Stochastic Processes

Renato Lo Cigno

Simulation and Performance Evaluation 2018-19

Stochastic Process: A Basic Definition

- A stochastic process is an ordered collection or family of random variables

$$\{X(t)|t \in T\}$$

- The values assumed by $X(t)$ are the state space of the process
- t is index of the process, often called *time* and can be continuous or discrete
- Recalling that an RV is a function that maps the events of the RV onto **R** the process is also often written as a function

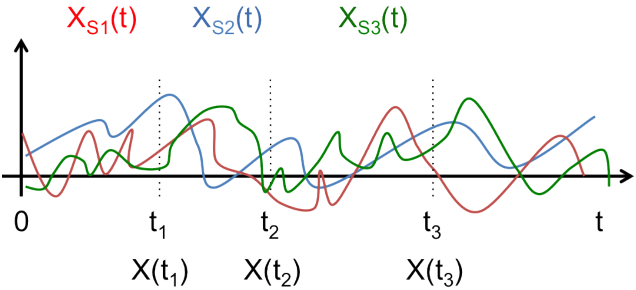
$$\{X(t, s)|s \in S, \; t \in T\}$$

- Given a fixed value of $t = t_1$ then the SP is a standard RV $X(t_1)$ that can assume values in S with a given pdf (or pmf) $f_{X(t_1)}(x)$
- For another point in time $t = t_2$ we have another RV $X(t_2)$ that can assume values in S with a given pdf (or pmf) $f_{X(t_2)}(x)$
  - In general $f_{X(t_1)}(x) \neq f_{X(t_2)}(x)$
- An SP $X(t)$ is thus characterized by the joint pdf (or CDF or pmf) of all possible RVs contained in it as a function of time

$$f_{\mathbf{X}}(\mathbf{x}; \mathbf{t})$$

We call a **realization** of an SP (or sample function) the (deterministic) function of time $X_{s_1}(t)$ that is the outcome of the SP when the sample space has been fixed to s_1

We call a **realization** of an SP (or sample function) the (deterministic) function of time $X_{s_1}(t)$ that is the outcome of the SP when the sample space has been fixed to s_1



Classes of SPs

- The state space S and the index (or time) value t can be continuous or discrete

- We have four classes of SPs
  - CS-CT: Continuous Space – Continuous Time: $s \in \mathbf{R},\ t \in \mathbf{R}$
  - CS-DT: Continuous Space – Discrete Time: $s \in \mathbf{R},\ t \in \mathbf{Z}$
  - DS-CT: Discrete Space – Continuous Time: $s \in \mathbf{Z},\ t \in \mathbf{R}$
  - DS-DT: Discrete Space – Discrete Time: $s \in \mathbf{Z},\ t \in \mathbf{Z}$

- Discrete Space SPs are normally called **chains**

- Temperature of a given area
  - CS-CT. Highly correlated in space and time (the temperature does not change by 10 C^0 in 1s nor in 1m!)

- Temperature of a given area
  - CS-CT. Highly correlated in space and time (the temperature does not change by 10 C^0 in 1s nor in 1m!)

- People (or jobs, or packets, or ...) waiting in a queue
  - DS-CT. People (or jobs, ..., objects in general) are countable so S is the number of objects in the queue; continuous time because objects can arrive and go at any moment
  - But if the "queue" is a bucket of water, then S is continuous (unless you want to count single molecules)

- Temperature of a given area
  - CS-CT. Highly correlated in space and time (the temperature does not change by 10 C^0 in 1s nor in 1m!)

- People (or jobs, or packets, or ...) waiting in a queue
  - DS-CT. People (or jobs, ..., objects in general) are countable so S is the number of objects in the queue; continuous time because objects can arrive and go at any moment
  - But if the "queue" is a bucket of water, then S is continuous (unless you want to count single molecules)

- The number of cars (or PCs, or apples, or wine bottles, ...again objects) produced in a single day
  - DS-DT. Time is discrete, because we are only interested in the number at the end of the day ...in other time instants we can even say that the system simply “does not exist”

- Chemical reactions
  - CS-CT. Highly dependent on temperature and affinity of chemicals. The randomness (normally ignored in high-school chemistry) comes from the random movement of molecules that determines how they "collide"

- Chemical reactions
  - CS-CT. Highly dependent on temperature and affinity of chemicals. The randomness (normally ignored in high-school chemistry) comes from the random movement of molecules that determines how they "collide"
- Disease diffusion
  - DS-CT. People get infected following very complex, time and state dependent random function that depend on the disease and how people interact

- Chemical reactions
  - CS-CT. Highly dependent on temperature and affinity of chemicals. The randomness (normally ignored in high-school chemistry) comes from the random movement of molecules that determines how they “collide”
- Disease diffusion
  - DS-CT. People get infected following very complex, time and state dependent random function that depend on the disease and how people interact
- The value of a market index (is economy a science?)
  - CS-CT. The process is very complex & Mandelbrot (yes, the "fractals" guy) has shown it is strictly unpredictable . . .
  - CS-DT. If you take daily closures

- Execution time of programs/tasks
  - CS-CT. Both time of arrival of programs and tasks are continuous, the execution time is random because of non-deterministic software execution (depends on parameters) and other tasks interaction.

- Execution time of programs/tasks
  - CS-CT. Both time of arrival of programs and tasks are continuous, the execution time is random because of non-deterministic software execution (depends on parameters) and other tasks interaction.
- Noise in electronics or telecommunications (not strictly the same)
  - CS-CT. Can affect voltage, current, EM fields intensity, frequency, phase, . . .

- Execution time of programs/tasks
  - CS-CT. Both time of arrival of programs and tasks are continuous, the execution time is random because of non-deterministic software execution (depends on parameters) and other tasks interaction.
- Noise in electronics or telecommunications (not strictly the same)
  - CS-CT. Can affect voltage, current, EM fields intensity, frequency, phase, ...
- Errors in digital transmissions
  - DS-DT. Errors are on bits or symbols, and bits and symbols are transmitted at discrete intervals

- Packet arrivals at routers/transmission media
  - DS-CT. If I count individual packets, they arrive at any point in time (sometimes)
  - DS-DT. If the system is slotted, so that for a given period of time T the packet is either present or not

UNIVERSITY OF TRENTO
Department of Information Engineering and Computer Science

- Packet arrivals at routers/transmission media
  - DS-CT. If I count individual packets, they arrive at any point in time (sometimes)
  - DS-DT. If the system is slotted, so that for a given period of time T the packet is either present or not

- The topology of a P2P overlay
  - DS-CT. The space is the graph representing the topology …the space has an enormous dimension, but it is countable, time is continuous unless we take snapshots

- Packet arrivals at routers/transmission media
  - DS-CT. If I count individual packets, they arrive at any point in time (sometimes)
  - DS-DT. If the system is slotted, so that for a given period of time T the packet is either present or not

- The topology of a P2P overlay
  - DS-CT. The space is the graph representing the topology ...the space has an enormous dimension, but it is countable, time is continuous unless we take snapshots

- Request arrivals at servers
  - DS-CT. If I count individual requests, they arrive at any point in time
  - DS-DT. If I count the number of requests per second (minute/hour/...)

A generic SP $\{X(t)|t \in T\}$ is an ordered collection of RVs

$$
\begin{aligned}
&\{X(t)|t \in T\} = \\
&\{X(t_n), X(t_{n-1}), \cdots, X(t_2), X(t_1)|t_n > t_{n-1} > \cdots > t_2 > t_1\} \\
&\forall \mathbf{x} = (x_n, x_{n-1}, \cdots, tx_1) \in \mathbf{R}^n; \\
&\forall \mathbf{t} = (t_n, t_{n-1}, \cdots, t_1) \in T^n;\ \ t_n < t_{n-1} < \cdots < t_1
\end{aligned}
$$

A generic SP $\{X(t)|t \in T\}$ is an ordered collection of RVs

$$
\begin{aligned}
&\{X(t)|t \in T\} = \\
&\{X(t_n), X(t_{n-1}), \cdots, X(t_2), X(t_1) | t_n > t_{n-1} > \cdots > t_2 > t_1\} \\
&\forall \mathbf{x} = (x_n, x_{n-1}, \cdots, tx_1) \in \mathbf{R}^n; \\
&\forall \mathbf{t} = (t_n, t_{n-1}, \cdots, t_1) \in T^n;\ t_n < t_{n-1} < \cdots < t_1
\end{aligned}
$$

Thus in general the SP $\{X(t)|t \in T\}$ can be described completely only through the global n-th order statistic:

$$F_{\mathbf{X}}(\mathbf{x}; \mathbf{t}) = \mathbf{P}[X_n(t_n) \leq x_n, X_{n-1}(t_{n-1}) \leq x_{n-1}, \cdots, X_1(t_1) \leq x_1]$$

A generic SP $\{X(t)|t \in T\}$ is an ordered collection of RVs

$$\{X(t)|t \in T\} =$$
$$\{X(t_n), X(t_{n-1}), \cdots, X(t_2), X(t_1)|t_n > t_{n-1} > \cdots > t_2 > t_1\}$$
$$\forall \mathbf{x} = (x_n, x_{n-1}, \cdots, tx_1) \in \mathbf{R}^n;$$
$$\forall \mathbf{t} = (t_n, t_{n-1}, \cdots, t_1) \in T^n;\ \ t_n < t_{n-1} < \cdots < t_1$$

Thus in general the SP $\{X(t)|t \in T\}$ can be described completely only through the global n-th order statistic:

$$F_{\mathbf{X}}(\mathbf{x}; \mathbf{t}) = \mathbf{P}[X_n(t_n) \leq x_n, X_{n-1}(t_{n-1}) \leq x_{n-1}, \cdots, X_1(t_1) \leq x_1]$$

- Such a complete description is a forbidding task ... but sometimes we can simplify this description

A generic SP $\{X(t)|t \in T\}$ is said strictly stationary if

$$F_{\mathbf{X}}(\mathbf{x};\mathbf{t}) = F_{\mathbf{X}}(\mathbf{x};\mathbf{t}+\tau) \tag{1}$$

for all vectors $\mathbf{x} \in \mathbf{R}^n$ and $\mathbf{t} \in T^n$, and all scalars τ (added to all components of $\mathbf{t}$) for which $t_i + \tau \in T$.

A generic SP $\{X(t)|t \in T\}$ is said strictly stationary if

$$F_{\mathbf{X}}(\mathbf{x};\mathbf{t}) = F_{\mathbf{X}}(\mathbf{x};\mathbf{t}+\tau) \tag{1}$$

for all vectors $\mathbf{x} \in \mathbf{R}^n$ and $\mathbf{t} \in T^n$, and all scalars τ (added to all components of $\mathbf{t}$) for which $t_i + \tau \in T$.

- Let $\mu(t) = E[X(t)]$ be the time dependent average (or ensemble average) of the a stationary process $\{X(t)|t \in T\}$. It is not difficult to see that $\mu(t) = \mu$: it is constant for all times

A generic SP $\{X(t)|t \in T\}$ is said strictly stationary if

$$F_{\mathbf{X}}(\mathbf{x}; \mathbf{t}) = F_{\mathbf{X}}(\mathbf{x}; \mathbf{t} + \tau) \tag{1}$$

for all vectors $\mathbf{x} \in \mathbf{R}^n$ and $\mathbf{t} \in T^n$, and all scalars τ (added to all components of $\mathbf{t}$) for which $t_i + \tau \in T$.

- Let $\mu(t) = E[X(t)]$ be the time dependent average (or ensemble average) of the a stationary process $\{X(t)|t \in T\}$. It is not difficult to see that $\mu(t) = \mu$: it is constant for all times
- With some more computations it is not difficult to prove that strict stationarity implies that all central moments and in particular the variance are constant in time

A generic SP $\{X(t)|t \in T\}$ is said strictly stationary if

$$F_{\mathbf{X}}(\mathbf{x};\mathbf{t}) = F_{\mathbf{X}}(\mathbf{x};\mathbf{t}+\tau) \qquad (1)$$

for all vectors $\mathbf{x} \in \mathbf{R}^n$ and $\mathbf{t} \in T^n$, and all scalars τ (added to all components of $\mathbf{t}$) for which $t_i + \tau \in T$.

- Let $\mu(t) = E[X(t)]$ be the time dependent average (or ensemble average) of the a stationary process $\{X(t)|t \in T\}$. It is not difficult to see that $\mu(t) = \mu$: it is constant for all times
- With some more computations it is not difficult to prove that strict stationarity implies that all central moments and in particular the variance are constant in time
- Still to fully describe a stationary process we need the full n-th order joint distribution

A process is independent if all its n-th order joint pdf, pmf, CDF are in product form

$$
\begin{aligned}
F_{\mathbf{X}}(\mathbf{x};\mathbf{t}) &= \prod_{i=1}^{n} F_{X_i}(x_i;t_i) \\
&= \prod_{i=1}^{n} \mathbf{P}[X_i(t_i) \leq x_i] \\
f_{\mathbf{X}}(\mathbf{x};\mathbf{t}) &= \prod_{i=1}^{n} f_{X_i}(x_i;t_i)
\end{aligned}
$$

- This is the simplest form of SP
- We often find it in measurement and experiments

A special case of Independent Process is the case where the time is discrete and all X_i are non-negative, independent and identically distributed (iid) RV

$$\{X_n | n \in \mathbf{Z}\} = \{X_n, X_{n-1}, \cdots, X_1\}$$

- Systems that can be repaired (or replaced) in negligible time: the sequence of times between failures is a renewal process (the name comes from this example)

- The assumption of independence, as well as the one of strict stationarity, are rarely met in reality, because they assume the world has no memory at all of the past
- A mild, but very useful and often met in reality, form of dependence is assuming that the evolution of the process depends only on its present state
- In other words that RV $X(t_{n+1})$ depends only on RV $X(t_n)$

- The assumption of independence, as well as the one of strict stationarity, are rarely met in reality, because they assume the world has no memory at all of the past
- A mild, but very useful and often met in reality, form of dependence is assuming that the evolution of the process depends only on its present state
- In other words that RV $X(t_{n+1})$ depends only on RV $X(t_n)$

A process $\{X(t)|t \in T\}$ is called **Markov Process** if given $t_0 < t_1 < \cdots < t_n$

$$\mathbf{P}[X(t) \leq x|X(t_n) = x_n, X(t_{n-1}) = x_{n-1}, X(t_0) = x_0] = \mathbf{P}[X(t) \leq x|X(t_n) = x_n] \tag{2}$$

- An MP is not strictly stationary: **Eq. (2) does not imply Eq. (1)**
- The distribution or RVs of an MP depends on its state, thus an MP is actually not stationary
- The so-called “Markov property” formalized by Eq. (2) implies only that the evolution of a markovian stochastic system depends only on its **current state at time** t, and not on its past history
- The history is completely summarized by the state at time t

- An MP is not strictly stationary: **Eq. (2) does not imply Eq. (1)**
- The distribution or RVs of an MP depends on its state, thus an MP is actually not stationary
- The so-called "Markov property" formalized by Eq. (2) implies only that the evolution of a markovian stochastic system depends only on its **current state at time** t, and not on its past history
- The history is completely summarized by the state at time t
- **Notice the dependence on time** t

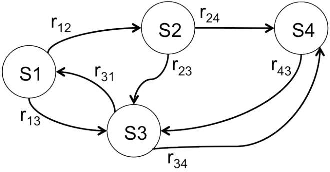
- In many real cases the marginal distributions of an MP are time invariant, i.e.:

$$\mathbf{P}[X(t) \leq x | X(t_n) = x_n] = \mathbf{P}[X(t - t_n) \leq x | X(t_0) = x_n] \quad (3)$$

  we cal this MP **time-homogeneous**, and we will normally restrict our analysis to this class of MPs
- Again, Eq. (3) does not imply Eq. (1), thus also a time-homogeneous MP is not strictly stationary
- However in a time-homogeneous MP the evolution of the process depends on the state and not on t, thus we can say that the **state of the process completely describes its past**
- This is indeed a very "normal" behavior that we expect from the real world

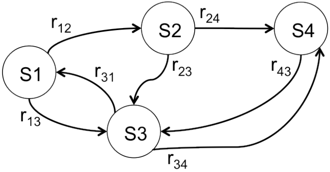
- If we restrict the state to be discrete we can represent graphically an MP

- If we restrict the state to be discrete we can represent graphically an MP

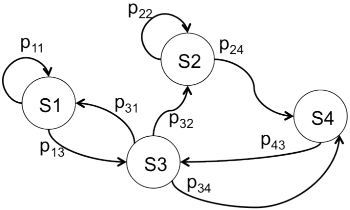


- In continuous time: CTMC

- If we restrict the state to be discrete we can represent graphically an MP

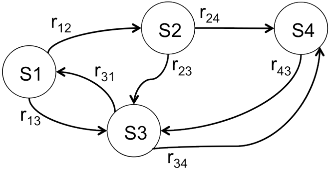


- In continuous time: CTMC

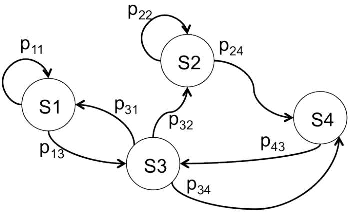


- In discrete time: DTMC

- If we restrict the state to be discrete we can represent graphically an MP



- In continuous time: CTMC



- In discrete time: DTMC
- The time spent by the system in a state is called *dwell time*

A time-homogeneous CTMC has exponential dwell times, let's see why.

A time-homogeneous CTMC has exponential dwell times, let's see why.

- Given that the past history of the process must be completely described by the current state, the dwell time Y in any state must satisfy the following property:

$$\mathbf{P}[Y \leq r + t | Y \geq t] = \mathbf{P}[Y \leq r]$$

A time-homogeneous CTMC has exponential dwell times, let's see why.

- Given that the past history of the process must be completely described by the current state, the dwell time Y in any state must satisfy the following property:

$$\mathbf{P}[Y \leq r + t | Y \geq t] = \mathbf{P}[Y \leq r]$$

- and by definition of conditional probability

$$\mathbf{P}[Y \leq r] = \frac{\mathbf{P}[t \leq Y \leq r + t]}{\mathbf{P}[Y \geq t]}$$

- Which is equivalent to

$$F_Y(r) = \frac{F_Y(t+r) - F_Y(t)}{1 - F_Y(t)}$$

- Which is equivalent to

$$F_Y(r) = \frac{F_Y(t+r) - F_Y(t)}{1 - F_Y(t)}$$

- Taking the derivative of both terms with respect to r we have

$$F_Y'(0) = \frac{F_Y'(t)}{1 - F_Y(t)}$$

- Which is equivalent to

$$F_Y(r) = \frac{F_Y(t+r) - F_Y(t)}{1 - F_Y(t)}$$

- Taking the derivative of both terms with respect to r we have

$$F'_Y(0) = \frac{F'_Y(t)}{1 - F_Y(t)}$$

- This is a differential equation that admits a unique solution

$$F_Y(t) = 1 - e^{F'_Y(0)t}$$

i.e., the dwell time must be exponentially distributed

A time-homogeneous DTMC has geometically distributed dwell times

- With a procedure similar to the one used for CTMC but using finite difference equations it is not difficult to show that the dwell time for this class of DTMCs must be geometrically distributed

A time-homogeneous DTMC has geometically distributed dwell times

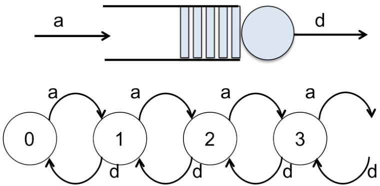
- With a procedure similar to the one used for CTMC but using finite difference equations it is not difficult to show that the dwell time for this class of DTMCs must be geometrically distributed

- The "time" between successive steps of a DTMC (DT process in general) "does not exist", i.e., it is not correct to try to represent a DT process in continuous time, as the process cannot represent instants between transition events

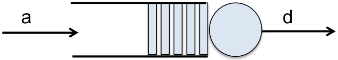
- Generalizing a time-homogeneous MC to dwell times that follow an arbitrary distribution we obtain a MC whose future evolution depends on the state, but also on the time t_d already spent in the state
- Clearly the MC is not time-homogeneous, but we have a guarantee that whenever the process returns to the same state, its behavior remains the same as in the previous visit

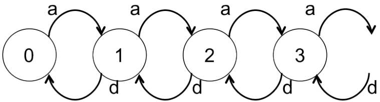
- Generalizing a time-homogeneous MC to dwell times that follow an arbitrary distribution we obtain a MC whose future evolution depends on the state, but also on the time t_d already spent in the state
- Clearly the MC is not time-homogeneous, but we have a guarantee that whenever the process returns to the same state, its behavior remains the same as in the previous visit
- An Event Driven simulation is always a semi-Markov process in discrete time where the program variables are the state and the events correspond to transitions between states

Example: A queueing system with exponential inter-arrival times with rate a, but generally distributed service times with rate d



Example: A queueing system with exponential inter-arrival times with rate *a*, but generally distributed service times with rate *d*





- A CPU serving tasks in batch (there is no reason why a task should have an exponential service time)
- Links serving packets of non-geometric size
- A post office/bank/train tiketing

Autocorrelation Function

- How can we measure if the RVs X_{t_i} of an SP $\{X\}$ are correlated or not?
- knowing that a process is independent is important
  ...specially if we are taking measures from a process and we want to reconstruct its structure
- Given $\{X\}$, take the function

$$R(t_1, t_2) = E[X(t_1) \cdot X(t_2)]$$

- We call this function the **Autocorrelation function** of the process $\{X\}$

- The autocorrelation computed for $t_2 = t_1$ is the second moment of $X(t_1)$

$$R(t_1, t_2) = E[X(t_1) \cdot X(t_1)] = E[X^2(t_1)] = \mathrm{Var}[X(t_1)] + \mu_{X(t_1)}^2$$

- The autocorrelation computed for $t_2 = t_1$ is the second moment of $X(t_1)$

$$R(t_1, t_2) = E[X(t_1) \cdot X(t_1)] = E[X^2(t_1)] = \mathrm{Var}[X(t_1)] + \mu^2_{X(t_1)}$$

- The autocorrelation is a function of the covariance

$$R(t_1, t_2) = \mathrm{Cov}(X(t_1), X(t_2)) + \mu_1 \mu_2$$

- The autocovariance of a stationary process $\{X\}$ depends only on $\tau = t_2 - t_1$

$$R(t_1, t_2) = R(\tau);\ \tau = t_2 - t_1$$

thus it is a one-dimensional function of τ

- The autocovariance of a stationary process $\{X\}$ depends only on $\tau = t_2 - t_1$

$$R(t_1, t_2) = R(\tau);\ \tau = t_2 - t_1$$

thus it is a one-dimensional function of τ

- The autocorrelation function of a stationary independent process is

$$R(\tau) = \begin{cases} 0, & \forall \tau \neq 0 \\ \mathrm{Var}[X(t)] + \mu^2 = \sigma^2 + \mu^2, & \tau = 0 \end{cases}$$

- We can normalize $R(\cdot)$ so that it is a function comprised between -1 and 1

$$R'(t_1, t_2) = \frac{R(t_1, t_2) - \mu_{X(t_1)}\mu_{X(t_2)}}{\sigma_{X(t_1)}\sigma_{X(t_2)}}$$

- We can normalize $R(\cdot)$ so that it is a function comprised between -1 and 1

$$R'(t_1, t_2) = \frac{R(t_1, t_2) - \mu_{X(t_1)}\mu_{X(t_2)}}{\sigma_{X(t_1)}\sigma_{X(t_2)}}$$

- Restricting to stationary processes

$$R'(\tau) = \frac{R(\tau) - \mu^2}{\sigma^2}$$

Consider a process $\{X\}$ with the following properties

1. $\mu(t) = E[X(t)] = \mu$ is time independent
2. $R(t_1, t_2) = R(0, t_2 - t_1) = R(\tau);\ \ t_2 \geq t_1 \geq 0$
3. $R(0) = E[X^2(t)] = \sigma^2 \leq \infty$ (the variance is finite and constant)

Consider a process $\{X\}$ with the following properties

1. $\mu(t) = E[X(t)] = \mu$ is time independent
2. $R(t_1, t_2) = R(0, t_2 - t_1) = R(\tau);\ \ t_2 \geq t_1 \geq 0$
3. $R(0) = E[X^2(t)] = \sigma^2 \leq \infty$ (the variance is finite and constant)

We call this process stationary in wide-sense and it is indeed a definition that comply with our intuition and it is easily measurable on data