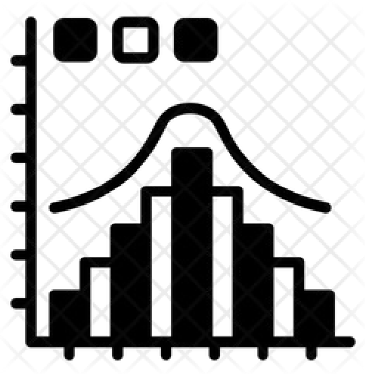
Poisson Distributions

Tushar B. Kute,
http://tusharkute.com

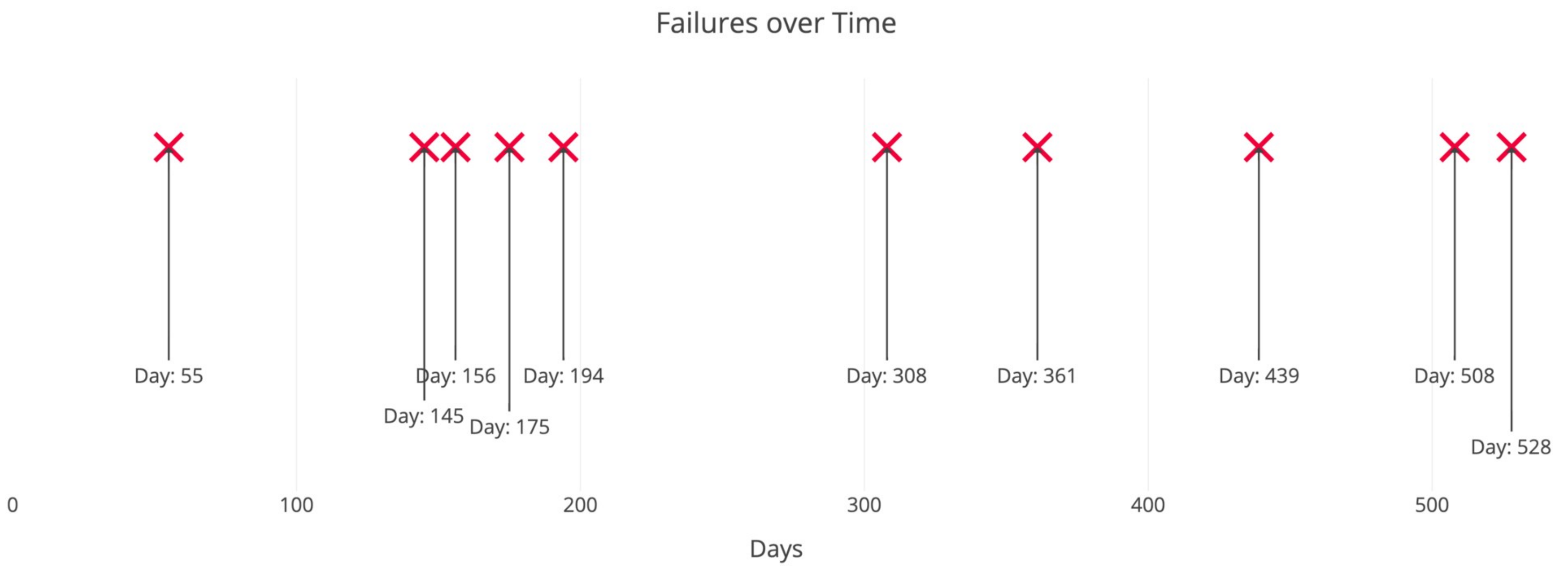


Poisson Process

- A Poisson Process is a model for a series of discrete event where the average time between events is known, but the exact timing of events is random.
- The arrival of an event is independent of the event before (waiting time between events is memoryless).
- For example, suppose we own a website which our content delivery network (CDN) tells us goes down on average once per 60 days, but one failure doesn't affect the probability of the next.
- All we know is the average time between failures.

Poisson Process

- This is a Poisson process that looks like:



Poisson Process

- The important point is we know the average time between events but they are randomly spaced (stochastic). We might have back-to-back failures, but we could also go years between failures due to the randomness of the process.
- A Poisson Process meets the following criteria (in reality many phenomena modeled as Poisson processes don't meet these exactly):
  - Events are independent of each other. The occurrence of one event does not affect the probability another event will occur.
  - The average rate (events per time period) is constant.
  - Two events cannot occur at the same time.

Poisson Process

- The last point — events are not simultaneous — means we can think of each sub-interval of a Poisson process as a Bernoulli Trial, that is, either a success or a failure.
- With our website, the entire interval may be 600 days, but each sub-interval — one day — our website either goes down or it doesn’t.

Poisson Process

- Common examples of Poisson processes are customers calling a help center, visitors to a website, radioactive decay in atoms, photons arriving at a space telescope, and movements in a stock price.
- Poisson processes are generally associated with time, but they do not have to be.
- In the stock case, we might know the average movements per day (events per time), but we could also have a Poisson process for the number of trees in an acre (events per area).

Poisson Distribution

- The Poisson Process is the model we use for describing randomly occurring events and by itself, isn't that useful.
- We need the Poisson Distribution to do interesting things like finding the probability of a number of events in a time period or finding the probability of waiting some time until the next event.

Poisson Distribution

- The Poisson Distribution probability mass function gives the probability of observing k events in a time period given the length of the period and the average events per time:

$$P(k \text{ events in time period}) = e^{-\frac{\text{events}}{\text{time}} * \text{time period}} * \frac{(\frac{\text{events}}{\text{time}} * \text{time period})^k}{k!}$$

Poisson Distribution

- This is a little convoluted, and events/time * time period is usually simplified into a single parameter, λ, lambda, the rate parameter.
- With this substitution, the Poisson Distribution probability function now has one parameter:

$$P(k \text{ events in interval}) = e^{-\lambda}\frac{\lambda^k}{k!}$$

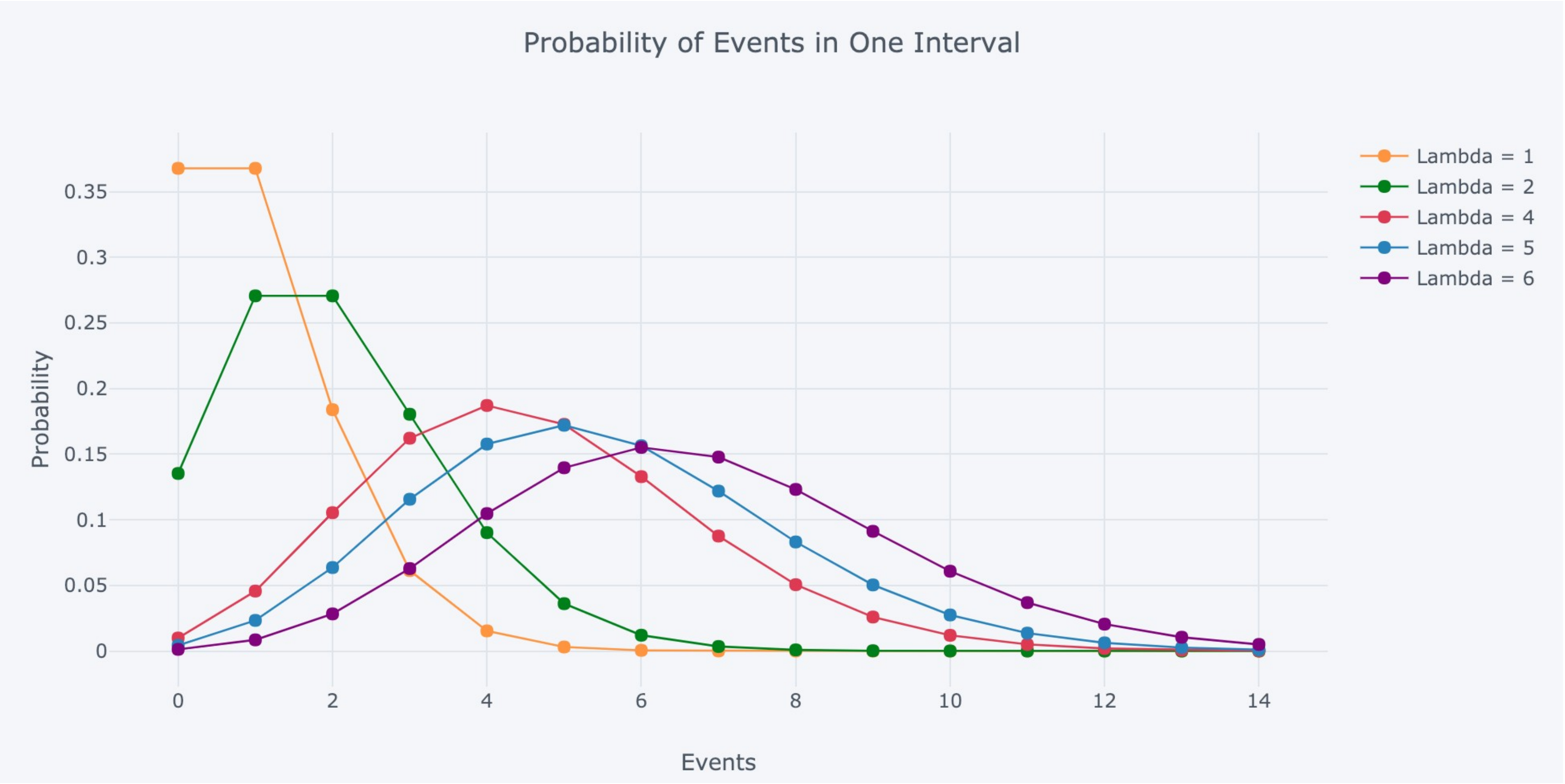
Poisson Distribution

- Lambda can be thought of as the expected number of events in the interval. (We'll switch to calling this an interval because remember, we don't have to use a time period, we could use area or volume based on our Poisson process).
- I like to write out lambda to remind myself the rate parameter is a function of both the average events per time and the length of the time period but you'll most commonly see it as directly above.

Poisson Distribution

- As we change the rate parameter, λ, we change the probability of seeing different numbers of events in one interval.
- The below graph is the probability mass function of the Poisson distribution showing the probability of a number of events occurring in an interval with different rate parameters.

Poisson Distribution



Probability Mass function for Poisson Distribution with varying rate parameter.

tusharkute.com

Poisson Distribution

- The most likely number of events in the interval for each curve is the rate parameter.
- This makes sense because the rate parameter is the expected number of events in the interval and therefore when it's an integer, the rate parameter will be the number of events with the greatest probability.
- When it's not an integer, the highest probability number of events will be the nearest integer to the rate parameter, since the Poisson distribution is only defined for a discrete number of events.

Poisson Distribution

- The discrete nature of the Poisson distribution is also why this is a probability mass function and not a density function. (The rate parameter is also the mean and variance of the distribution, which do not need to be integers.)
- We can use the Poisson Distribution mass function to find the probability of observing a number of events over an interval generated by a Poisson process.

Poisson Distribution: Example

- In my childhood, my father would often take me into our yard to observe (or try to observe) meteor showers.
- We were not space geeks, but watching objects from outer space burn up in the sky was enough to get us outside even though meteor showers always seemed to occur in the coldest months.

Poisson Distribution: Example

- The number of meteors seen can be modeled as a Poisson distribution because the meteors are independent, the average number of meteors per hour is constant (in the short term), and — this is an approximation — meteors don’t occur simultaneously.
- To characterize the Poisson distribution, all we need is the rate parameter which is the number of events/interval * interval length.

Poisson Distribution: Example

- From what I remember, we were told to expect 5 meteors per hour on average or 1 every 12 minutes.
- Due to the limited patience of a young child (especially on a freezing night), we never stayed out more than 60 minutes, so we'll use that as the time period. Putting the two together, we get:

$$\frac{1 \text{ meteor}}{12 \text{ minutes}} * 60 \text{ minutes} = 5 \text{ meteors expected} = \lambda$$

Rate parameter for the meteor shower situation.

Poisson Distribution: Example

- What exactly does “5 meteors expected” mean? Well, according to my pessimistic dad, that meant we’d see 3 meteors in an hour, tops.
- At the time, I had no data science skills and trusted his judgment. Now that I’m older and have a healthy amount of skepticism towards authority figures, it’s time to put his statement to the test.
- We can use the Poisson distribution to find the probability of seeing exactly 3 meteors in one hour of observation:

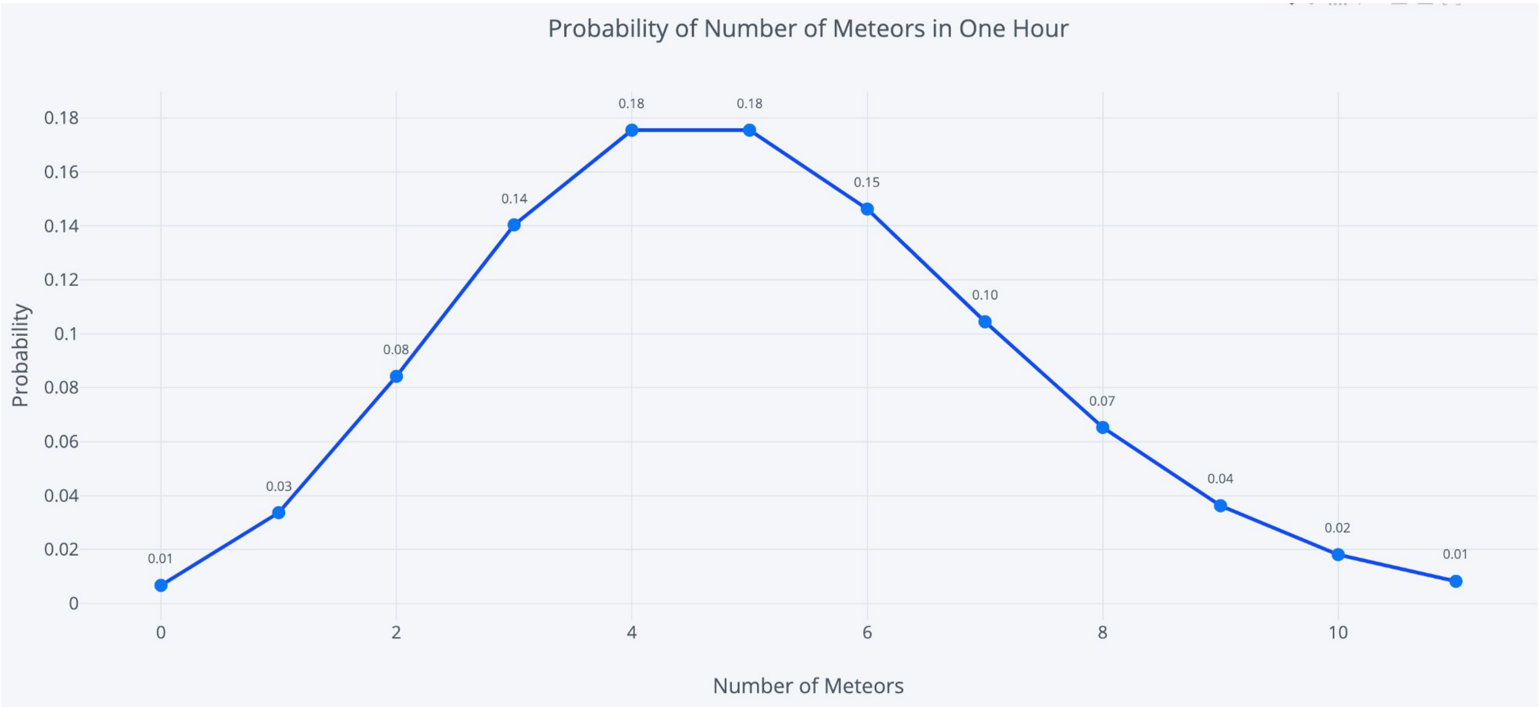
$$P(3 \text{ meteors in 1 hour}) = e^{-5}\frac{5^3}{3!} = 0.1404 = 14.0\%$$

Poisson Distribution: Example

- 14% or about 1/7. If we went outside every night for one week, then we could expect my dad to be right precisely once!
- While that is nice to know, what we are after is the distribution, the probability of seeing different numbers of meteors.

Poisson Distribution: Example

- The below graph shows the Probability Mass Function for the number of meteors in an hour with an average time between meteors of 12 minutes (which is the same as saying 5 meteors expected in an hour).



Poisson Distribution: Example

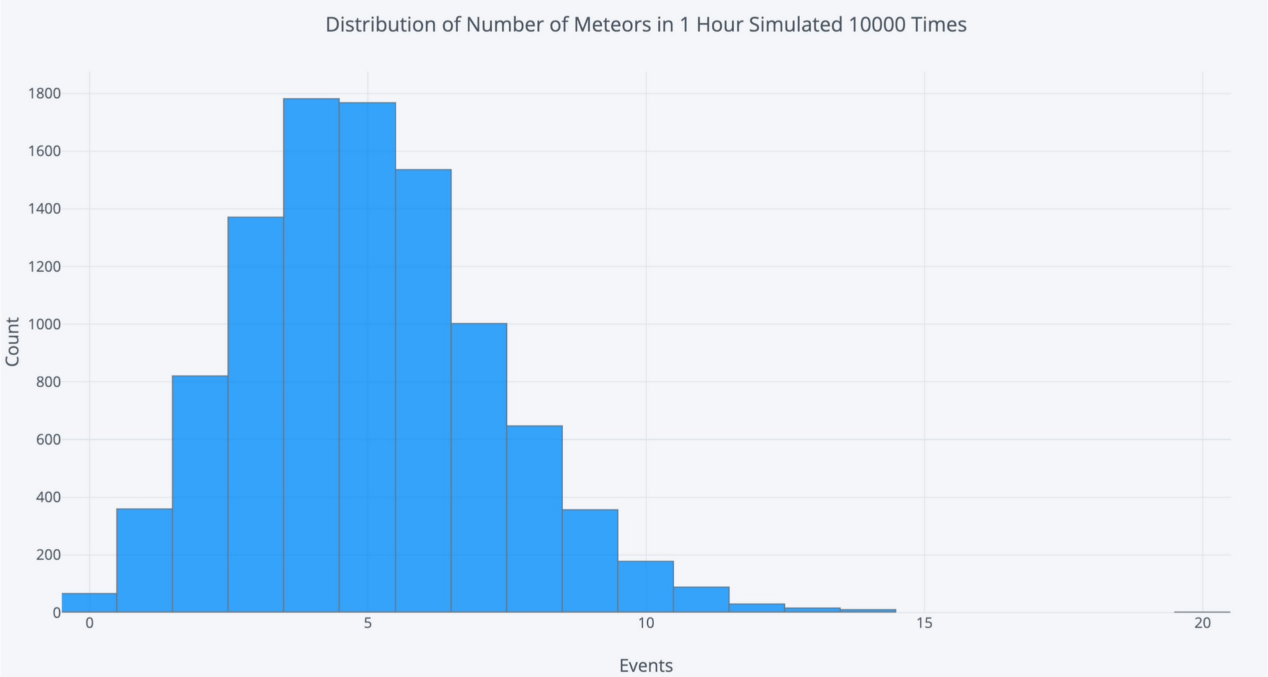
- This is what “5 expected events” means! The most likely number of meteors is 5, the rate parameter of the distribution. (Due to a quirk of the numbers, 4 and 5 have the same probability, 18%).
- As with any distribution, there is one most likely value, but there are also a wide range of possible values.
- For example, we could go out and see 0 meteors, or we could see more than 10 in one hour. To find the probabilities of these events, we use the same equation but this time calculate sums of probabilities

Poisson Distribution: Example

- We already calculated the chance of seeing exactly 3 meteors as about 14%. The chance of seeing 3 or fewer meteors in one hour is 27% which means the probability of seeing more than 3 is 73%.
- Likewise, the probability of more than 5 meteors is 38.4% while we could expect to see 5 or fewer meteors in 61.6% of observation hours.
- Although it's small, there is a 1.4% chance of observing more than 10 meteors in an hour!

Poisson Distribution: Example

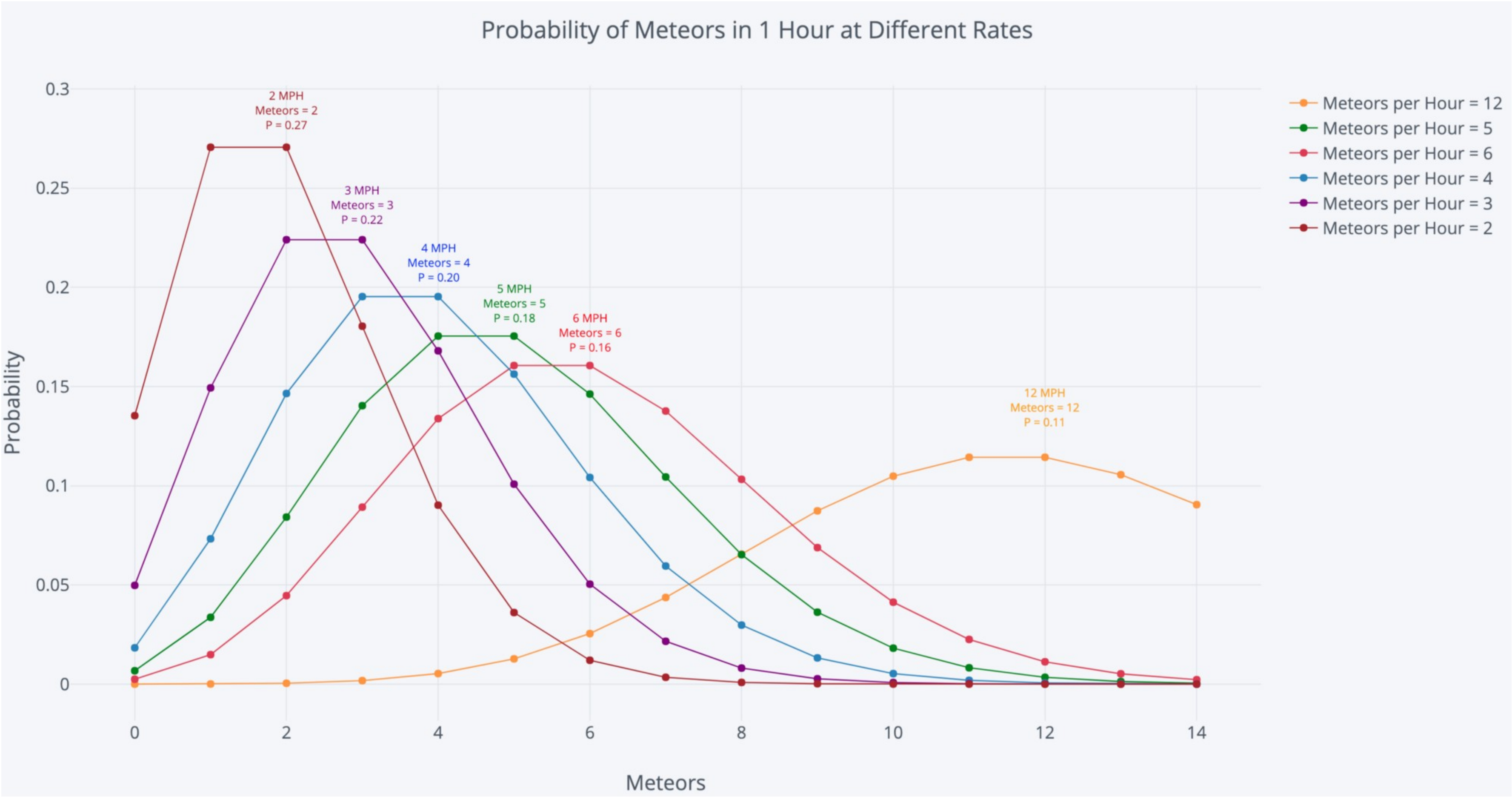
- To visualize these possible scenarios, we can run an experiment by having our sister record the number of meteors she sees every hour for 10,000 hours. The results are shown in the histogram below:



The rate parameter

- The rate parameter, λ, is the only number we need to define the Poisson distribution.
- However, since it is a product of two parts (events/interval * interval length) there are two ways to change it: we can increase or decrease the events/interval and we can increase or decrease the interval length.
- First, let's change the rate parameter by increasing or decreasing the number of meteors per hour to see how the distribution is affected.
- For this graph, we are keeping the time period constant at 60 minutes (1 hour).

The rate parameter

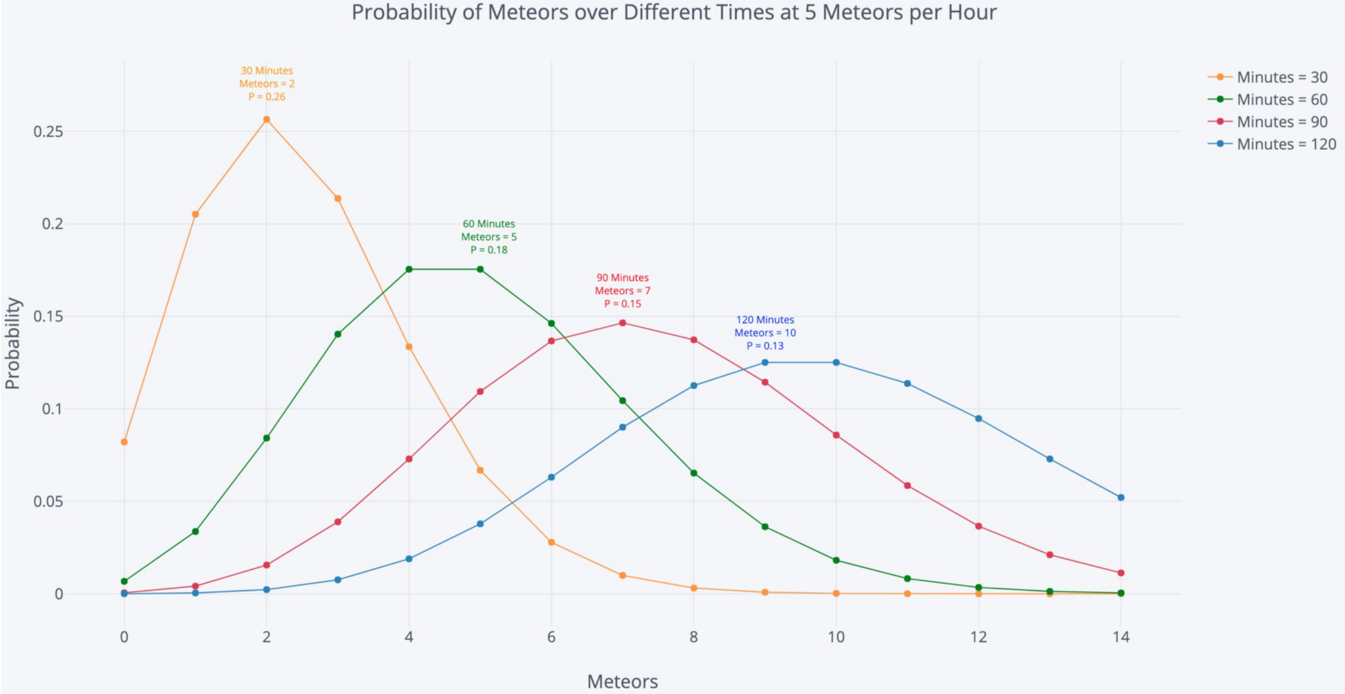


The rate parameter

- In each case, the most likely number of meteors over the hour is the expected number of meteors, the rate parameter for the Poisson distribution.
- As one example, at 12 meteors per hour (MPH), our rate parameter is 12 and there is an 11% chance of observing exactly 12 meteors in 1 hour.
- If our rate parameter increases, we should expect to see more meteors per hour.

The rate parameter

- Another option is to increase or decrease the interval length. Below is the same plot, but this time we are keeping the number of meteors per hour constant at 5 and changing the length of time we observe.



Poisson Distribution

- In each case, the most likely number of meteors over the hour is the expected number of meteors, the rate parameter for the Poisson distribution.
- As one example, at 12 meteors per hour (MPH), our rate parameter is 12 and there is an 11% chance of observing exactly 12 meteors in 1 hour.
- If our rate parameter increases, we should expect to see more meteors per hour.

Thank you

@mitu_skillologies

/mITuSkillologies

@mitu_group

/company/mitu-skillologies

MITUSkillologies

Web Resources
https://mitu.co.in
http://tusharkute.com

contact@mitu.co.in

tushar@tusharkute.com