Stochastic processes, the Poisson process

Javier Campos
Universidad de Zaragoza
http://webdiis.unizar.es/~jcampos/

Outline

- Why stochastic processes?
- The Poisson process
  - Exponential distribution
  - Properties of exponential r.v. and Poisson process

Etimología (extraído de Wikipedia)

Estocástico

Palabra proveniente del griego: στοχαστικός, que significa “hábil en conjeturar”.

Según el DRAE significa "perteneciente o relativo al azar“.

Se denomina estocástico a aquel sistema que funciona, sobre todo, por el azar. Las leyes de causa-efecto no explican cómo actúa el sistema (y de modo reducido el fenómeno) de manera determinista, sino en función de probabilidades.

En Investigación de Operaciones: Modelos Probabilísticos y Estocásticos es lo mismo.

Etimology (taken from wikipedia)

Stochastic

From greek: στοχαστικός, for “skilled in conjecture”.

According to Spanish Dictionary (DRAE), it means “relative to chance“.

We call *stochastic* a system that behaves, mainly, by chance or at random. The cause and effect laws do not explain how the system acts in a deterministic way, but according to probabilities.

In Operational Research: Probabilistic Models and Stochastic Models is the same thing.

Why stochastic processes?

- Complex systems (biological/computer/manufacturing/telecom./…) are
  - Dynamic: they can pass through a **succession** of states as time progresses.
  - Influenced by events which we consider here as **random** phenomena.

- *Definition*: A **stochastic process** is a family of random variables

$$\{X(t) \in \Omega \mid t \in T\}$$

each defined on some (the same for each) sample space Ω for a parameter space T.

Why stochastic processes?

- T, Ω may be either discrete or continuous.
  - Discrete state and continuous state processes:
    - A process is called discrete or continuous state depending upon the values its states can take, i.e., whether the values (Ω) are finite and countable, or any value on the real line.
  - Discrete and continuous (time) parameter processes:
    - A process is called discrete or continuous (time) parameter process depending on whether the index set T is discrete or continuous.

Why stochastic processes?

- *T* is normally regarded as time
  - real time: continuous
  - every month or after job completion: discrete

- Ω is the set of values each $X(t)$ may take
  - bank balance: discrete
  - number of active tasks: discrete
  - time delay in communication network: continuous

Why stochastic processes?

- Example:
  - Suppose we observe $n(t)$, the number of bacterium at a given culture as a function of time, then the process

    $$\{n(t),\ t \in [0,\infty)\}$$

    is a stochastic process, where $n(t)$ is a random variable, and $n(t) \in \{0,1,2,...\}$
  - The values assumed by the random variable are called **states**, and the set of all possible values forms the state space of the process.
  - In this example time is continuous and state space is discrete.

Why stochastic processes?

- Description of a stochastic process:
  - Probabilistic description of a random variable *X* is given by its probability density function (pdf)

$$f_X(x) = \frac{d}{dx} \underbrace{P\{X \le x\}}_{\text{Probability Distribution Function (PDF) also called cumulative distribution function (cdf)}}, \quad -\infty < x < \infty$$

  - Probabilistic description of a stochastic process is given by the **joint pdf** of any set of random variables selected from the process.
    - Thus, in the general case, the detailed description of a stochastic process is unfeasable.

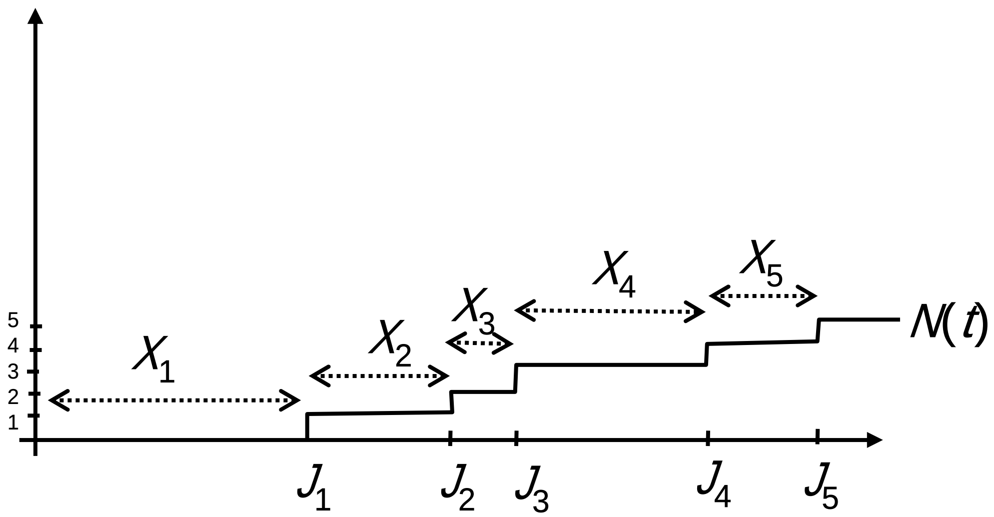
Outline

- Why stochastic processes?
- The Poisson process
  - Exponential distribution
  - Properties of exponential r.v. and Poisson process

The Poisson process

- A large class of stochastic processes are *renewal processes*. This class of processes are used to model independent identically distributed occurrences.
- *Definition*: Let $X_1,X_2,X_3,...$ be independent identically distributed and positive random variables, and set $J_n = X_1+...+X_n$.

  Then process $N(t)$, $t \geq 0$, where $N(t) = \max\{ n \mid J_n \leq t \}$ is called a **renewal process**.

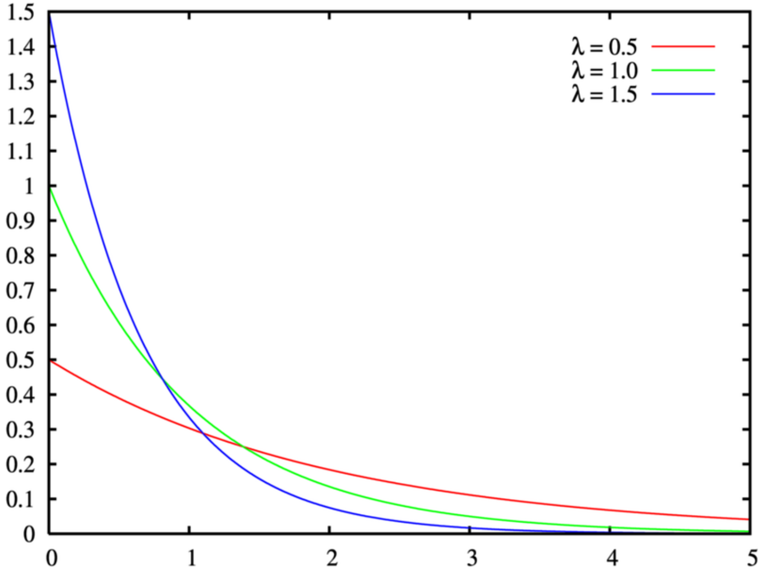


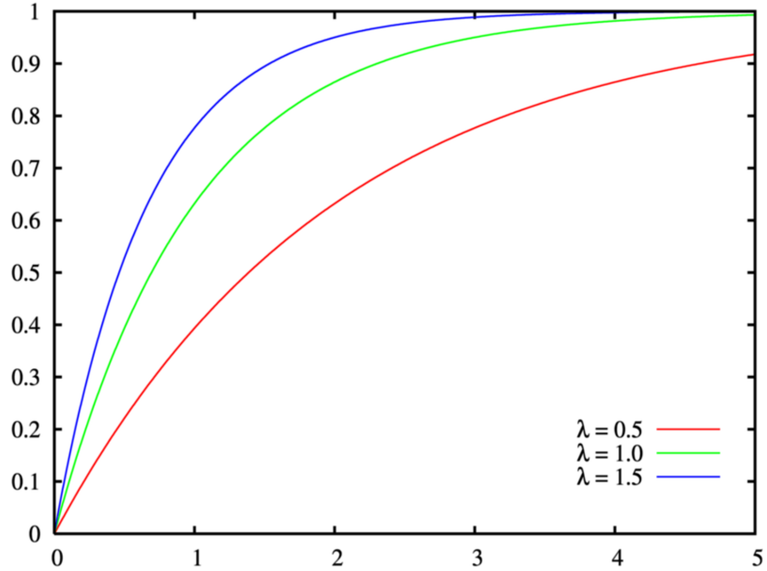
The Poisson process

- *Definition*: The (time-homogeneous, one-dimensional) **Poisson process** is a special case of a renewal process where the time between occurrences is exponentially distributed.

- The pdf and PDF of an **exponentially distributed random variable** *X* are:

$$f_X(x) = \lambda e^{-\lambda x} \quad (x \geq 0) \qquad F_X(x) = P(X \leq x) = 1 - e^{-\lambda x} \quad (x \geq 0)$$





Exponential distribution

- Properties of exponential distribution
  - The mean value and variance

$$E[X] = \frac{1}{\lambda} \qquad V[X] = \frac{1}{\lambda^2}$$

  - The minimum of exponentials is exponential (sum of rates)

$$f_X(x) = \lambda e^{-\lambda x} \quad (x \geq 0)$$

$$f_Y(y) = \mu e^{-\mu y} \quad (y \geq 0)$$

$$Z = \text{mín}\{X, Y\}$$

$$f_Z(z) = (\lambda + \mu) e^{-(\lambda+\mu) z} \quad (z \geq 0)$$

Exponential distribution

Proof:

$$F_X(x) = P(X \le x) = 1 - e^{-\lambda x} \quad (x \ge 0)$$

$$F_Y(y) = P(Y \le y) = 1 - e^{-\mu y} \quad (y \ge 0)$$

$$Z = \text{mín}\{X, Y\}$$

$$F_Z(z) = P(Z \le z) = P(\text{mín}(X, Y) \le z)$$

$$F_Z(z) = P(\text{mín}(X, Y) \le z) = 1 - P(\text{mín}(X, Y) > z) =$$

$$= 1 - P(X > z, Y > z) = 1 - P(X > z)P(Y > z) =$$

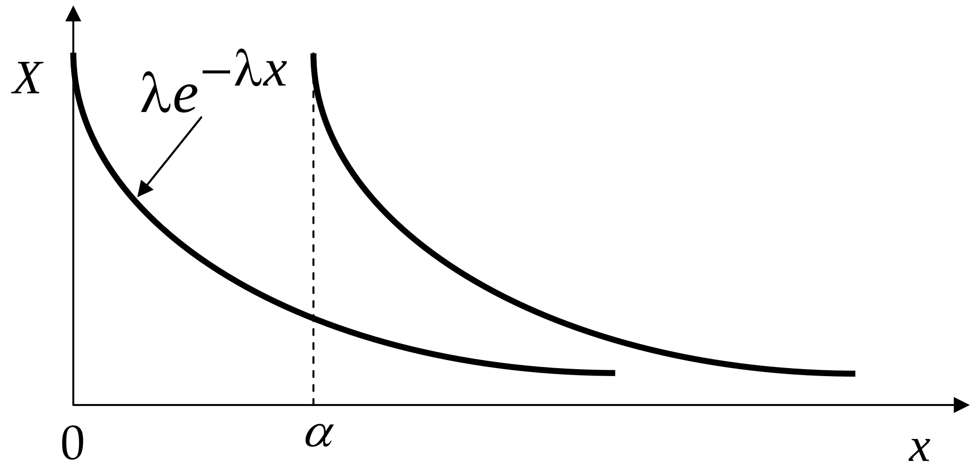
$$= 1 - (1 - P(X \le z))(1 - P(Y \le z)) = 1 - e^{-\lambda z} e^{-\mu z} = 1 - e^{-(\lambda + \mu) z}$$

$$\Rightarrow Z \ \text{ es una variable exponencial de tasa } \lambda + \mu$$

Exponential distribution

– Memoryless property

$$P\{X \geq x+\alpha \mid X \geq \alpha\} = P\{X \geq x\}$$



$$\text{(Right) } P(T > 40 \mid T > 30) = P(T > 10).$$

It does *not* mean

$$\text{(Wrong) } P(T > 40 \mid T > 30) = P(T > 40).$$

The Poisson process

- Properties of Poisson process
  - Residual life
    - If you pick a random time point during a Poisson process, what is the time remaining R to the next instant (arrival)?
    - E.g. when you get to a bus stop, how long will you have to wait for the next bus?
    - If process is Poisson, R has the same distribution as X (the time between ocurrences) by the memoryless property of exponential
    - it doesn't matter when the last bus went!

      contrast constant interarrival times in a perfectly regular bus service

The Poisson process

– Infinitesimal definition of Poisson process

- P(arrival in $(t,t + \Delta t)$) $= P(R \le \Delta t) = \mathrm{P}(X \le \Delta t)$ for all t
  $= 1 - e^{-\lambda \Delta t}$
  $= \lambda\Delta t + o(\Delta t)$
- Therefore
  – Probability of an arrival in $(t, t + \Delta t)$ is $\lambda\Delta t + o(\Delta t)$ regardless of process history before t
  – Probability of more than one arrival in $(t, t + \Delta t)$ is $o(\Delta t)$ regardless of process history before t

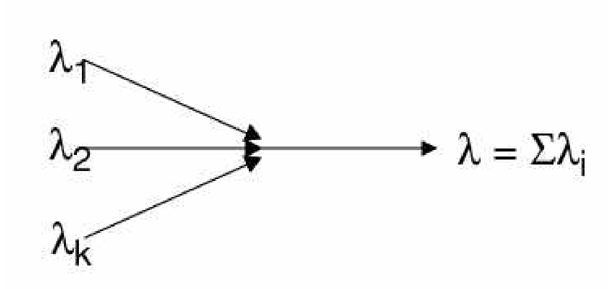
– The Poisson distribution

- Distribution of number of arrivals in time t

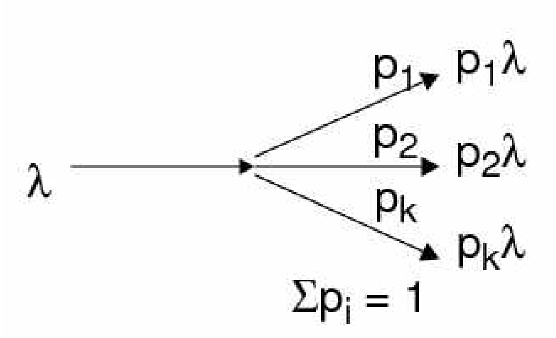
$$P\big(N(t) = n\big) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

The Poisson process

– Superposition property (merging)



– Decomposition property (splitting)



The Poisson process

- Central limit theorem for counting processes:
  - Let $A_1(t)$, …, $A_k(t)$ be independent counting processes (with arbitrary distributions), then

$$X(t) = \frac{\sum_{i=1}^{k} A_i(t)}{k}$$

    is a Poisson process when k → ∞ (under certain "technical conditions")
  - Interpretation: independently of the behaviour of individual countings, the average counting behaviour is Poisson if population is big