Network Engineering

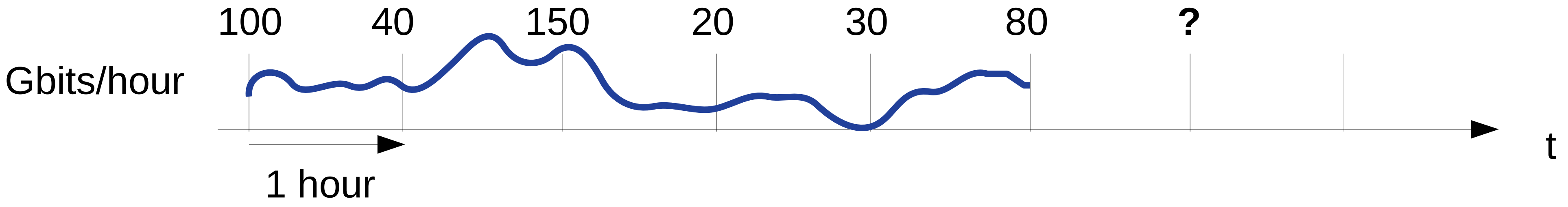
Lecture 2. Stochastic Processes

Boris Bellalta
boris.bellalta@upf.edu

Wireless Networking
https://www.upf.edu/web/wnrg

What is a random phenomenon?

- We cannot know in advance the ‘value’ of future outcomes.
  - Example: the number of bits transmitted in the next hour over a communication link.
- Random/stochastic process: the entity that generates the sequence (temporal) of random outcomes.



We can characterize it

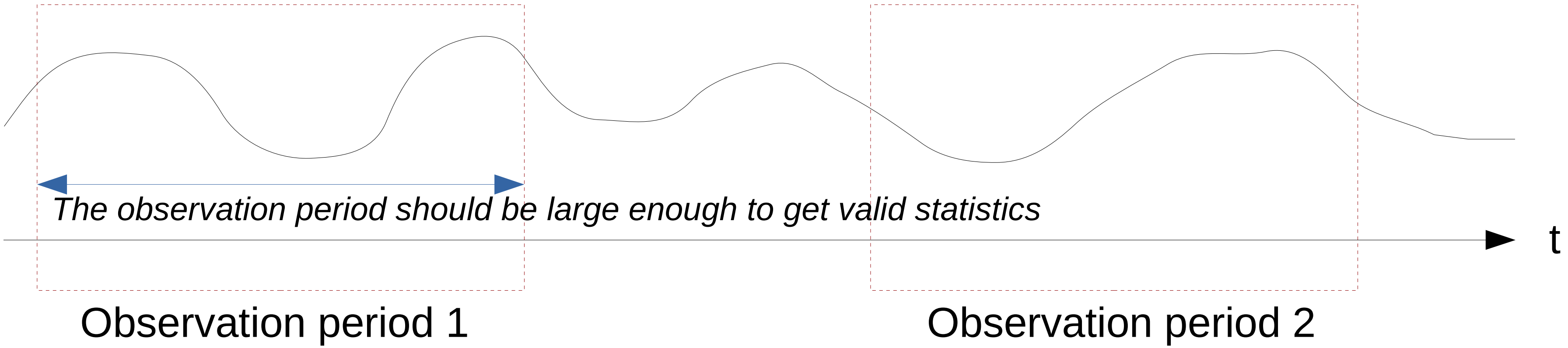
- We can study the stochastic process (i.e., the sequence of outcomes) and characterize it
  - Range of values: state space
  - Mean, variance, median, etc.
  - Histogram / stationary probability distribution

Example

- We monitor the number of associated STAs in an AP per minute, and get the following temporal sequence:
  - x(t) = [… 10 9 8 10 11 9 7 10 5 8 ?…]
- We don’t know what will be the next value.
- However, we can characterize the stochastic process by computing some key parameters (mean, variance, probability distribution, etc.), and so have some ‘knowledge’ of what could be the next value.

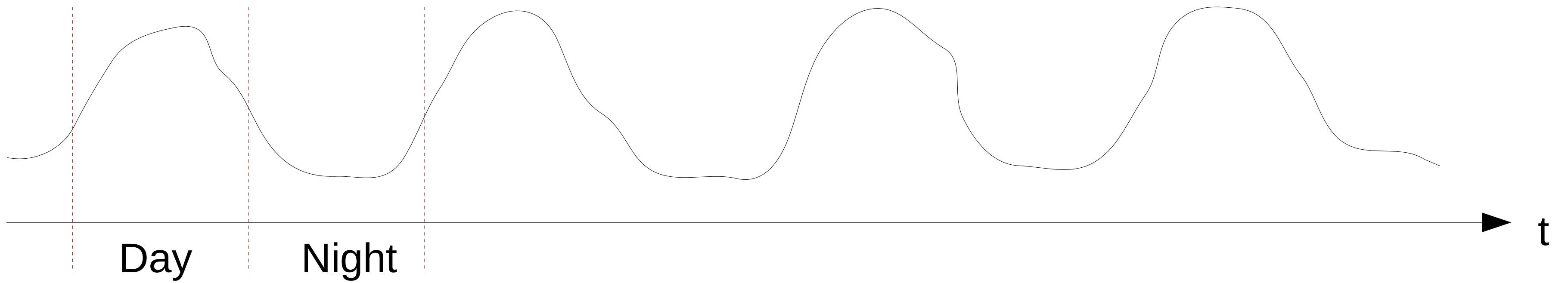
Stationary process

- The statistics of the stochastic process do not change with the time.
- Example: We will get the same ‘statistics’ from different observation periods if the stochastic process is stationary.

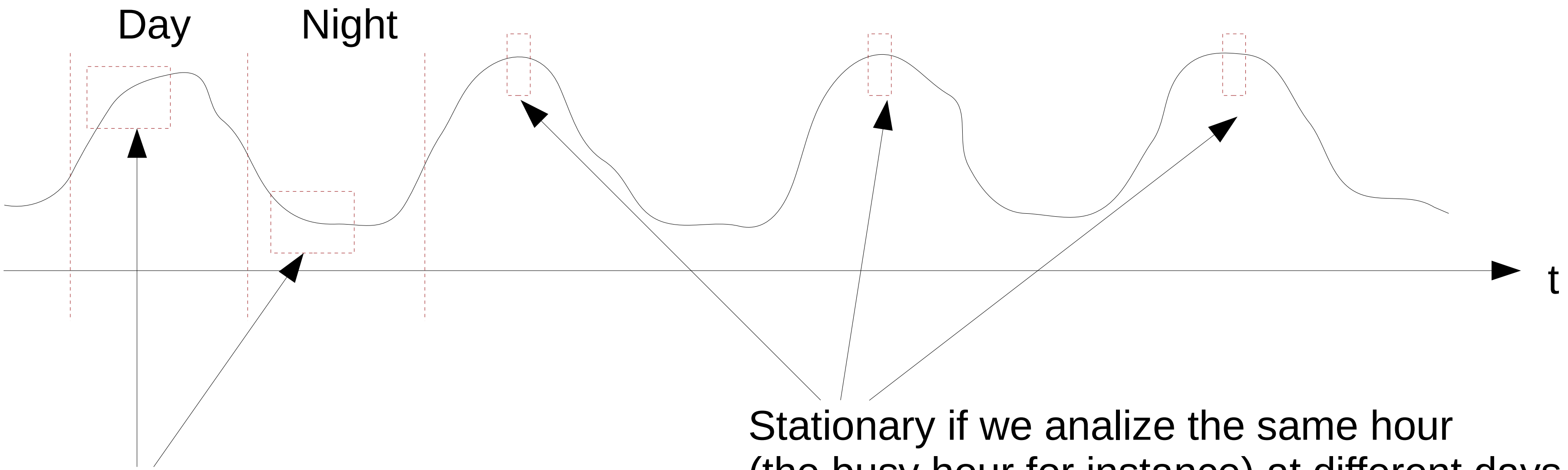


Stationary process

- Non-stationary processes are challenging as previous knowledge may not be useful to describe future outcomes.
- We will assume all stochastic processes are stationary (at least during the time of interest).
- Example: traffic load of a link during some days



Stationary process



Not stationary
(different statistics at different observation periods)

Stationary if we analize the same hour
(the busy hour for instance) at different days.

We do a subsampling of the original stochastic process to create a new one
(including only the busy hours of each day)

Stationary dynamics: assumption!

- In reality (in networking), it is hard to find stationary phenomena:
  - Traffic loads keep increasing all the time.
  - User activity depends on new apps / services, external events.
- However, we must assume it as the mathematical tools we will use later require it.
  - Not too bad, but we must be aware the results obtained using those tools are just an approximation (very good in most cases).

Characterizing a random variable

Let X be a random variable that takes values from the set $\mathcal{X}$, which is called the state space of X. Each possible value $x \in \mathcal{X}$ has assigned a probability, and will be referred as $\mathrm{P}\{X = x\}$. For an stationary stochastic process that is stationary, the $\mathrm{P}\{X = x\}$ does not change with the time.

A first consideration about X is related with $\mathcal{X}$:

- If the range of values that X can take is finite, we say that X is a random variable with a **discrete state space**.
- If the range of values that X can take is infinite, we say that X is a random variable with a **continuous state space**.

Characterizing a random variable

A second consideration is when the random variable X takes a new value.

- If X can take a new value at any arbitrary time, we say that X is a **continuous time** random variable.
- If X can take a new value only at specific instants of time, we say that X is a **discrete time** random variable.

Finally, a third consideration is about the dependence between present, past and future values.

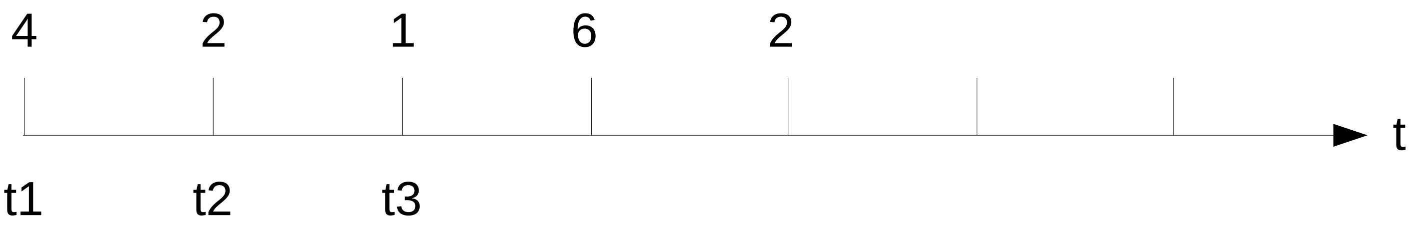
- We will say that the stochastic process that generates X is an **independent stochastic process**.
- If there are some dependencies between the present and past or future values, we will refer to it as a **dependent stochastic process**.

Example 1

- Random variable: X
- State space: χ=[{1}{2}{3}{4}{5}{6}]
  - Discrete
- Value of X: x (it must belong to the state space)
- P(X=x)=1/6, for all x
  - P(X=1) = 1/6; P(X=2) = 1/6; … ; P(X=6) = 1/6

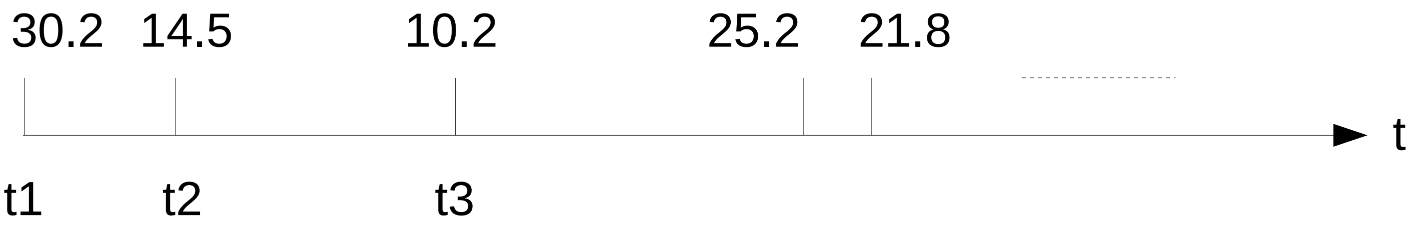
Since we assume X(t) is stationary, the prob. of occurrence of each possible value is the same regardless 't'.

- Discrete in time (it gets new values only at specified time instants)



Example 2

- Random variable: A
- State space: α=[10,40]
  - Continuous (any value between 10 and 40, included)
- Value of A: a (it must belong to the state space)
- P(A=a)=1/30 (uniform distribution, 1/(b-a), with b=40 and a = 10)
- Continuous in time (the value may change at any given instant of time)



2.2.1 Histogram

The histogram is a function that given a value of X, x (discrete), or a range of values of X, $[x_1, x_2]$ (continuous), gives the probability that the x value, or a value inside the chosen range of values, appears. In detail:

- If X has a discrete state space, the histogram of X is the $P\{X = x\}, \forall x \in \mathcal{X}$.
- If X has an infinite state space, the histogram of X is the $P\{x_1 < X \leq x_2\}$, $\forall x_1,\ x_2 \in \mathcal{X}$.

If the process is stationary, we will refer to the histogram as the '**stationary probability distribution**'

2.2.2 Expected Value

If the random variable is discrete, we can compute its expected value, $E[X]$, as follows:

$$E[X] = \sum_{\forall x \in \mathcal{X}} x P\{X = x\} \tag{2.4}$$

In case the random variable is continuous, we have to use its pdf:

$$\mathrm{E}[X] = \int_{-\infty}^{+\infty} x f_X(x) dx \tag{2.5}$$

2.2.3 Variance

The variance is an indicator of the dispersion of the values that X can take. If the random variable X is discrete, the variance can be calculated as follows

$$\mathrm{V}[X] = \sum_{\forall x \in \mathcal{X}} P\{X = x\} \left(x - \mathrm{E}[X]\right)^2 \tag{2.6}$$

In case the random variable X is continuous, the variance is:

$$\mathrm{V}[X] = \int_{-\infty}^{+\infty} f_X(x) \left(x - \mathrm{E}[X]\right)^2 dx \tag{2.7}$$

2.2.4 Coefficient of Variation

The coefficient of variation is a metric that depends on the both the variance and the expected value. It is computed as:

$$\mathrm{CV}[X] = \frac{\sqrt{\mathrm{V}[X]}}{\mathrm{E}[X]} \tag{2.8}$$

Note that $\sqrt{\mathrm{V}[X]}$ is known as the standard deviation of X.

2.2.5 Moments

The mth moment of a random variable X is defined as:

$$\mathrm{E}[X^m] = \sum_{\forall x \in \mathcal{X}} x^m P\{X = x\} \tag{2.9}$$

in case it is discrete. If it is continuous, it is defined as:

$$\mathrm{E}[X^m] = \int_{-\infty}^{+\infty} x^m f_X(x) dx \tag{2.10}$$

Note that the first moment is the expected value and that the variance can be easily obtained from the first and second moments:

$$\mathrm{V}[X] = \mathrm{E}[X^2] - \mathrm{E}^2[X] \tag{2.11}$$

Example 1: Number of packets received by Router R1 in intervals of Δ seconds

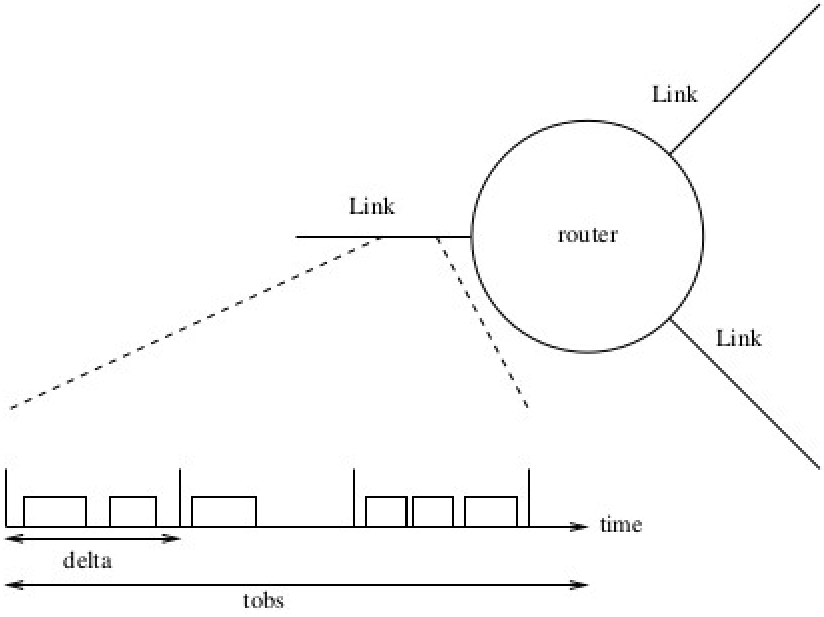


Figure 2.1: Measuring the transmitted packets between R2 and R1.

Consider the router R1 in Figure 4.1, depicted also in Figure 2.1. We measure the number of packets that arrive to R1 from R2, and count the number of packets that arrive to R1 in intervals of Δ seconds. The observation time is $T_{\text{obs}} = 10\Delta$, and the collected data is shown in Table 2.1.

Interval (Δ)	Packets (X)	Interval	Packets (X)
1	5	6	2
2	4	7	1
3	4	8	5
4	3	9	3
5	1	10	1

Table 2.1: Results from the experiment

As it can be observed, the number of packets that arrive at each interval Δ is a random variable, and we call it X. Now, we are going to characterize X by computing its *histogram*, *expected value* and *variance*.

Solution:

To obtain the histogram we need to compute the probability $P\{X = x\}$, for all possible values of x, i.e., $\forall x \in \mathcal{X}$.

$$P\{X = x\} = \frac{\text{Number of samples equal to } x}{N} \tag{2.13}$$

where N is the total number of samples. In our case, $N = 10$. We can observe that X only takes five different values: $\{1,\ 2,\ 3,\ 4,\ 5\}$. The resulting histogram is depicted in Table 2.2.

x	Appearances	Probability
1	3	3/10
2	1	1/10
3	2	2/10
4	2	2/10
5	2	2/10

Table 2.2: Histogram

To compute the expected value of X, we use (2.4).

$$E[X] = \sum_{\forall x \in \mathcal{X}} x\mathrm{P}\{X = x\} = \tag{2.14}$$

$$= \frac{3}{10} + 2\frac{1}{10} + 3\frac{2}{10} + 4\frac{2}{10} + 5\frac{2}{10} = \frac{29}{10} = 2.9 \text{ packets} \tag{2.15}$$

Now, we compute the variance using (2.6).

$$\mathrm{Var}[X] = \sum_{\forall x \in \mathcal{X}} \mathrm{P}\{X = x\}(x - E[X])^2 = \tag{2.16}$$

$$= (1 - 2.9)^2 \frac{3}{10} + (2 - 2.9)^2 \frac{1}{10} + (3 - 2.9)^2 \frac{2}{10} + \tag{2.17}$$

$$+ (3 - 2.9)^2 \frac{2}{10} + (3 - 2.9)^2 \frac{2}{10} = 2.29 \text{ packets}^2 \tag{2.18}$$

From the variance and the expected value, we compute the coefficient of variation:

$$\mathrm{CV}[X] = \frac{\sqrt{\mathrm{Var}[X]}}{E[X]} = 0.5218 \tag{2.19}$$

Observe that the variance can be computed using the first and second moment,

$$\mathrm{Var}[X] = E[X^2] - E^2[X] \tag{2.20}$$

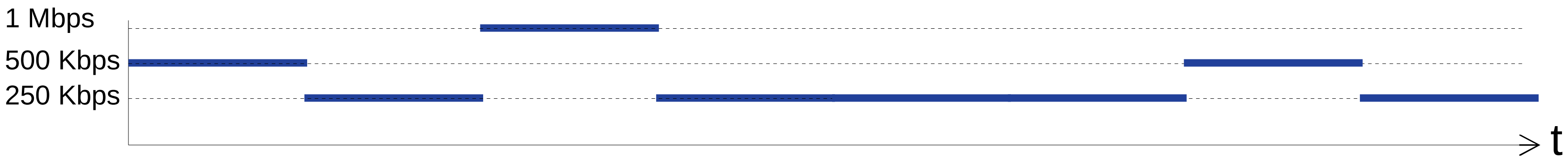
where the second moment is

$$\mathrm{E}[X^2] = \sum_{\forall x \in \mathcal{X}} x\mathrm{P}\{X = x\} = \tag{2.21}$$

$$= \frac{3}{10} + 2^2\frac{1}{10} + 3^2\frac{2}{10} + 4^2\frac{2}{10} + 5^2\frac{2}{10} = 10.7 \text{ packets}^2 \tag{2.22}$$

Exercise

- A video flow adapts its rate every second depending on the contents:
  - Fast changes on the images → high rate; otherwise the opposite.
- Calculate the stationary probability distribution, first and second moments of the following data:



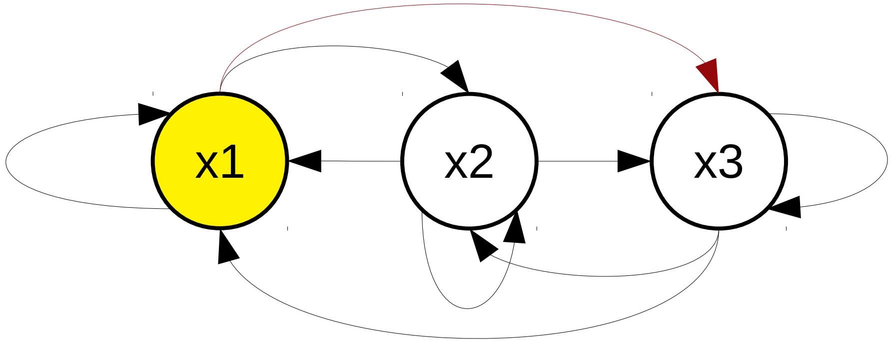
- <u>From the observed data</u>: what is the state space? Is it discrete or continuous? Is it continuous or discrete in time?

Independent stochastic process: representation

- **Definition**: Future values of the random variable do not depend on current or past values.
- Let's consider a stochastic process X(t) with state space χ=[{x1}{x2}{x3}], that changes at discrete instants of time.
- Let's represent the values of the state space by circles, and by arrows the transition probabilities:

Prob. that next value of X(t), X(t+1), is x3 given current value ix x1

… x1,x2,x1,x3,x3,x2,**x1**,?,...



Important: we are assuming next value only depends (if it depends) on current value

Independent stochastic process: representation

- Let's represent transition probabilities as p_{ij} (prob to move from state i to state j)
- In independent stochastic processes transition probabilities correspond to the stationary probability distribution of the process:
  - $p_{x1x1}=P(X=x1)$; $p_{x1x2}=P(X=x2)$; $p_{x1x3}=P(X=x3)$ → They sum 1!
  - $p_{x2x1}=P(X=x1)$; $p_{x2x2}=P(X=x2)$; $p_{x2x3}=P(X=x3)$
  - $p_{x3x1}=P(X=x1)$; $p_{x3x2}=P(X=x2)$; $p_{x3x3}=P(X=x3)$
- Therefore, future values DOES NOT depend on current value → Of course, it is an independent stochastic process.

Dependent stochastic process: representation

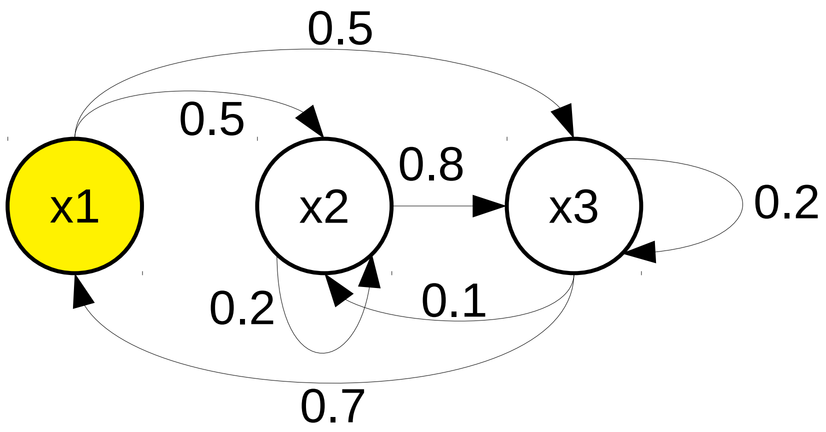
- **Definition**: Future values of the random variable depend on current or past values.
- In dependent stochastic processes transition probabilities does not correspond to the stationary probability distribution of the process.
- For example, considering X(t), and the following transition probabilities:
  - $p_{x1x1}=0$; $p_{x1x2}=0.5$; $p_{x1x3}=0.5$
  - $p_{x2x1}=0$; $p_{x2x2}=0.2$; $p_{x2x3}=0.8$
  - $p_{x3x1}=0.7$ $p_{x3x2}=0.1$; $p_{x3x3}=0.2$

→ Probability transition matrix, P

- Future values depend on current state! The prob to observe x3 (0.5) is different if we are now in state x1 than if we are in state x2 (0.8)

Dependent stochastic process: representation

- We can represent both states and transition probabilities:
  - Probabilities equal to 0 are generally omitted



In dependent stochastic processes, we usually will get the process's state space, and the probability transition matrix, as they are related with some phenomena. Then, we will have to find the stationary probability distribution.

Dependent state processes: past dependence

- We can have stochastic processes in which the next value depends on current value and other past values.
- They are however hard to be represented graphically, as the concept of state must include the path to reach it.
- Similarly, transitions have to include current and previous state, i.e.,
  - $P_{i,\{j,q\}}=P(X(t+1)=i \mid X(t)=j,X(t-1)=q)$
- In this subject we will consider only dependent stochastic processes in which next value depends on the current one.

Dependence: general formulation

- In general: $P_{xt+1, \{xt,xt-1,xt-2,\ldots\}} = P(X(t+1)=x_{t+1} \mid X(t)=x_t, X(t-1)=x_{t-1}, X(t-2)=x_{t-2}, \ldots)$

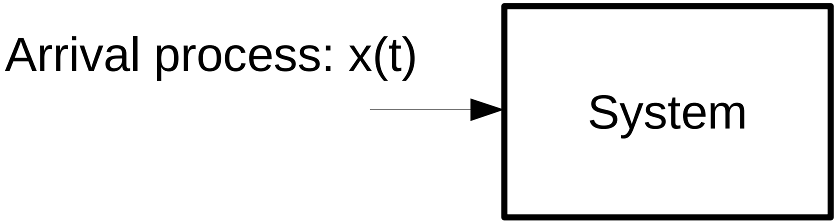
- Independent processes: $P_{xt+1} = P(X(t+1)=x_{t+1})$
- Dependent processes (on current value): $P_{xt+1,xt} = P(X(t+1)=x_{t+1} \mid X(t)=x_t)$
  - Markovian processes

Example of independent and dependent processes

- Independent process: tossing a coin, rolling a dice…
- Dependent process:
  - $X(t)$ = Number of people waiting in a queue (at the supermarket)
  - State space: $X = [0, 1, 2, \dots Q]$
    - Discrete
  - A new value of $X(t)$ appears when a customer arrives to the queue or departs.
    - Continuous in time
  - $X(t+1)$ will depend on value $X(t)$. If $X(t)=x$, $X(t+1)$ can take two values only:
    - $X(t+1)=x+1$
    - $X(t+1)=x-1$

Models

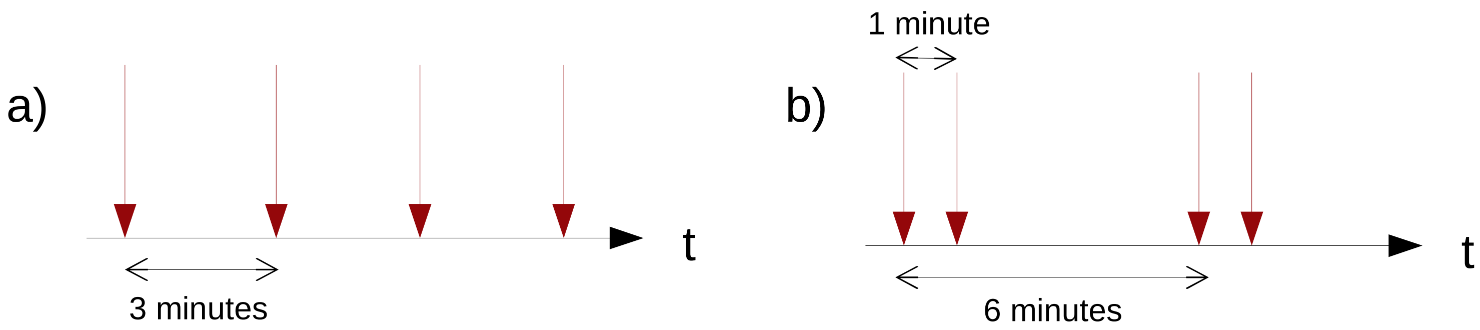
- From real data, we can build models that can be used to statistically reproduce the behavior of a given phenomenon.
  - Those models can be used to study many different systems.



- What to do when there is no data?
  - We have to build them from scratch: i.e., assume the properties and parameters of the models.

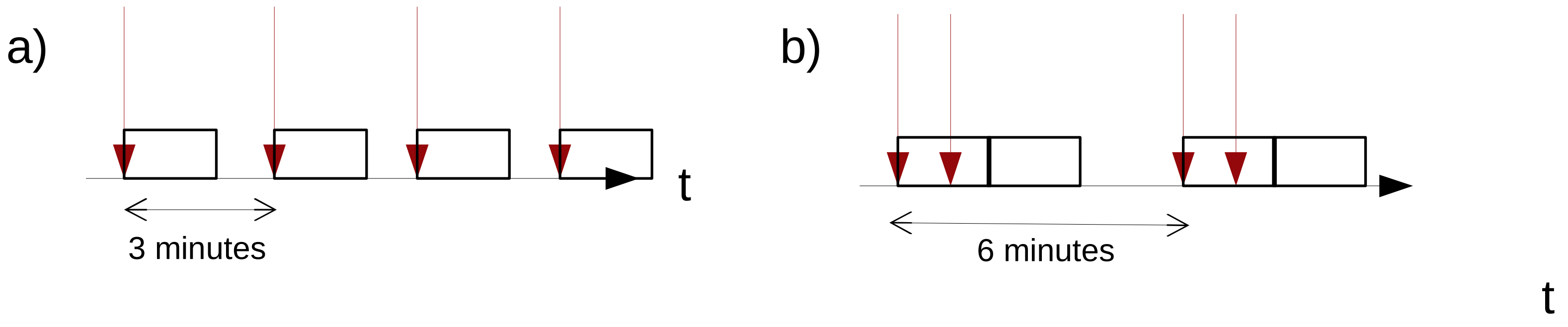
Example: effect of the arrival process in the system perf.

- We have a coffee machine (the system). It makes a coffee in exactly 2 minutes (since the button is press).
- We want to know what is the probability that when a new person arrives to the coffee machine finds it idle.
- People (i.e., requests for coffee) follow two different arrival processes:



Example: effect of the arrival process in the system perf.

- Solution:
  - In case a), the probability that a new person finds the coffee machine idle is 1 (prob = 4 persons find it idle / 4 persons arrive = 1)
  - In case b), the probability that a new person finds the coffee machine idle is 0.5 (prob = 2 persons find it idle / 4 persons arrive = 0.5)



Using a proper 'arrival process' matters!!!