05.Basic Probability Distribution

We might want our model to compute the probability of an event to happens.

Discrete Distributions

Finite discrete

Finite discrete distributions are encoded by tables that associate to each outcome the corresponding probability.

Geometric

The Geometric distribution is a discrete random variable, with infinite support starting from k = 0, that is characterized by a probability parameter q.

$$P(X_{Geom<q>} = k) = (1-q)^k q$$

$$F_{Geom<q>}(k) = 1 - (1-q)^{k+1}$$

If a success can be obtained with probability q, then $P(X_{Geom<q>} = k)$ corresponds to the probability of having exactly k failures before the positive event in a sequence of k+1 independent trials.

$$\text{average: } \mu_{Geom<q>} = \frac{1-q}{q}$$

$$\text{variance: } \sigma^2_{Geom<q>} = \frac{1-q}{q^2}$$

It is also important in one of the most important queue distribution system as it follows a geometric distribution

Poisson

The Poisson distribution is characterized by a parameter λ > 0, and has the following expression:

$$P(X_{Pois<\lambda>} = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$F_{Pois<\lambda>}(k) = e^{-\lambda} \sum_{i=0}^{k} \frac{\lambda^i}{i!}$$

$$mod$$

Discrete distributions in Performance Evaluation

In performance evaluation, we use the finite discrete distributions to encode the probability of making a specific choice from a set of alternatives.
Geometric and Poisson distributions are also found as the distribution of number of jobs in a queue for some special cases of single server systems.

Deterministic

A deterministic random variable is used to represent actions that always take the same time. Its parameter is the constant value τ it returns.

$$f_{Det<\tau>}(t) = \delta(t - \tau)$$

$$F_{Det<\tau>}(t) = I(t \geq \tau)$$

The average corresponds to its value, and the variance is zero.

$$\text{average: } \mu_{Det<\tau>} = \tau$$

It is used to model arrivals at fixed intervals. Deterministic service times represent actions that always take the same constant time.

Uniform

The Uniform distribution returns values equally distributed between two extremes a and b.

There aren't many real scenarios that can occur with an uniform distribution; the most common one can be the computational latency of the reading of a disk in a hard-drive. Use only if it really needed.

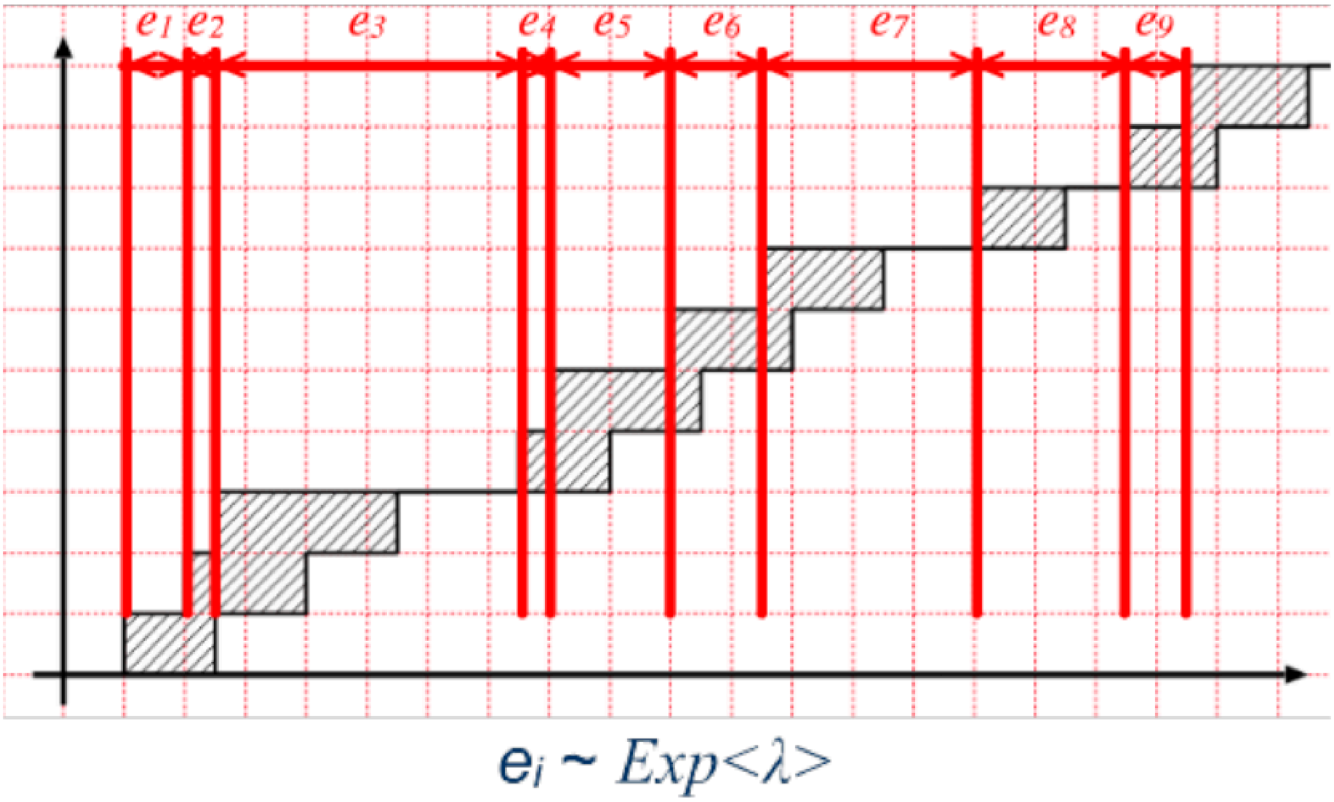
Exponential

The Exponential distribution is the most used in performance evaluation. It is characterized by a parameter that defines its rate. It is called exponential as its PDF or CDF have an exponential form.

It is characterized by the memory less property. It means that every time that the probability of the event occur it is independent from the probability at the beginning.
As we will see, models that are based on it, can be analyzed with simple computations in most of the cases.
Used in inter-arrival times, it indicates that the events can occur completely randomly and independently. Exponential service times model actions that are equally likely to finish at any time at a given rate (end of a call, end of a repair, ...).

Poisson process

When the time between events in a sequence is exponentially distributed, the corresponding process is called Poisson process.



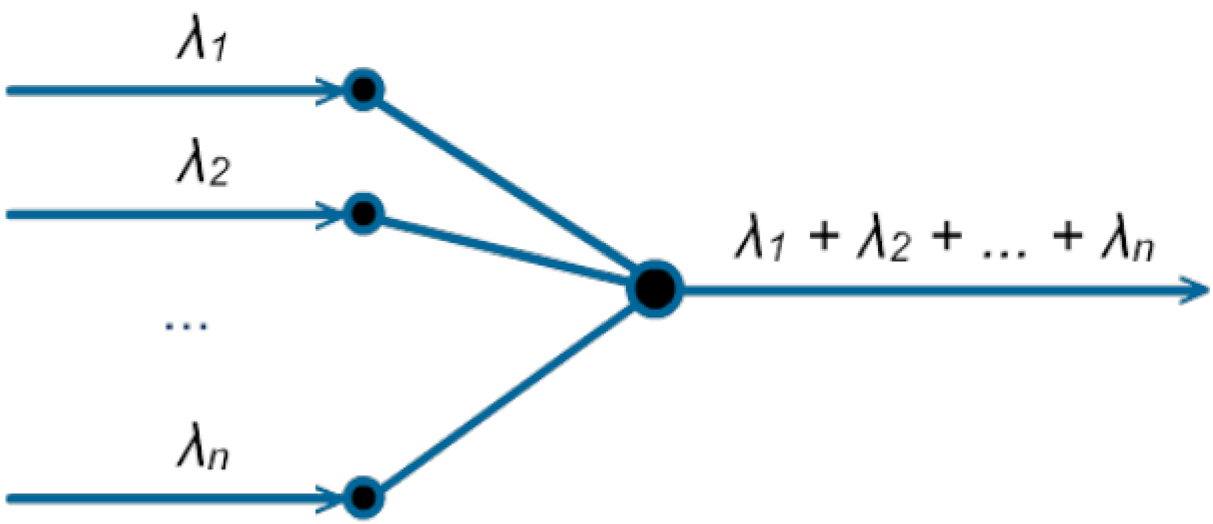
This is because, if we fix a time interval T, the number of events with an exponential inter-arrival time of parameter l that occur in that interval is distributed according to a Poisson distribution with parameter λT:

Usually Markovian arrivals (services) - Poisson arrivals (services) - exponential inter-arrival times (service times) are three different ways of addressinh the same concept. We have seen why exponential and Poisson, later we will see also why the term “Markovian” is used.

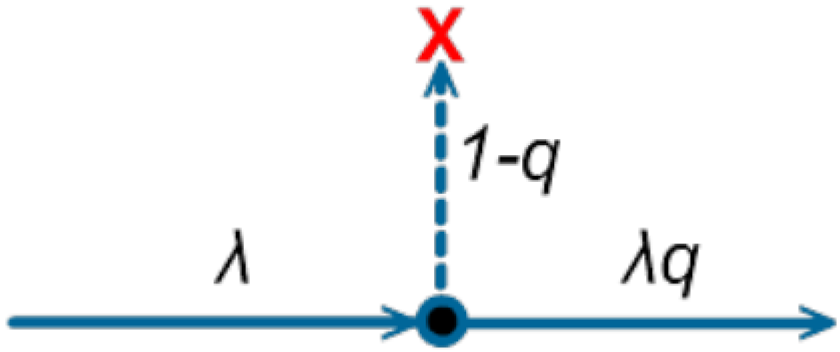
Properties of Poisson processes

The minimum of a set of exponential random variables is still exponential with a rate equal to the sum of the rates.

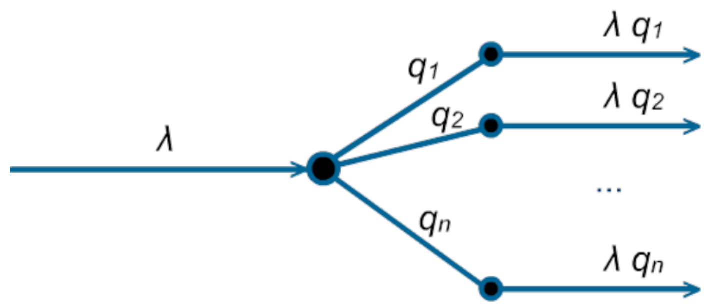
For this reason, if we combine together n Poisson processes at different rates, we obtain a new Poisson process whose parameter corresponds the sum of the individual rates.



If from a Poisson process characterized by rate λ we randomly discard samples with probability (1-q), the remaining events still occur according to a Poisson process, with a reduced parameter λ . q.



Repeating the previous rule, if we split a Poisson process into n fluxes with probability q1 ... qn , we obtain a set of independent Poisson process whose rate is the product between the original rate and the corresponding probability.



Hyper-exponential

The Hyper-exponential distribution corresponds to the combination of two or more exponential random variables that are selected according to a set of independent discrete probabilities.