (3): 369–78.

Pishro Nik, Hossein. n.d. "Basic Concepts of the Poisson Process." *Introduction to Probability, Statistics, and Random Processes.* https://www.probabilitycourse.com/chapter11/11_1_2_basic_concepts_of_the_poisson_process.php.