Exercise

M servers are subject to attacks during a period of time T (for instance 1 year). Subdivide the interval T in N subinterval of size T/N and in each of this suppose that an attack can occur with probability λ T/N. Simulate the attacks to the M servers and represent each of them with a line which makes jumps of 1 at each attack event.

Using the same objects ("movable/resizable rectangle", histogram, etc.) of the previous homework 3 draw vertically on the line chart the 2 histograms representing the distribution of the number of attacks at the end of the period and one internal istant for comparison.

Study what happens asymptotically, for N large, and a number of systems M a sufficient to give shape to a simulated distribution. Make some personal considerations about the shape and the average of the distributions that you see.

To simulate the accumulation of a “security score” I wrote a function which create a random variable for M systems (between 0 and 10) and set fixed values for N attacks (in this case 400), the constant average rate of occurrence (in this case 100) and time interval T (in this case 1) such that the penetration probability p = λ T/N is very small. After, I simulated the score trajectory for each system and if the random probability that attack i happens is less than p, the system is penetrated, so the score trajectory is set to -1, otherwise it is set to 1.

For N large and specific parameters (λ, T, M) the shape of the histogram closely resembles a Poisson distribution, so the distribution represented in the exercise is a Poisson distribution. In fact, the Poisson distribution is appropriate when events are rare, independent, and occur with a constant average rate (λ) within a fixed interval (T/N). So, we can conclude that the assumptions of the simulation align with the conditions for a Poisson distribution that includes assuming that the attacks are rare, independent, and occur at a relatively constant rate.

Research

Find out on the web about a Poisson point process. See if you can see any analogy with this Exercise and verify whether your distributions come close (for N, M sufficiently large) to the theoretical asymptotic distribution.

In probability, statistics and related fields, a Poisson point process is a type of random mathematical object that consists of points randomly located on a mathematical space with the essential feature that the points occur independently of one another. The Poisson point process is also called a Poisson random measure, Poisson random point field or Poisson point field. When the process is defined on the real line, it is often called simply the Poisson process.

This point process has convenient mathematical properties, which has led to its being frequently defined in Euclidean space and used as a mathematical model for seemingly random processes.

Its name derives from the fact that if a collection of random points in some space forms a Poisson process, then the number of points in a region of finite size is a random variable with a Poisson distribution.

The point process depends on a single mathematical object, which may be a constant. In this case, the constant, known as the rate or intensity, is the average density of the points in the Poisson process located in some region of space. The resulting point process is called a homogeneous or stationary Poisson point process.

[Poisson point process - Wikipedia](https://en.wikipedia.org/wiki/Poisson_point_process)

In a Poisson point process, events (in this case the attacks) occur randomly and independently in time, with a constant average rate of occurrence (λ in the exercise) within a specified time interval (T). The interval is divided into N subintervals of equal size (T/N), and in each subinterval, an event (attack) can occur with a certain probability, λ \* T/N.

For these reasons the exercise has analogies with the Poisson point process in the sense of:

* Randomness: The attacks are random and occur with a probability λ during each subinterval of size T/N;
* Independence: the attacks on different servers are assumed to be independent.
* Constant Rate: There is a constant probability λ of an attack occurring in each subinterval.