UNIVERSITY OF CAGLIARI

DIEE - Department of Electrical and Electronic Engineering

# STOCHASTIC MODELS
# -
# Stochastic Processes

alessandro.pilloni@unica.it

# Stochastic Processes

- A stochastic process is defined as a collection of indexed random variables (or family of r.v.) defined on a common sample space $S_X$

$$X_t = (S_X, p_{X_t}) \quad \forall \quad t \in T \subset \mathbb{R}$$

Not necessarily IID

- A stochastic process is denoted in short by

$$\{X_t, t \in T\} \equiv \{X_t\}$$

Examples: Every quantity associated with repeated experiment

- Bankroll [€] in a gambling game at time $t$;
- #services request at time $t$;
- Etc…

Notation:

- $X_t$: is the r.v. of the $t$-th experiment. It takes values from $S_X$
- $x_t \in S_X$: is a possible outcome at time t. It is a-priory unknown.
- $T \subseteq \mathbb{I} \subseteq \mathbb{R}$ : is the index set such that $t \in T$ and $t$ is $t$-th time or $t$-th experiment
- $p_{X_t}(t)$, or $f_{X_t}(t)$ denote, resp. the pmf, or pdf associated with $X_t$

# Stochastic Processes Classification:

With respect to the index set $T$:

- if $T$ is countable the SP is called discrete-time SP (DT-SP)

i.e. $X_t$ is the #call for $T \subseteq \mathbb{N}_0$

- Otherwise, it is called continuous-time SP (CT-SP)

i.e. $X_t$ is #call for $T \subseteq \mathbb{R}_{\geq 0}$

- Further note that every CT-SP can be approximated as a DT-SP by means of discretization procedures (see later the differences between DT and CT markov chains…)

Alternatively SP are also classified on the basis of the sample-space cardinality $|S_X|$

- If $S_X$ is countable and finite → finite-discrete-space SP
- If $S_X$ is countable and infinite → infinite-discrete-space SP
- If $S_X$ is uncountable → continuous-space SP

# Examples

Example 3: $\{X_t, t \subseteq \mathbb{N}\}$ represents the person's height on his/her birthday

$X_t \in \mathbb{R}_{>0}$ continuous state space and discrete time SP

Example 4: $\{X_t, t \in \mathbb{R}_{\geq 0}\}$ is the number of people in a queue at time $t$

$X_t \in \{0,1,2, \dots k, \dots\} \subseteq \mathbb{N}_{\geq 0}$ continuous-time, discrete state-space

**From a SP $\{X_t\}$ other SPs with different properties can also be defined!!!**

Example 5: Let $T_k \in \mathbb{R}_{\geq 0}$ be the r.v. describing the $k$-th arrival time of Example 4.

- Thus from Example 4 can also be defined, e.g.:

**Arrival time SP** $\{T_k, k \in \mathbb{N}_{\geq 0}\} \rightarrow$ continuous state space discrete time SP

**Interarrival time SP** $\{\Delta T_k = T_k - T_{k-1}, k \in \mathbb{N}_{\geq 0}\} \rightarrow$ continuous state space discrete-time SP

**Note: Example 5 further show that from a SP can be defined other SPs**

# Stochastic Processes (cont'd)

- **Definition:** We call **realization** $R = \{x_0, x_1, x_2, \dots, x_\tau, \dots\}$ a series of indexed outcomes of the process $\{X_t\}$, for $0 \leq t \leq \tau$

- Clearly each index $t$ must belong to the index-set, i.e., $t \in T$

Example 1: Tossing a fair coin every second $t = 0, 1, 2, \dots$

- Let Head=1, and Tail=0, then $\{X_t, t \in \mathbb{N}\}$
- This process has infinite many realizations.

| $t$ | $X_t$ | $p_{X_t}(t)$ |
|---|---|---|
| 0 | 0 | 0.5 |
| | 1 | 0.5 |
| 1 | 0 | 0.5 |
| | 1 | 0.5 |
| ⋮ | 0 | 0.5 |
| | 1 | 0.5 |

Example of possible realization

$$R = \{x_1, x_2, x_3, x_4, \dots,\}, \{0, 1, 0, 0, 1, \dots\}$$

| $t$ | $x_t$ |
|---|---|
| 0 | 0 |
| 1 | 1 |
| 2 | 0 |
| 3 | 0 |
| 4 | 1 |
| … | … |

# Examples (cont'd)

Example 2: Tossing a coin at time $t = 0$, then leave it in the same position.

**Notice that:** Although this process $\{X_t, t \in T\}$ has only 2 possible realizations **R1** and **R2**, it shares the same description in term of $\{X_t\}$ and of $p_{X_t}(t)$ of Example 1.

R1: $x_n = x_{n-1} = \cdots = x_1 = x_0 = 0$

R2: $x_n = x_{n-1} = \cdots = x_1 = x_0 = 1$ $\quad \forall\, t_k > 0, \quad k = 1, 2, \ldots, n$

Indeed

| $t$ | $X_t$ | $p_{X_t}(t)$ |
|---|---|---|
| 0 | 0 | 0.5 |
| | 1 | 0.5 |
| 1 | 0 | 0.5 |
| | 1 | 0.5 |
| … | 0 | 0.5 |
| | 1 | 0.5 |

$$p_{X_0}(0) = \frac{1}{2}$$

$$p_{X_1}(0) = \Pr(X_1 = 0) = \Pr(X_1 = 0 | X_0 = 0)\Pr(X_0 = 0) + $$
$$+ \Pr(X_1 = 0 | X_0 = 1)\Pr(X_0 = 1) = 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = \frac{1}{2}$$

$$p_{X_k}(0) = \sum_{\forall\, x_{k-1}} \Pr(X_k = 0 | X_{k-1} = x_{k-1})\Pr(X_{k-1} = x_{k-1}) = \frac{1}{2}$$

**Thus, more sophisticated concepts and tools than pmf/pdf need to be introduced to study a SP**

# Joint probability functions and Stochastic Processes

- The more sophisticated tool allowing to study SPs is the Joint Probability
- Similarly, to what seen for discrete r.v., for a discrete-space SPs the Joint pmf is

$$p_{X_{t_0},X_{t_1},\ldots,X_n}(x_0,x_1,\ldots,x_n) = \Pr(X_{t_0}=x_0, X_{t_1}=x_1,\ldots,X_{t_n}=x_n)$$

$$\forall\, x_i \in S_X, \quad i=0,1,2,\ldots,n \qquad\qquad \forall\, t_0 < t_1 < t_2 < \cdots < t_n$$

- Analogously, for a continuous-space SPs, its Joint pdf is as next

$$f_{X_{t_0},X_{t_1},X_{t_2},\ldots,X_{t_n}} : S_X \times S_X \times \cdots \times S_X \mapsto [0,1]$$

$$F_{X_{t_0},X_{t_1},\ldots,X_n}(x_0,x_1,\ldots,x_n) = \int_{-\infty}^{x_0}\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n} f_{X_{t_0},X_{t_1},\ldots,X_{t_n}}(x_{t_0},x_{t_1},\ldots,x_{t_n})\, dx_{t_0}dx_{t_1}\ldots dx_{t_n}$$

$$= \Pr(X_{t_0} \le x_0, X_{t_1} \le x_1,\ldots,X_{t_n} \le x_n)$$

# Joint CDF and Stochastic Processes

- For brevity in notation:
- Let $X = \left[X_{t_1}, X_{t_2}, X_{t_3}, \dots, X_{t_n}\right]$ be random vector belonging to $\{X_t, t \in T\}$
- Let $x = [x_1, x_2, x_3, \dots, x_n]$ be a realization for $X$ where each $x_i \in S_X$
- The Joint pmf/pdf and Joint cdf associated to a SP can be re-written as follows

Discrete-state SP:

$$p_X(x) = p_{X_{t_0}, X_{t_1}, \dots, X_{t_n}}(x_0, x_1, \dots, x_n)$$

$$P_X(x) = P_{X_{t_0}, X_{t_1}, \dots, X_{t_n}}(x_0, x_1, \dots, x_n)$$

Continuous-state SP:

$$f_X(x) = f_{X_{t_0}, X_{t_1}, \dots, X_{t_n}}(x_0, x_1, \dots, x_n)$$

$$F_X(x) = F_{X_{t_0}, X_{t_1}, \dots, X_{t_n}}(x_0, x_1, \dots, x_n)$$

# Marginalization of Joint CDF

- By exploting the Law of Total Probability one can always derive every single probability distribution by properly marginalizing the Joint pmf/pdf/cdf

Discrete-state SPs

$$P_{X_{t_i}}(x_i) = \Pr(X_{t_i} \leq x_i) = \Pr(X_{t_0} \leq \infty, X_{t_1} \leq \infty, \ldots, X_{t_i} \leq x_i, \ldots, X_{t_n} \leq \infty)$$

$$= \sum_{\forall x_{t_0} \in S_X} \sum_{\forall x_{t_1} \in S_X} \cdots \sum_{\forall x_{t_i} \leq x_i} \cdots \sum_{\forall x_{t_n} \in S_X} p_{X_{t_0}, X_{t_1}, \cdots, X_{t_n}}(x_{t_0}, x_{t_1}, \ldots, x_{t_i}, \ldots, x_{t_n})$$

Continuous-state SPs

$$F_{X_{t_i}}(x_i) = \Pr(X_{t_i} \leq x_i) = \Pr(X_{t_0} \leq \infty, X_{t_1} \leq \infty, \ldots, X_{t_i} \leq x_i, \ldots, X_{t_n} \leq \infty)$$

$$= \int_{\forall x_{t_0} \in S_X} \int_{\forall x_{t_1} \in S_X} \cdots \int_{x_{t_i} \leq x_i} \cdots \int_{\forall x_{t_n} \in S_X} f_{X_{t_0}, X_{t_1}, \ldots, X_{t_n}}(x_{t_0}, x_{t_1}, \ldots, x_{t_n})\, dx_{t_0} dx_{t_1} \ldots dx_{t_n}$$

**Marginalization allows to remove the influence of one or more events**

# Let's see if the Joint PMF of Example 1 and 2 differs

- Example 1: Tossing a coin at every second $t = 0,1,2,\ldots$

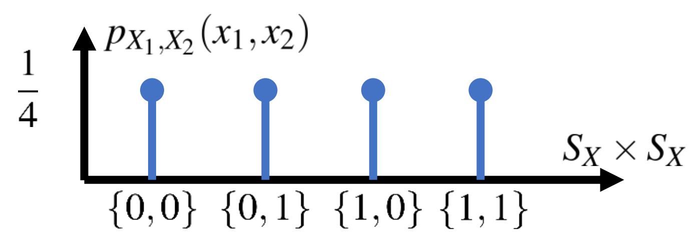
- Since each toss is a Bernoulli trial, then it follows

$@\ t = i$

$$p_{X_i}(0) = \Pr(X_i = 0) = \frac{1}{2} \qquad p_{X_i}(1) = \Pr(X_i = 1) = \frac{1}{2}$$

- The Joint pmf of $X_0$ and $X_1$ for this SP, because of the independence, is

$$p_{X_0,X_1}(i,j) = \Pr(X_0 = i) \cdot \Pr(X_1 = j) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \quad \forall i,j = 0,1$$




- By marginalizing, and because of independence, one obtains

$$p_{X_0}(0) = \Pr(X_0 = 0, X_1 \leq \infty) = \sum_{x_1=0}^{1} \Pr(X_0 = 0 | X_1 = x_1) \cdot p_{X_1}(x_i) = \sum_{x_1=0}^{1} \Pr(X_0 = 0) \cdot p_{X_1}(x_i) =$$

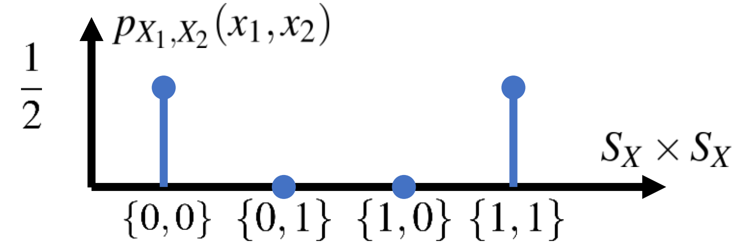
$$= \Pr(X_0 = 0) \cdot \Pr(X_1 \leq 1) = \Pr(X_0 = 0) \cdot P_{X_1}(1) = \frac{1}{2} \cdot 1 = \frac{1}{2}$$

# Let's see if the Joint PMF of Example 1 and 2 differs (cont'd)

- Example 2: Tossing a coin at $t = 0$, then leave it in the same position for $t = 0,1,2,\ldots$
- This SP consists of a single Bernoulli trial at $t = 0$, then $X_t$ depends only by $X_0$
- The resulting Joint pmf of $X_0$ and $X_1$ for this SP

$$p_{X_0,X_1}(0,0) = \Pr(X_0 = 0, X_1 = 0) = \Pr(X_1 = 0 | X_0 = 0) \cdot \Pr(X_0 = 0) = 1 \cdot \frac{1}{2} = \frac{1}{2}$$

$$\Pr(X_0 = i, X_1 = j) : \begin{cases} 0.5 & j = i,\ i = 1,2 \\ 0 & j \neq i \end{cases}$$




- By marginalizing one obtain

$$p_{X_1}(0) = \sum_{x_0=0}^{1} p_{X_0,X_1}(x_0, 0) = \sum_{x_0=0}^{1} p_{X_0,X_1}(X_1 = 0 | X_0 = x_0) p_{X_0}(x_0) = 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = \frac{1}{2}$$

$$p_{X_1}(1) = \sum_{x_0=0}^{1} p_{X_0,X_1}(x_0, 1) = \sum_{x_0=0}^{1} p_{X_0,X_1}(X_1 = 1 | X_0 = x_0) p_{X_0}(x_0) = 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = \frac{1}{2}$$

**Joint pmf/cdf allows to discrimintate among SP with the same marginal pmf/pdf/cdf**

# Independent SPs

- Let $X = [X_0, X_1, X_2, \dots, X_n]$ be a random vector of $\{X_t, t \in T\}$
- Let $x = [x_0, x_1, x_2, \dots, x_n]$ be a vecter which entries are such that $x_i \in S_X$

- Definition: The SP $\{X_t\}$ is called independent their random variables are independent.

- Thus, the Joint cdf, and thus the Joint pdf/pmf have the so-called product-form

- DS-SP:

$$P_X(x) = \Pr(X_0 \leq x_0, X_1 \leq x_1, \dots, X_n \leq x_n) = P_{X_{t_0}}(x_0) \cdot P_{X_{t_1}}(x_1) \cdots P_{X_{t_n}}(x_n)$$

- CS-SP:

$$F_X(x) = \Pr(X_0 \leq x_0, X_1 \leq x_1, \dots, X_n \leq x_n) = F_{X_{t_0}}(x_0) \cdot F_{X_{t_1}}(x_1) \cdots F_{X_{t_n}}(x_n)$$

- **Note:** This is a desired property that is often verified in client-servers system, see later the results on *ergodic queuing networks*

# Let's see if the either Example 1 or 2 are Independent SPs

- Example 1: Tossing a coin at every time $t = 0,1,2,\ldots$

$$\Pr(X_0 = 0, X_1 = 0, \ldots, X_n = 0) =$$

$$= \Pr(X_0 = 0)\Pr(X_1 = 0)\cdots\Pr(X_n = 0) =$$

$$= \left(\frac{1}{2}\right)^n$$

**Independent SP**

$$p_{X_0,X_1}(i,j) = p_{X_0}(i)\cdot p_{X_1}(j) \quad \forall\, i,j$$

- Example 2: Tossing a coin at $t = 0$, then leave as it is it for all $t = 1,2,\ldots$

$$\Pr(X_0 = 0, X_1 = 0, \ldots, X_n = 0) =$$

$$= \Pr(X_n = 0|X_{n-1} = 0)\cdots\Pr(X_1 = 0|X_0 = 0)\Pr(X_0 = 0) =$$

$$= 1\cdots 1\cdot 1\cdot\frac{1}{2} = \frac{1}{2}$$

**Not Independent SP**

$$p_{X_0,X_1}(i,j) \neq p_{X_0}(i)\cdot p_{X_1}(j)$$

$$\frac{1}{2} \neq \left(\frac{1}{2}\right)^2$$

# Stricly VS Wide Sense Stationary SPs

**Definition:** A SP is strictly-sense stationary (SSS) if its **Joint probability** does not change when **shifted in time by** $m$, namely, for all random vectors

$$\bar{X}_t = [X_{t_1}, X_{t_2}, \ldots, X_{t_n}] \qquad \bar{x}_t = [x_{t_1}, x_{t_2}, \ldots, x_{t_n}]$$

- and for all $m \in \mathbb{N}$, it yields

$$\begin{aligned} \text{DS-SP:} \quad & P_{\bar{X}_t}(\bar{x}_t) = P_{\bar{X}_{t+m}}(\bar{x}_t) \\ \text{CS-SP:} \quad & F_{\bar{X}_t}(\bar{x}_t) = F_{\bar{X}_{t+m}}(\bar{x}_t) \end{aligned} \qquad (1)$$

- thus **pdf/pmf/cdf** do not depend by $m$.

IMPORTANT: A different stationarity definition is the WSS (wide sense stationarity).
This holds if, although (1) doesn't hold, it yields
$\mathrm{E}[\bar{X}_t] = \mathrm{E}[\bar{X}_{t+m}]$ and $\mathrm{E}[\bar{X}_t^2] < \infty \quad \forall\, t$

- Consequently

$$\mu = \mathsf{E}[\bar{X}_t] = \mathsf{E}[\bar{X}_{t+m}]\ ,\ \mathsf{Var}[\bar{X}_t] = \mathsf{Var}[\bar{X}_{t+m}] \ (2)$$

- the **mean** and **variance** are also independent to any time shift operation

- **Interpretation:** A SP is **SSS**, if what I can infer today, at $t = 0$, about the **next three days**, has the same **Joint probability** of what I could $25$ years later, for the **next three days** provided that the a-priory information a time $0$ and $25$ are the same.
- **Note:** **SSS** implies **WSS** **if and only if** the **mean** and **variance exist** and are **bounded**
- **Note:** The "**strict stationary**" concept is close to that of "**causality**" for **dynamical systems**

# Examples (cont'd)

- Example 1: Tossing a coin at every time $t = 0,1,2,\ldots$

- Example 2: Tossing a coin at time $t = 0$, then leave it for all $t = 0,1,2,\ldots$

Since the pmf doesn't depend by $t$, then they would not depend by $m$ as well,
Thus, both, **for sure**, are strictly stationary!

- Example 1: let $m = 25, x_0 = x_{25} = 1$, and due to independence, note that

$$p_{X_0,X_1}(1,1) = p_{X_1}(1)p_{X_0}(1) = \frac{1}{2}\cdot 1 \qquad p_{X_{25},X_{26}}(1,1) = p_{X_{26}}(1)p_{X_{25}}(1) = \frac{1}{2}\cdot 1$$

- Example 2: let $m = 25, x_0 = x_{25} = 1$, then

$$p_{X_0,X_1}(1,1) = p_{X_1|X_0}(1|1)p_{X_0}(1) = 1\cdot 1 \qquad p_{X_{25},X_{26}}(1,1) = p_{X_{26}|X_{25}}(1|1)p_{X_{25}}(1) = 1\cdot 1$$

# Examples (cont'd)

- Example 3: Consider a DT SPs with continuous- sample-space

$$\{X_t, t \in \mathbb{N}_{\geq 0}\}$$

- where

$$X_t = A_t \qquad\qquad A_t \sim Uni_c(0,1)$$

- Each $X_t$ is IID and time-invariant because $X_t \sim Uni_c(0,1)$

- Let, e.g., $x_0 = 0.3$ and $x_{25} = 0.3$, then $F_{X_0}(0.3) = \frac{\Pr(x_0 \leq 0.3 | x_0 = 0.3) \cdot \Pr(x_0 = 0.3)}{\Pr(x_0 = 0.3)} = 1$
  Due to **independence**, the Joint cmf satisfies

$$F_{X_0,X_1}(0.3, 0.7) = F_{X_1}(0.7) \cdot F_{X_0}(0.3) = 0.7 \cdot 1$$

$$F_{X_{25},X_{26}}(0.3, 0.7) = F_{X_{26}}(0.7) \cdot F_{X_{25}}(0.3) = 0.7 \cdot 1$$

This SP is SSS since the Joint pdf doesn't depend either by $m$ or by $t$

# Examples (cont'd)

- Example 4: Consider a DT SPs with continuous- sample-space

$$\{X_t, t \in \mathbb{N}\}$$

- where

$$X_t = A_t \cdot t \qquad A_t \sim Uni_c(0,1)$$

Interpretation: $\{X_t\}$ can be viewed as the time-line quantization in the presence of a moltiplicative IID uniform error, e.g.,

$$X_1 = a_1,\ X_2 = 2a_2, \ldots,\ X_t = t \cdot a_t, \ldots$$

- Then, let $x_1 = 0.3$ and $x_{25} = 0.3$, due to **independence**, it results:

$$F_{X_1,X_2}(0.3, 0.7) = F_{X_2}(0.7) \cdot F_{X_1}(0.3)) = \Pr\left(Uni_c(0,1) \leq \frac{0.7}{2}\right) \cdot 1 = \frac{0.7}{2} \cdot 1$$

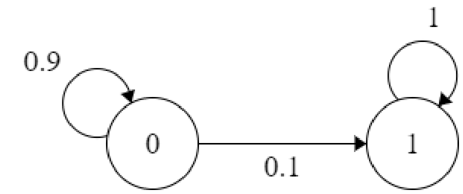
$$F_{X_{25},X_{26}}(0.3, 0.7) = F_{X_{26}}(0.7) \cdot F_{X_{25}}(0.3) = \Pr\left(Uni_c(0,1) \leq \frac{a_1}{26}\right) \cdot 1 = \frac{0.7}{26} \cdot 1 \ \equiv \ \frac{0.7}{m+1}\Big|_{m=25}$$

Since the Joint pdf depends by $m$ the SP is **not SSS**
Moreover, the process is also **NOT WSS** because:

$$E[X_t^2]_{t\to\infty} = E[t^2 \cdot A_t^2]_{t\to\infty} = t^2 \cdot E[A_t^2] = t^2 \int_0^1 a_t^2 \cdot da_t = t^2 \cdot \frac{1}{3}\Big|_{t\to\infty} = +\infty$$

Example 5: Every year a manufacturing system could be in two states

Let the transition probability from the two states, every year, be




$$X_t = 0 \rightarrow Healthy \qquad X_t = 1 \rightarrow Faulty$$

Let $\{X_t, t\}$ be the process describing the system's status at each year and $X_0 = 0$

$$\Pr(X_{t+1} = 0 | X_t = 0) = \pi = 0.9 \qquad \Pr(X_{t+1} = 1 | X_t = 1) = 1$$

- Although probabilities are time-varying, i.e. note that

$$p_{X_t}(0) = (\pi)^t \sim Geo(\pi) \quad \text{with} \quad X_0 = 0 \qquad (1)$$

- after a shifting of $m > 0$ steps, and suppose $x_0 = x_{25} = 0$, it yields

$$p_{X_0,X_1}(0,0) = p_{X_1|X_0}(0|0)p_{X_0}(1) = \pi \cdot 1 \qquad p_{X_{25},X_{26}}(0,0) = p_{X_{26}|X_{25}}(0|0)p_{X_{25}}(0) = \pi \cdot 1$$

- Since the **joint pmf** is independent by a time-shift $m$ this **SP is SSS**

Note: the dependence by $t$ of the Joint PMF is not a problem

- Furthermore note that **SSS** clearly implies $\mathrm{E}[X_t] = \mathrm{E}[X_{t+m}]$

- **Moreover, because of** $\lim_{t \to \infty} \mathrm{E}[X_t^2] = 0 \cdot \pi^t + 1(1 - \pi^t) = 1 < +\infty$ **, it is also WSS**

- **Note:** This is also a consequence the SP may assume a finite number of states only

# Mean Ergodic Stochastic Processes

**INTERPRETATION:** A WSS SP is mean-ergodic (or simply ergodic) if its statistical properties (mean & variance) are not only bounded due to stationarity, but they can also be deduced by analyzing a **single sufficiently long SP's realization in place of many.**

**Definition:** Let $R_i = [x_{i,1}, x_{i,2}, \dots, x_{i,n}] \in \mathbb{R}^n$ be the $i$-th realization of a WSS SP of $n$ samples

Let $\mu_X = \lim_{t\to\infty} E[X_t]$ be its **long-term statistical mean,** due to stationarity: $\mu_X < \infty$

Consider the **limit of the sample mean** associated with $R_i$ as $t \to \infty$

DT-SP $$\hat{\mu}_{i,X} = \lim_{n\to\infty} \frac{\sum_{t=1}^{n} x_{i,t}}{n}$$

CT-SP $$\hat{\mu}_{i,X} = \lim_{t_{i,n}\to\infty} \frac{\sum_{k=1}^{n} x_{i,k} \int_{t_{i,t-1}}^{t_{i,t}} dt}{t_{i,n} - t_{i,0}}$$

The stochastic process $\{X_t\ , t \in\ T\}$ is said to be mean-ergodic if the following holds true

- 1) The limit **exists** $\forall\ i$ (existence implies boundedness);
- 2) The limit is **unique** (all sample means converges to same value, i.e. $\hat{\mu}_{i,X} = \hat{\mu}_X\ \forall\ i$)
- 3) $\hat{\mu}_X = \mu_X$ (all sample means converges to the statistical mean)

# Mean Ergodic Stochastic Processes (cont'd)

- In few word, a SP is ergodic if its mean, variance, etc can be deduced by studying, in place of many long realizations of it, a single, sufficiently long, realization.

- Example 1: Tossing a coin at every time $t = 0,1,2,\ldots$

Mean $$\mu_X = \lim_{t\to\infty} \mathsf{E}[X_t] = \lim_{t\to\infty}\left(\sum_{k=0}^{1} k\cdot\frac{1}{2}\right) = \lim_{t\to\infty}\frac{1}{2} = \frac{1}{2} \quad \forall\, t \geq 0$$

**Ergodic in the mean**

Every sufficiently long realization satisfies $\hat{\mu}_{i,X} = \lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n} x_t = \frac{1}{2}$

- Example 2: Tossing a coin at $t = 0$, then leave it for all $t = 0,1,2,\ldots$

Mean $$\mu_X = \lim_{t\to\infty} \mathsf{E}[X_t] = \sum_{k=0}^{1} k\cdot\frac{1}{2} = \frac{1}{2} \quad \forall\, t \geq 0$$

whereas: $$\hat{\mu}_{i,X} = \lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n} x_t = \begin{cases} 0 & if \quad X_0 = 0 \\ 1 & if \quad X_0 = 1 \end{cases}$$

**Not Ergodic in the mean**
The limit exists but it's not unique. It depends on the realization!

# Testing ergodicity with Example 1 and 2

```
rng(1) % Fix the seed of the random generator
n=10^5; % Realization length
```

Example 1: Tossing a coin at every time $t = 0,1,2,\dots$

```
% Generate 4 realization of length n
Ri_n_EX1=unidrnd(2,n,4)-1;
sample_mean_i_EX1=mean(Ri_n_EX1) % ->:0.4999 0.5013 0.4993 0.5001
```

Example 2: Tossing a coin at $t = 0$, then leave it for all $t = 0,1,2,\dots$

```
% Generate 4 realization of length n
Ri_n_EX2=zeros(n,4);     % Matrix allocation 4 x n
Ri_0=unidrnd(2,1,4)-1;   % Observation at t=0

for i=1:n
    Ri_n_EX2(i,:)=Ri_0;
end
sample_mean_i_EX2=mean(Ri_n_EX2) % ->:  1 0 0 0
```

# Testing ergodicity with Example 5

Example 5: A manufacturing process could be in two states

$$X = 0 \rightarrow Healthy \qquad\qquad X = 1 \rightarrow Faulty$$

where

$$\Pr(X_{t+1} = 0 | X_t = 0) = \pi = 0.9 \qquad\qquad \Pr(X_{t+1} = 0 | X_t = 1) = 1$$

$$p_{X_t}(0) = \pi^t \quad , \quad p_{X_t}(1) = 1 - \pi^t \quad \Longrightarrow \mathrm{E}[X_t] = 0 \cdot \pi^t + 1 \cdot (1 - \pi^t) = (1 - \pi^t)\big|_{t \to \infty} = 1$$

```
% Generate 4 realization of length n
Ri_n_EX5=zeros(n,4); % Matrix allocation 4xn

Ri_0_EX5=binornd(1,0.1,1,4) % Observation at t=0; Bin(1,0.1)==Ber(p)
Ri_n_EX5(1,:)=Ri_0_EX5;

for j=2:n
    for i=1:4
        if Ri_n_EX5(j-1,i)==0
            Ri_n_EX5(j,i)=binornd(1,0.1,1,1);
        else
            Ri_n_EX5(j,i)=1;
        end
    end
end
sample_mean_i=mean(Ri_n_EX5) % -> 0.9999  1.0000  1.0000  1.0000
```

# Markovian stochastic processes

- Definition: A SP $\{X_t, t \in T\}$ is Markovian if, for any time $t = t_0 < t_1 < t_2 < \cdots < t_n$, the conditional probability of moving to the next state at time $t_{n+1}$ depends only on the last observed state $x_n$ and not by the entire history of events $x_{n-1}, x_{n-2}, \dots, x_1, x_0$.

- [DT-SP] A Discrete-time SP satisfies the Markov property if

$$\Pr(X_{t_{n+1}} = x_{n+1} | X_{t_n} = x_n, \dots, X_{t_0} = x_0) = \Pr(X_{t_{n+1}} = x_{n+1} | X_{t_n} = x_n)$$

$$\forall\, t_0 < t_1 < \cdots < t_n < t_{n+1} \qquad \forall\, x_i \in S_X$$

- [CT-SP] Let $x_{[t_0,t]}$ be the history of observed events of a CT-SP from $t_0$to $t$, which latest sample is such that $x(t) = x_i \in S_X$, a continuous-time SP satisfy the Markov property if

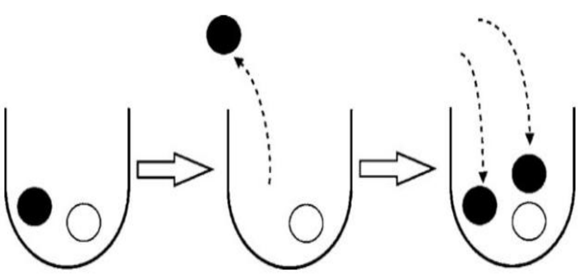
$$\Pr(X_{t+dt} = x_{t+dt} | X_{[t_0,t]} = x_{[t_0,t]}) = \Pr(X_{t+dt} = x_j | X_t = x_t)$$

$$\forall\, dt > 0 \qquad \forall\, x_i \in S_X$$

# Examples of Time Nonhomogenous Markovian SP

Example 6 (Pólya's urn model): Consider an urn with 1 black and 1 white ball.

- At each second $t = 1,2,\ldots$ a ball is drawn.
- After each drawn, 2 balls of the same color of the extracted one are added
- Let $\{B_t, t \in \mathbb{N}_{\geq 0}\}$ the **SP of counting** the **#black balls in the urn** at time $t \geq 0$



Although the total number of black balls at time $t$ depends by the history of extracted balls, the conditional probability **satisfies** the **Markov property** (it is a *time-inhomogenous Markov process*)

$$\Pr(B_{t+1} = j+1 | B_t = j) = \frac{j}{2+t} = \frac{B_t}{2+t} \quad \forall t = 0,1,\ldots, j = 1,\ldots$$

Further note that: Althouth $\{B_t, t\}$ **is Markov**, it is **neither SSS**

$$\Pr(B_{t+1} = j+1 | B_t = j) \neq \Pr(B_{t+1+m} = j+1 | B_{t+m} = j),$$

**nor WSS** (intuition says $\mathrm{E}[B_t]|_{t\to\infty} = +\infty$ and that $\mathrm{E}[B_t] \neq \mathrm{E}[B_{t+m}] > \mathrm{E}[B_t]$) and thus also

$$\mathrm{Var}[B_t]\Big|_{t\to\infty} = +\infty$$

# Example 6 (cont'd)

Example 6 (Pólya's urn model): Let us proof it. To do so let us introduce the next concept:

**Law of Total Expectation:** $\mathsf{E}[A] = \mathsf{E}[\mathsf{E}[A|B_i]] = \sum_{\forall i} \mathsf{E}[A|B_i] \cdot \Pr(B_i)$

**Proof:**

$$\mathsf{E}[A] = \sum_{\forall a} a \cdot \overbrace{\left( \sum_{\forall i} \Pr(A = a|B_i) \cdot \Pr(B_i) \right)}^{\Pr(A)} = \sum_{\forall i} \overbrace{\left( \sum_{\forall a} a \cdot \Pr(A = a|B_i) \right)}^{\mathsf{E}[A|B_i]} \cdot \Pr(B_i) = \sum_{\forall i} \mathsf{E}[A|B_i] \cdot \Pr(B_i)$$

Thus, note that

$$\mathsf{E}[B_{t+1}|B_t] = \sum_{\forall b_{t+1}} b_{t+1} \cdot \Pr(B_{t+1} = b_{t+1}|B_t) = \overbrace{(B_t + 1)}^{b_{t+1}=b_t+1} \cdot \frac{B_t}{2+t} + \overbrace{B_t}^{b_{t+1}=b_t} \cdot \left(1 - \frac{B_t}{2+t}\right)$$

Then

$$= \frac{B_t^2 + B_t + 2B_t + t \cdot B_t - B_t^2}{2+t} = \frac{(3+t)}{2+t} \cdot B_t$$

$$\mathsf{E}[B_0] = 1$$

$$\mathsf{E}[B_1] = \mathsf{E}[\mathsf{E}[B_1|B_0]] = \frac{(3+t)}{2+t} \cdot B_t \bigg|_{t=0, B_0=1} \times \Pr(B_0 = 1) = \frac{3}{2} \times 1 \equiv \overbrace{1}^{\mathsf{E}[B_0]} + \overbrace{\frac{1}{2}}^{\mathsf{E}[B_1 - B_0]}$$

$$\mathsf{E}[B_2] = \mathsf{E}[\mathsf{E}[B_2|B_1]] = \mathsf{E}[B_2|B_1 = 1] \cdot \Pr(B_1 = 1) + \mathsf{E}[B_2|B_1 = 2] \cdot \Pr(B_1 = 2)$$

$$= \frac{(4) \cdot 1}{3} \bigg|_{t=1, B_1=1} \overbrace{\frac{1}{2}} + \frac{(4) \cdot 2}{3} \bigg|_{t=1, B_1=2} \overbrace{\frac{1}{2}} = \frac{6}{3} = 2 \equiv \overbrace{1}^{\mathsf{E}[B_0]} + \overbrace{\frac{1}{2} + \frac{1}{2}}^{\sum_{k=0}^{1} \mathsf{E}[B_{k+1} - B_k]}$$

# Example 6 (cont'd)

Example 6 (Pólya's urn model): Let us proof it. To do so let us introduce the next concept:

**Law of Total Expectation:** $\mathsf{E}[A] = \mathsf{E}[\mathsf{E}[A|B_i]] = \sum_{\forall\, i} \mathsf{E}[A|B_i] \cdot \Pr(B_i)$

**Proof:**

$$\mathsf{E}[A] = \sum_{\forall\, a} a \cdot \left( \sum_{\forall\, i} \Pr(A = a|B_i) \cdot \Pr(B_i) \right) = \sum_{\forall\, i} \left( \sum_{\forall\, a} a \cdot \Pr(A = a|B_i) \right) \cdot \Pr(B_i) = \sum_{\forall\, i} \mathsf{E}[A|B_i] \cdot \Pr(B_i)$$

$$\mathsf{E}[B_{t+1}|B_t] = \sum_{\forall\, b_{t+1}} b_{t+1} \cdot \Pr(B_{t+1} = b_{t+1}|B_t) = \overbrace{(B_t + 1)}^{b_{t+1} = b_t + 1} \cdot \frac{B_t}{2+t} + \overbrace{B_t}^{b_{t+1} = b_t} \cdot \left(1 - \frac{B_t}{2+t}\right)$$

$$= \frac{B_t^2 + B_t + 2B_t + t \cdot B_t - B_t^2}{2+t} = \frac{(3+t)}{2+t} \cdot B_t$$

- Thus, by induction it results

$$\mathsf{E}[B_{t+1}] = \mathsf{E}[\mathsf{E}[B_{t+1}|B_t]] = \overbrace{1}^{\mathsf{E}[B_0]} + \overbrace{t \cdot \frac{1}{2}}^{\sum_{k=0}^{t} \mathsf{E}[B_{k+1} - B_k]} = \frac{2+t}{2} \implies \mathsf{E}[B_t] = \frac{2 + (t-1)}{2} = \frac{1+t}{2}$$

$$\mathsf{E}[B_t] \neq \mathsf{E}[B_{t+m}] \quad \forall\, m \in \mathbb{N} \qquad\qquad \lim_{t\to\infty} \mathsf{E}[B_t] = +\infty$$

# Examples of time-homogenous Markovian SP

- Example 5: Every year a manufacturing system could be in two states

$$X_t = 0 \rightarrow Healthy \qquad\qquad X_t = 1 \rightarrow Faulty$$

- Let the transition probability from the two states, every year, be

$$\Pr(X_{t+1} = 0 | X_t = 0) = \pi = 0.9 \qquad\qquad \Pr(X_{t+1} = 1 | X_t = 1) = 1$$

- Let $\{X_t, t\}$ be the process describing the system's state at each year
- It is easy conclude **this SP is markovian** because the conditional probabilities to move to the **faulty state depend only** by the **actual state**, and **not** by **whole history**

- This implies **memorylessness**, and it's confirmed by noticing the probability the system will move to «1» at a time $t$ from $X_0 = 0$ follows $\pi^t(1-\pi)$, namely $Geo(1-\pi)$, i.e.

$$T = \overbrace{\min(t \geq 0 \; : \; X_t = 1 \mid X_0 = 0) \sim Geo(1-\pi)}^{\text{r.v. describing the minimum time to leave 0}} \implies \mathsf{E}[T] = \frac{1}{1-\pi} = \frac{1}{1-0.9} = 10 \text{ years}$$

**Example of related problem:** When planning the maintenance time $\bar{t}$ to ensure that the **probability of the system remain healthy never drops** below **70%**?

This is equivalent to maximize $\bar{t}$ such that:

$$\max \bar{t} > 0 \; : \; \underbrace{\Pr(T > \bar{t}) > 0.7}_{\text{pr. still in 0 after } \bar{t}} \implies \Pr(T \leq \bar{t}) \leq 0.3$$

$$\bar{t} = 0\text{y} \; : \; \Pr(T \leq 0) = 0.1$$

$$\bar{t} = 1\text{y} \; : \; \Pr(T \leq 1) = 0.19$$

$$\bar{t} = 2\text{y} \; : \; \Pr(T \leq 2) = 0.271$$

$$\bar{t} = 3\text{y} \; : \; \Pr(T \leq 3) = 0.3439$$

# The Markov property ⇒ Memorylessness of event occurence

Every **Markov SP** has **memoryless events!!!**

In fact, the **memoryless property** is referred to random variables (**not to processes**),

**Note:** Having **memoryless events** has 3 fundamental implications

1. No state memory: All past observed states are not relevant in determining the future possible state (outcomes)
2. No state age memory: How long the process has been in a given state is also irrelevant.
3. Memoryless interevent-time: Let $t_i$ be the time the next event will then

$$\Pr(\, t_i \; > \; s + t \mid t_i > s \,) = \Pr(\, t_i > t \,)$$

Example: Consider the process **counting** the #purchases on an online-shop $\{X_t, t \in \mathbb{R}_{\geq 0}\}$, since purchases can be assumed independent, at given time $t$ then the r.v. $X_t \sim Pois(\lambda t)$.

Prove Condition 3 holds.

- Let $t_i$ be the r.v. associated with the time passed since the last purchase at time $t_{i-1}$, then

$$\Pr(\, t_i > t_{i-1} + dt \mid t_i > t_{i-1} \,) = \frac{\Pr(\, t_i > t_{i-1} + dt \; \cap \; t_i > t_{i-1} \,)}{\Pr(t_i > t_{i-1})}$$

$$= \frac{\Pr(\, t_i > t_{i-1} + dt \,)}{\Pr(t_i > t_{i-1})} = \frac{e^{-\lambda(t_{i-1}+dt)}}{e^{-\lambda t_{i-1}}} = e^{-\lambda dt} \implies \overbrace{\Pr(t_i > dt)}^{\text{comp. cdf of a Exp. r.v.}}$$

# Generalized Semi-Markov Processes (GSMP)

- The GSMPs is a class of SP where the requirement of absence of state age is relaxed
- Here, transition probabilities are **functions** of the **time $d\tau$ spent in the actual state**
- Here, under certain **assumptions**, it can be found an ad-hoc **"state transformations"** making every GSMPs **explicitly rewritten** as a Markov process

Example: Consider the DT-SP

$$X_{k+1} = X_k - X_{k-1} \qquad \text{with} \qquad X_k \sim Ber(\pi) \qquad \text{and} \qquad X_{-1} = X_0 = 0$$

- Is this a Markov process?

**NO** ($X_{k+1}$ depends by $X_k$ and $X_{k-1}$) but it is **GSMP because** $X_{k+1}$ depends by something happened at time $k$ and $k-1$

- Is it possible to show the Markov property hold?
- Let $Y_k = X_{k-1}$ and $Z_k = \begin{bmatrix} Y_k \\ X_k \end{bmatrix}$ note that ➡

$$\overbrace{\begin{bmatrix} Y_{k+1} \\ X_{k+1} \end{bmatrix}}^{Z_{k+1}} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} \cdot \overbrace{\begin{bmatrix} Y_k \\ X_k \end{bmatrix}}^{Z_k}$$

A CT-GSMP is a Markov process **if and only if** its conditional probability satisfies:

$$\Pr(\, X_{t+dt} = x_j \mid X_t = x_i \,) = \lambda_{ij} dt$$

*(cf. with CT Markov chains where $\lambda_{ij}$ is the mean rate of events allowing to move from $x_i$ to $x_j$)*

where $\lambda_{ij}$ must be a **constant** depending **only** by state $x_i$ and $x_j$

# Counting Processes

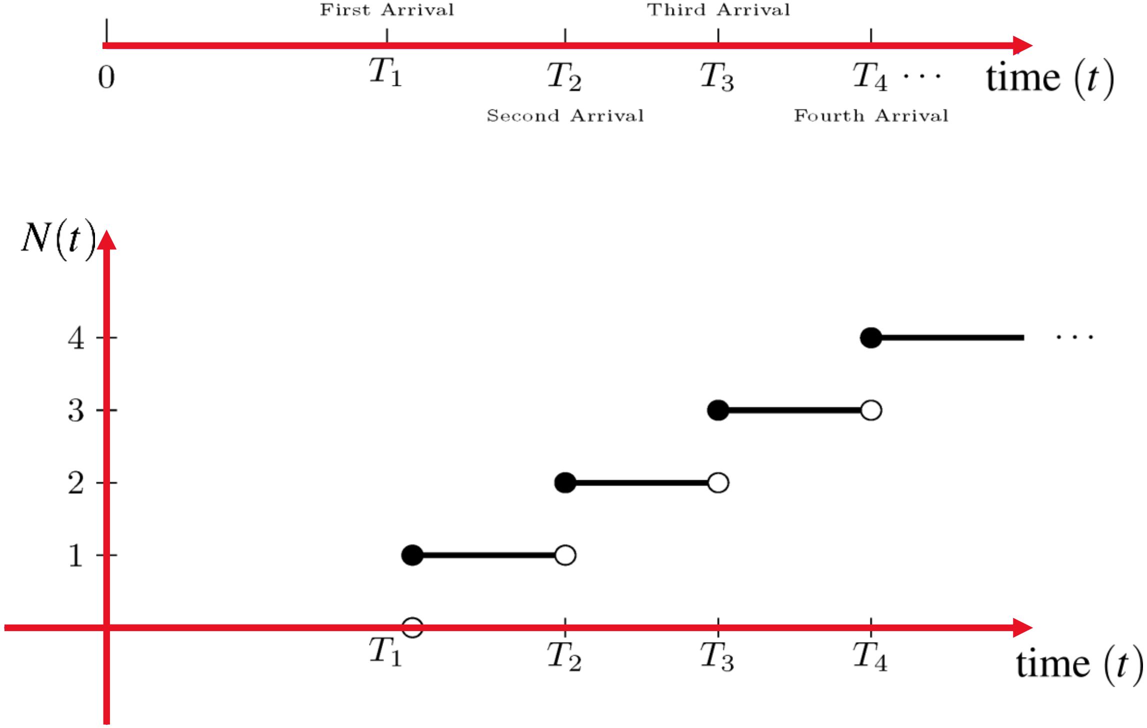
A stochastic process $\{N(t), t \geq 0\}$ is a counting process if it count the occurrences of something over as the time passes

Its sample space is **always discrete** and **non-negative**,

even though the events can occur either in a continuous or discrete time setting.

Examples of counting process:

- Arrivals to a queue
- Waking up during night
- Traffic accidents in a day
- Laps during a Gran Prix

# Counting Processes (cont'd)

- Definition: A counting process is a stochastic process $\{N(t), t \geq 0\}$ which values are always **non-negative**, **integer**, and **non-decreasing**, namely

1. $N(t) \geq 0$
2. $N(t) \in \mathbb{N}_{\geq 0} \equiv \{0\} \cup \mathbb{N}$
3. *If* $t' \leq t''$ *then* $N(t') \leq N(t'')$

- Remark: Counting SPs $N(t)$ are **not independent SP** because

$$\Pr(N(u) = j | N(v) = k) \neq \Pr(N(u) = j) \qquad u > v$$

- but they **may have** independent increments

- Example: If $N(t)$ is a counting process, and $t'' > t'$, then $N(t'')$ on $[0, t'')$, depends on the number of arrivals $N(t'' - t')$

# Counting process with **independent** and **stationary** increments

- Definition: A counting process $\{N(t), t \geq 0\}$ has independent increments if

$$\forall \quad s > t > u > v > 0$$

- the r.v. $(N(s) - N(t))$ and $(N(t) - N(u))$ and $(N(u) - N(v))$ are independent.

- **In other works:** the numbers of arrivals in non-overlapping (disjoint) time-intervals,

$$(v,u] \ , \ (u,t] \ , \ (t,s]$$

are independents to each other.

- Definition: A counting process $\{N(t), t \geq 0\}$ has stationary increments if $\forall \ t \geq 0$ and $h > 0$, the distribution of the random variables

$$N(t+h) - N(t) \sim N(h)$$

- depends only on $h > 0$ and **not** on $t$.

**A CT counting process with stationary increments is a Markov SP**

# Orderly counting process

- Property: A counting process $\{N(t), t \geq 0\}$ is **orderly** if the probability that two or more arrivals occurs at once is negligible.

**Events cannot occur simultaneously**

- Mathematically, for a CT counting process to be orderly, one has

$$\lim_{dt \to 0} \Pr(\ N(t+dt) - N(t) > 1 \mid N(t+dt) - N(t) \geq 1\ ) = 0$$

- Interpretation: The probability a second arrival occurs approaches 0 as $dt \to 0$.

- Example: Consider a SP counting independent arrivals at a mean rate $\lambda$ (

  (i.e. a Poisson process ☺)

$$\Pr(\ N(t+dt) - N(t) > 1 \mid N(t+dt) - N(t) \geq 1\ )$$

$$\underbrace{= \Pr(N(dt) > 1)}_{\text{(due to staz. incr.)}} = 1 - \frac{(\lambda dt)^0}{0!} e^{-\lambda dt} - \frac{(\lambda dt)^1}{1!} e^{-\lambda dt} \Big|_{dt \to 0} \to 0$$

# Poisson Processes

- **Definition:** A CT Counting Process $\{N(t), t \in [0, \infty)\}$ is defined as a Poisson process with rate $\lambda > 0$ **if and only if**:

i. $N(0) = 0$

ii. it has independent increments

iii. $\forall\, t \geq 0, \forall\, h > 0$, the number of occurrences in an interval of length $h$, denoted by $N(t+h) - N(t)$ has a Poisson distribution with parameter $\lambda h$ (i.e. it has stationary increment and satisfy the orderliness property)

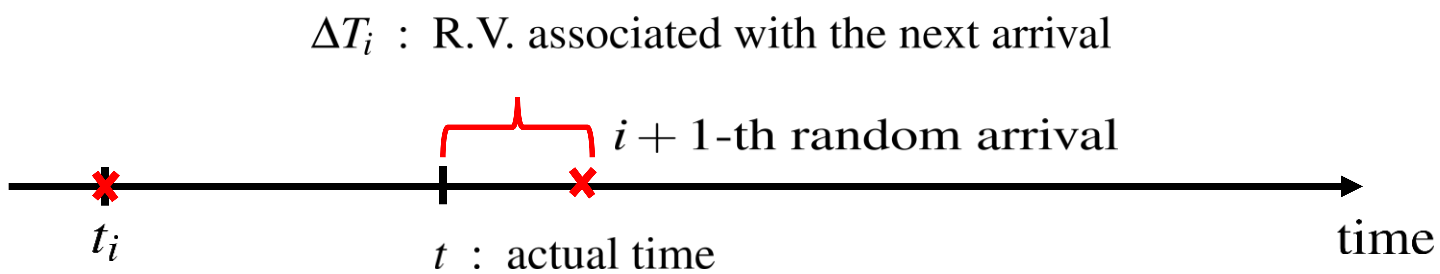
- In a **Poisson Process** the **occurrence of an event**, is commonly referred as "**arrival**". More precisely, at each time each r.v. $N(t) \sim Pois(\lambda t)$, i.e.:

$$p_{N(t)}(n) = \Pr(N(t) = n) = \begin{cases} \frac{(\lambda t)^n}{n!} \cdot e^{-\lambda t} & \lambda \in \mathbb{R}^+, n \in \mathbb{N}_{\geq 0} \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_{N(t)} = \mathsf{E}[N(t)] = \lambda t \qquad \sigma^2_{N(t)} = \mathsf{Var}[N(t)] = \lambda t$$

# Properties of a Poisson Process (cont'd)

- **IMPORTANT:** The inter-arrivals time of a Poisson process with parameter $\lambda$ are independent and thus, they are exponentially distributed
- **Proof:** Let $t$ be the actual time where $N(t) = N(t_i)$, and let $\Delta T_i = t_{i+1} - t$ be the random time we need for the next arrival, this is clearly a r.v. since unknown at time $t$.




- Because of **increments** are **independent**, and **stationary**, then

$$\big(N(t_i + dt) - N(t_i)\big) \text{ and } N(dt)$$

- have the same distribution, that is $Pois(\lambda dt)$, thus

$$\Pr(\Delta T_i > dt) = \Pr(N(dt) = 0) = \frac{(\lambda dt)^0}{0!} = \mathrm{e}^{-\lambda dt} \Rightarrow F_\Delta(t) = \Pr(\Delta T \le dt) = 1 - e^{-\lambda dt}$$

$$\Rightarrow f_{\Delta T_i}(dt) = \frac{dF_{\Delta T_i}(dt)}{dt} = \lambda \cdot \mathrm{e}^{-\lambda dt} \implies \Delta T_i \sim Exp(\lambda dt)$$

# Properties of a Poisson Process (cont'd)

- In a Poisson process the following are equivalent
- Poisson arrivals

$$p_{N(t)}(n) = \Pr(N(t) = n) \; = \begin{cases} \frac{(\lambda t)^n}{n!} \cdot e^{-\lambda t} & \lambda \in \mathbb{R}^+, n \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$$

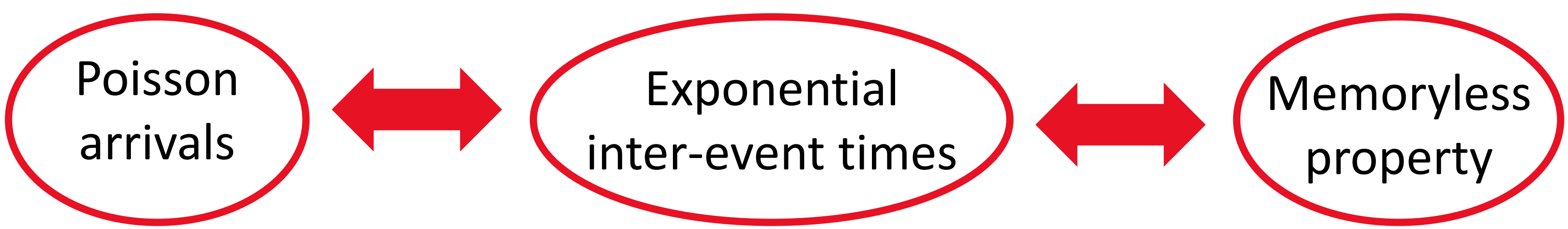
- Exponential inter-event times

$$f_{\Delta T_i}(dt) = \frac{dF_{\Delta T_i}(dt)}{dt} = \lambda \cdot \mathrm{e}^{-\lambda dt}$$

- Memoryless property

$$\Pr(\Delta t > t + dt \mid \Delta t > t) \; = \; \Pr(\Delta t > dt)$$

**Interpretation:** After waiting 1 minute without a call, the probability of a call arriving in the next 2 minutes is the same as was a minute ago, of getting a call in the following 2 minutes



# Poisson process  (see Ass 3, Ex. 6)

```
n=1e6;
lambda=40;
x=rand(n,1);

y = -lambda^-1*log(1-x);

t=zeros(n,1);
for i=2:n
    t(i)=t(i-1)+y(i-1);
end

z= poissrnd(lambda*y);
```

- Simulation of a Poisson process $\lambda = 40$ arrival/sec

- Inter-event time process

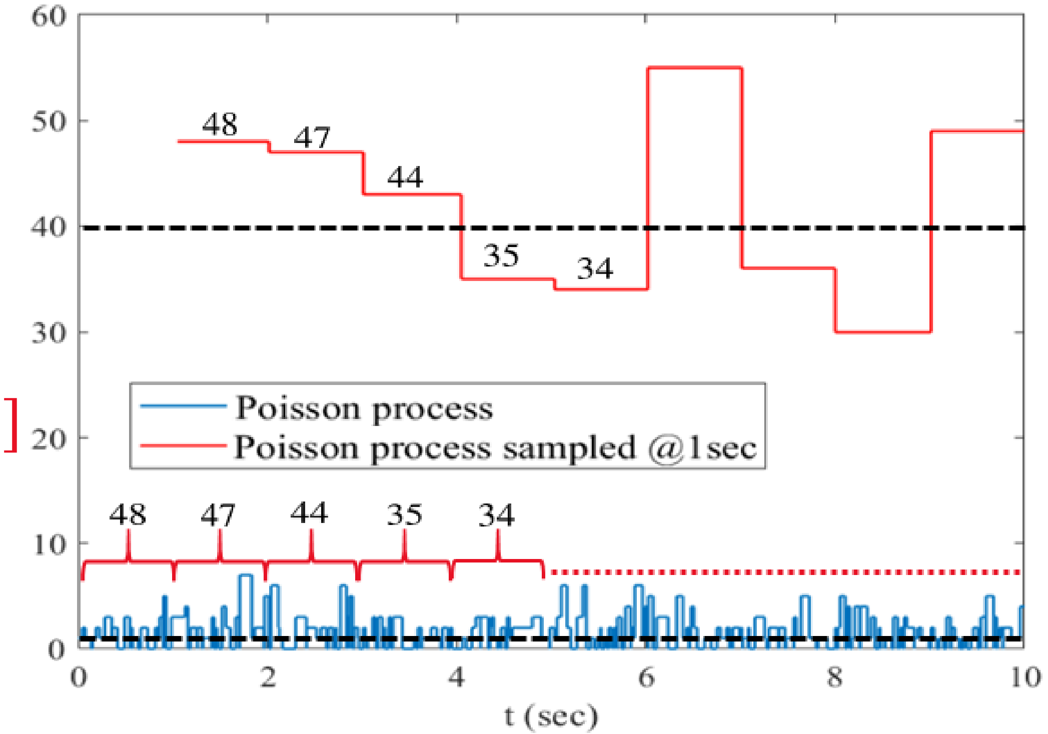
$$\{\Delta t_k, t_k \in \mathbb{R}_{\geq 0}\} \quad : \quad \Delta t_k \sim Exp(\lambda) \implies \mu_{\Delta t_k} = \mathsf{E}[\Delta t_k] = \frac{1}{\lambda} = 0.025sec$$

- Expected arrival rate in arrival/sec

$$\begin{aligned} \mathsf{E}[N(\Delta t)]|_{\Delta t = 1sec} &= \lambda \cdot \Delta t \\ &= 40 \frac{\text{arrival}}{\text{sec}} \cdot 1\text{sec} \\ &= 40\text{arrival} \end{aligned}$$

- Expected arrival rate in arrival/$\mathsf{E}[\Delta t_k]$

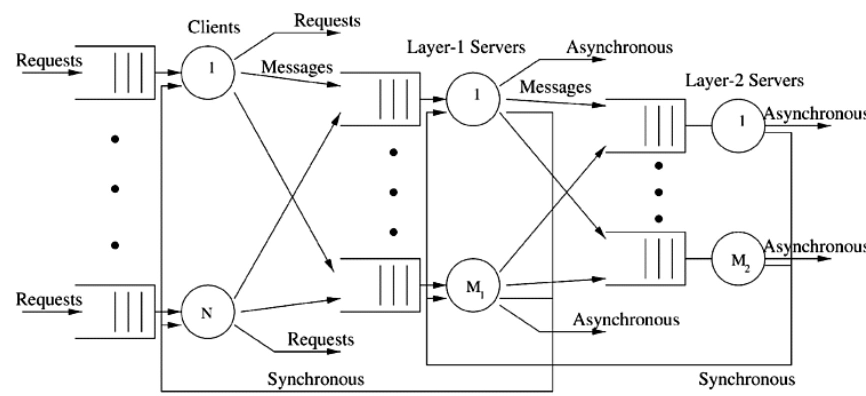
$$\begin{aligned} \mathsf{E}[N(\Delta t)]|_{\Delta t = \mu_{\Delta t_k}} &= \lambda \cdot \mathsf{E}[\Delta t] \\ &= 40 \frac{\text{arrival}}{\text{sec}} \cdot 0.025\text{sec} \\ &= 1\text{arrival} \end{aligned}$$



- Question: Is the process of queuing of cars at a traffic light Poisson distributed?
- If the traffic light is far YES, because the arrivals would be independent.

Otherwise, if it is sufficiently close to another traffic light, then the cars will arrive in small groups (i.e. exists correlation) and thus they are not independent.

- Nevertheless, we are going to see even **coupled queuing systems**, known as Queueing Networks (**QNs**), under certain conditions (see **ergodic QNs**), exhibit, at least at the **steady-state** (as $t \to \infty$), **certain independence characteristics** that allow for the easy characterization of the overall system performance.




Multi-services system topology as a QN

In **ICT client-server architectures**, most resource allocation problems involve **independent arrivals from outside** and **independent services**.

This is sufficient to characterize and allocate resources to meet certain **QoS** requirements.

# Superposition of Poisson processes

- **Theorem:** Let $\{N_i(t), t \in \Re_{\geq 0}\}$ with $i = 1,2, \dots, n$, be $n$ Poisson processes with rates $\lambda_i > 0$. Let us consider their super-position

$$\left\{ N(t) = \sum_{i=0}^{n} N_i(t), t \in [0, \infty) \right\}$$

- Then, $N(t)$ is again a Poisson process that counts all those events, and with a rate

$$\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_n \qquad \blacksquare$$

- The inter-event time of $N(t)$ can be viewed as the time of the winner among the competition of $n$ exponential r.v. $\Delta t_i$, no matter he is the winner $N_i(t)$

$$T = \min\{T_1, T_2, \dots, T_n\}$$

- Assume no events are occurred in $[0, t]$, and $\Delta t$ is the **1st-arrival time**

$$\Pr(T > t) = \Pr(\min\{T_1, T_2, \dots, T_n\} > t) = \Pr(T_1 > t, \dots, T_n > t) = \Pr(T_1 > t) \cdots \Pr(T_n > t)$$

$$= \Pr(N_1(t) = 0) \cdots \Pr(N_n(t) = 0) \cdots = e^{-(\sum_{i=1}^{n} \lambda_i)t}$$ (due to independence)

- **Thus**

$$\mathrm{E}[T] = \frac{1}{\lambda} = \frac{1}{\lambda_1 + \lambda_2 + \cdots + \lambda_n}$$

# Examples (cont'd)

- Example: Consider the superposition of 2 Poisson processes with rates $\lambda_1$, $\lambda_2$.
- We are interested on the competition between the 2 processes, and $N_1(t)$ won
- Which is the probability the next event belongs to process $N_1(t)$?

- Assume, for simplicity, no events are occurred in $[0, \mathrm{t}]$.

$$\Pr(T_1 < T_2) = \Pr(T_1 < t_2, T_2 = t_2 \; \forall \, t_2 \in S_{T_2}) = \qquad \text{(by marginalizing the Joint CDF)}$$

$$= \int_0^{t_2} \int_0^{+\infty} f_{T_1,T_2}(t_1,t_2) dt_1 dt_2 = \int_0^{t_2} \int_0^{+\infty} f_{T_1}(t_1) f_{T_2}(t_2) dt_1 dt_2$$

$$= \int_0^{+\infty} f_{T_2}(t_2) \overbrace{\left( \int_0^{t_2} f_{T_1}(t_1) dt_1 \right)}^{\Pr(T_1 < t_2 | T_2 = t_2) \equiv F_{T_1}(t_2)} dt_2 = \int_0^{\infty} \left( \lambda_2 \mathrm{e}^{-\lambda_2 t_2} \right) \left( 1 - \mathrm{e}^{-\lambda_1 t_2} \right) dt_2$$

$$= \int_0^{\infty} \left( \lambda_2 \mathrm{e}^{-\lambda_2 t_2} + \lambda_2 \mathrm{e}^{-(\lambda_1+\lambda_2) t_2} \right) dt_2 = e^{-\lambda_2 t_2} \Big|_{\infty}^{0} + \lambda_2 \frac{e^{-(\lambda_1+\lambda_2)t_2}}{-(\lambda_1+\lambda_2)} \Bigg|_{\infty}^{0} = \frac{\lambda_1}{\lambda_1+\lambda_2}$$

Cf. same proof of slide 80 in SM2_Probabity Theory

- On the other hand

$$\Pr(T_2 < T_1) = 1 - \Pr(T_1 < T_2) = \frac{\lambda_2}{\lambda_1 + \lambda_2}$$

# Splitting of Poisson processes

- In many networking applications it is of interest to study the effect of splitting of packet arrival process.
- Two types of splitting are in general considered:

Random splitting

Regular (or Erlang) splitting

- Let us explain the differences by means of an example:

a) Let us consider the arrival Poisson process of packets $\{X(t), t \in [0, \infty)\}$ with parameter $\lambda$ to a certain switch called Switch X.

b) Some of these packets are then forwarded to Switch A the others to Switch B.

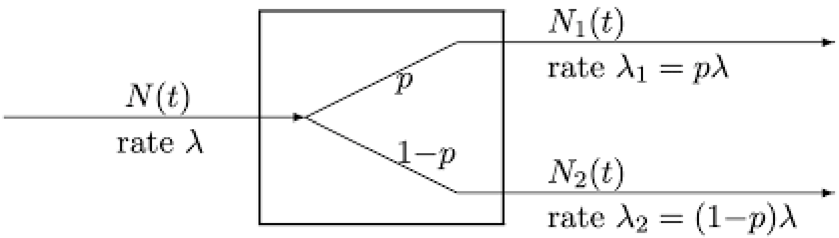
c) The counting processes of packets forwarded to A and B are designated $\{X_A(t), t \in [0, \infty)\}$ and $\{X_B(t), t \in [0, \infty)\}$ , respectively.

# Random splitting of Poisson processes

Under the random splitting, every packet that arrives at Switch X is forwarded **at random** to Switch A with probability $\pi$, or to Switch B with probability $1 - \pi$.

Assumed the forwarding is done instantly upon arrival at Switch X, so

$$X(t) = X_A(t) + X_B(t)$$




then the next fundamental result of "Queuing theory" takes place:

- $\{X_A(t), t \in [0, \infty)\}$ has a Poisson fashion with rate $\pi\lambda$
- $\{X_B(t), t \in [0, \infty)\}$ has Poisson fashion with rate $(1 - \pi)\lambda$.

Note that processes $X(t)$, $X_A(t)$ and $X_B(t)$ counts the number of arrivals in the time interval $[0, t)$ to each of these devices.

**Interpretation:** The rate of process $X(t)$ is scaled by $\pi$, because of only the $\pi\%$ of the overall arrivals are forwarded to Switch A

# Regular splitting of Poisson processes

- Under a regular splitting the 1st packet that arrives at Switch X is forwarded to A; 2nd to B; 3rd to A, etc...

- Let $\{X_{X-A}(t), t \in [0, \infty)\}$ be the process of packet forwarded from Switch X to A, and $\{X_{X-B}(t), t \in [0, \infty)\}$ that from Switch X to B, where

$$X(t) = X_{X-A}(t) + X_{X-B}(t)$$

- Now, $X_{X-A}(t)$ and $X_{X-B}(t)$ are no more Poisson distributed

- Since arrival-times on $X(t)$, namely, $t_1, t_2, \ldots$, are exponential with rate $\lambda$, then time $k$-th arrivals in a row to each s $Erlang(k, \lambda)$.

- Thus, $X_{X-A}(t)$ and $X_{X-B}(t)$ follow a **counting SP** which **inter-arrival times** are **Erlang** with **parameters** $\lambda$, and $k = 2$

$$\Delta t_{X-A},\ \Delta t_{X-B} \sim \mathrm{Erl}(2, \lambda)$$

$$f_{\Delta t_{X-A}}(t) \ = \ f_{\Delta t_{X-B}}(t) \ = \lambda^2 \cdot t \cdot \mathrm{e}^{\lambda t}$$

# Regular splitting of Poisson processes (cont'd)

- Proof: Consider a Poisson counting process $\{X_t, t \in \Re_{\geq 0}\}$.
- Thus we have:

$$\Pr(X_t = i) = \mathrm{e}^{-\lambda t} \cdot \frac{(\lambda t)^i}{i!}$$

- Assume no arrivals to Switch X in the interval $t \in [0, \mathrm{t}]$.
- Consider the r.v. $\Delta t_k = t_k - t_{k-1}$ which represents the inter-event time between the random arrivals at Switch X at times $t_k$ and $t_{k-1}$
- Since the **splitting is regular**, **each Switch (A or B)** have two wait exactly **2 Exponential inter-event times** after **having received the last packet**.

$$\Pr(\Delta t_{X-B} > t) = \Pr(\Delta t_1 + \Delta t_2 > t) = \Pr(X_t \leq 1) = \Pr(X_t = 0) + \Pr(X_t = 1) =$$
$$= \mathrm{e}^{-\lambda t} + (\lambda t)\mathrm{e}^{-\lambda t}$$

$$\Rightarrow \quad F_{\Delta t_1 + \Delta t_2}(t) = \Pr(\Delta t_1 + \Delta t_2 \leq t) = 1 - \mathrm{e}^{-\lambda t} - (\lambda t)\mathrm{e}^{-\lambda t}$$

The sum of 2 Exp. R.V. is $Erlang(2, \lambda)$

$$\Rightarrow \quad f_{\Delta t_1 + \Delta t_2}(t) = \frac{dF_{\Delta t_1 + \Delta t_2}(t)}{dt} = \lambda \mathrm{e}^{-\lambda t} - \lambda \cdot (\mathrm{e}^{-\lambda t} + t(-\lambda \mathrm{e}^{-\lambda t})) = \boxed{\lambda^2 t \mathrm{e}^{-\lambda t}}$$