Mathematics for Informatics 4a

José Figueroa-O'Farrill

Lecture 19

28 March 2012

The story of the film so far...

- We have been studying **stochastic processes**; i.e., systems whose time evolution has an element of chance
- In particular, **Markov processes**, whose future only depends on the present and not on how we got there
- Particularly tractable examples are the **Markov chains**, which are discrete both in "time" and "space":
  - random walks
  - branching processes
- It is interesting to consider also Markov processes $\{X(t) \mid t \geqslant 0\}$ which are continuous in time
- We will only consider those where the $X(t)$ are discrete random variables; e.g., integer-valued
- Main examples will be:
  - Poisson processes
  - Birth and death process, such as queues

Continuous-time Markov processes

- We will be considering **continuous-time** stochastic processes $\{X(t) \mid t \geqslant 0\}$, where each $X(t)$ is a *discrete* random variable taking integer values
- Such a stochastic process has the **Markov property** if for all $i, j, k \in \mathbb{Z}$ and real numbers $0 \leqslant r < s < t$,

$$\mathbb{P}\left(X(t) = j \mid X(s) = i, X(r) = k\right) = \mathbb{P}\left(X(t) = j \mid X(s) = i\right)$$

- If, in addition, for any $s, t \geqslant 0$,

$$\mathbb{P}\left(X(t+s) = j \mid X(s) = i\right)$$

  is *independent* of s, we say $\{X(t) \mid t \geqslant 0\}$ is **(temporally) homogeneous**

Counting processes

- A stochastic process $\{N(t) \mid t \geqslant 0\}$ is a **counting process**, if $N(t)$ represents the total number of events that have occurred up to time t; that is,
  - $N(t) \in \{0, 1, 2, \ldots\}$
  - If $s < t$, $N(s) \leqslant N(t)$
  - For $s < t$, $N(t) - N(s)$ is the number of events which have taken place in $(s, t]$.
- $\{N(t) \mid t \geqslant 0\}$ has **independent increments** if the number of events taking place in disjoint time intervals are independent; that is,
  - the number $N(t)$ of events in $[0, t]$, and
  - the number $N(t+s) - N(t)$ of events in $(t, t+s]$

  are independent
- $\{N(t) \mid t \geqslant 0\}$ is **(temporally) homogeneous** if the distribution of events occurring in any interval of time depends only on the length of the interval: $N(t_2 + s) - N(t_1 + s)$ has the same distribution as $N(t_2) - N(t_1)$, for all $t_1 < t_2$ and $s > 0$

Poisson processes

Definition

$\{N(t) \mid t \geqslant 0\}$ is **Poisson** with rate $\lambda > 0$ if

- $N(0) = 0$
- it has independent increments
- for all $s, t \geqslant 0$ and $n \in \mathbb{N}$,

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

- Since $\mathbb{P}(N(s+t) - N(s) = n)$ does not depend on s, Poisson process are (temporally) *homogeneous*
- Hence, taking $s = 0$, one sees that $N(t)$ is Poisson distributed with mean $\mathbb{E}(N(t)) = \lambda t$

Inter-arrival times

- Let $\{N(t) \mid t \geqslant 0\}$ be a Poisson process with rate $\lambda > 0$
- Let T_1 be the time of the first event
- Let $T_{n>1}$ be the time between the $(n-1)$st and the nth events
- $\{T_1, T_2, \dots\}$ is the sequence of **inter-arrival times**
- $\mathbb{P}(T_1 > t) = \mathbb{P}(N(t) = 0) = e^{-\lambda t}$, whence T_1 is exponentially distributed with mean $\frac{1}{\lambda}$
- The same is true for the other T_n, e.g.,

$$\begin{aligned}
& \mathbb{P}(T_2 > t \mid T_1 = s) = \mathbb{P}(0 \text{ events in } (s, s+t] \mid T_1 = s) \\
\text{(indep. incr.)} \quad & = \mathbb{P}(0 \text{ events in } (s, s+t]) \\
\text{(homogeneity)} \quad & = \mathbb{P}(0 \text{ events in } (0, t]) \\
& = e^{-\lambda t}
\end{aligned}$$

- The $\{T_n\}$ for $n \geqslant 1$ are i.i.d. exponential random variables with mean $\frac{1}{\lambda}$

Waiting times

- The **waiting time** until the nth event ($n \geqslant 1$) is $S_n = \sum_{i=1}^n T_i$
- $S_n \leqslant t$ if and only if $N(t) \geqslant n$, whence

$$\mathbb{P}(S_n \leqslant t) = \mathbb{P}(N(t) \geqslant n) = \sum_{j=n}^{\infty} \mathbb{P}(N(t) = j) = \sum_{j=n}^{\infty} e^{-\lambda t}\frac{(\lambda t)^j}{j!}$$

- Differentiating with respect to t, we get the probability density function

$$f_{S_n}(t) = \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}$$

which is a "gamma" distribution
- $\mathbb{E}(S_n) = \sum_i \mathbb{E}(T_i) = \frac{n}{\lambda}$

Example

Suppose that trains arrive at a station at a Poisson rate $\lambda = 1$ per hour.

1. What is the expected time until the 6th train arrives?
2. What is the probability that the elapsed time between the 6th and 7th trains exceeds 1 hour?

1. We are asked for $\mathbb{E}(S_6)$ where S_6 is the waiting time until the 6th train, hence $\mathbb{E}(S_6) = \frac{6}{\lambda} = 6$ hours.
2. We are asked for $\mathbb{P}(T_7 > 1) = e^{-\lambda} = e^{-1} \simeq 37\%$

Time of occurrence is uniformly distributed

- If we know that exactly one event has occurred by time t, how is the time of occurrence distributed?
- For $s \leqslant t$,

$$
\begin{aligned}
\mathbb{P}(T_1 < s \mid N(t) = 1) &= \frac{\mathbb{P}(T_1 < s \text{ and } N(t) = 1)}{\mathbb{P}(N(t) = 1)} \\
&= \frac{\mathbb{P}(N(s) = 1 \text{ and } N(t) - N(s) = 0)}{\mathbb{P}(N(t) = 1)} \\
\text{(indep. incr.)} \qquad &= \frac{\mathbb{P}(N(s) = 1)\mathbb{P}(N(t) - N(s) = 0)}{\mathbb{P}(N(t) = 1)} \\
\text{(homogeneity)} \qquad &= \frac{\lambda s e^{-\lambda s} e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}} \\
&= \frac{s}{t}
\end{aligned}
$$

- i.e., it is uniformly distributed

Example (Stopping game)

Events occur according to a Poisson process $\{N(t) \mid t \geqslant 0\}$ with rate λ. Each time an event occurs we must decide whether or not to stop, with our objective being to stop at the last event to occur prior to some specified time τ. That is, if an event occurs at time t, $0 < t < \tau$ and we decide to stop, then we lose if there are any events in the interval $(t, \tau]$, and win otherwise. If we do not stop when an event occurs, and no additional events occur by time τ, then we also lose. Consider the strategy that stops at the first event that occurs after some specified time s, $0 < s < \tau$. *What should s be to maximise the probability of winning?*
We win if and only if there is precisely one event in $(s, \tau]$, hence

$$\mathbb{P}(☺) = \mathbb{P}(N(\tau) - N(s) = 1) = e^{-\lambda(\tau - s)}\lambda(\tau - s)$$

We differentiate with respect to s to find that the maximum occurs for $s = \tau - \frac{1}{\lambda}$, for which $\mathbb{P}(☺) = e^{-1} \simeq 37\%$

Example (Bus or walk?)

Buses arrive to a stop according to a Poisson process $\{N(t) \mid t \geqslant 0\}$ with rate λ. Your bus trip home takes b minutes from the time you get on the bus until you reach home. You can also walk home from the bus stop, the trip taking you w minutes. You adopt the following strategy: upon arriving at the bus stop, you wait for at most s minutes and start walking if no bus has arrived by that time. (Otherwise you catch the first bus that comes.) *Is there an optimal s which minimises the duration of your trip home?*
Let T denote the duration of the trip home from the time you arrive at the bus stop. Conditioning on the mode of transport,

$$
\begin{aligned}
\mathbb{E}(T) &= \mathbb{E}(T \mid N(s) = 0)\mathbb{P}(N(s) = 0) + \mathbb{E}(T \mid N(s) > 0)\mathbb{P}(N(s) > 0) \\
&= (s + w)e^{-\lambda s} + (\mathbb{E}(T_1) + b)(1 - e^{-\lambda s})
\end{aligned}
$$

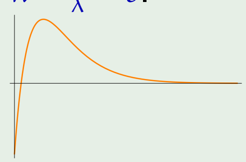
where T_1 is the first arrival time. Recall $\mathbb{E}(T_1) = \frac{1}{\lambda}$.

Example (Bus or walk? — continued)

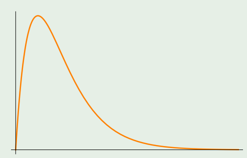
Therefore,

$$
\begin{aligned}
\mathbb{E}(T) &= (s + w)e^{-\lambda s} + (\tfrac{1}{\lambda} + b)(1 - e^{-\lambda s}) \\
&= \tfrac{1}{\lambda} + b + \left(s + w - \tfrac{1}{\lambda} - b\right) e^{-\lambda s}
\end{aligned}
$$

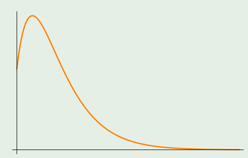
whose behaviour (as a function of s) depends on the sign of $w - \frac{1}{\lambda} - b$.



$w - \frac{1}{\lambda} - b < 0$



$w - \frac{1}{\lambda} - b = 0$



$w - \frac{1}{\lambda} - b > 0$

The minimum is either at $s = 0$ (walk without waiting) or at $s = \infty$ (wait for the bus no matter how long)

Poisson processes as Markov processes I

- A Poisson process $\{N(t) \mid t \geqslant 0\}$ is an example of a continuous-time homogeneous Markov process
- The states are the natural numbers $\mathbb{N} = \{0, 1, 2, \dots\}$
- The possible transitions are between states n and $n+1$
- The transition probabilities are for $0 \leqslant s < t$,

$$\begin{aligned}
\mathbb{P}(N(t) = n+1 \mid N(s) = n) &= \frac{\mathbb{P}(N(t) = n+1 \text{ and } N(s) = n)}{\mathbb{P}(N(s) = n)} \\
&= \frac{\mathbb{P}(N(s) = n \text{ and } N(t) - N(s) = 1)}{\mathbb{P}(N(s) = n)} \\
\text{(indep. incr.)} \qquad &= \frac{\mathbb{P}(N(s) = n)\mathbb{P}(N(t) - N(s) = 1)}{\mathbb{P}(N(s) = n)} \\
\text{(homogeneity)} \qquad &= \mathbb{P}(N(t-s) = 1) \\
&= \lambda(t-s)e^{-\lambda(t-s)}
\end{aligned}$$

Poisson processes as Markov processes II

- The inter-arrival time T_n is the time the system spends in state $(n-1)$ before making a transition to n
- They are exponentially distributed with mean $\frac{1}{\lambda}$
- Now let us consider a *general* continuous-time Markov process, not necessarily Poisson
- Let T_n denote the time the system spends in state n before making a transition to a different state
- The Markov property says that for $s, t \geqslant 0$,

$$\mathbb{P}(T_n > s + t \mid T_n > s) = \mathbb{P}(T_n > t)$$

  i.e., $\mathbb{P}(T_n > s + t \mid T_n > s)$ cannot depend on s, since how long the system has been in state n cannot matter
- Thus T_n is *memoryless* and this means that T_n is exponentially distributed

Continuous and memoryless = exponential (I)

- Let X be a *continuous* random variable taking non-negative values
- We say that X is **memoryless** if for all $s, t \geqslant 0$,

$$\mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t)$$

- This is equivalent to

$$\mathbb{P}(X > t) = \frac{\mathbb{P}(X > s + t \text{ and } X > s)}{\mathbb{P}(X > s)} = \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)}$$

  or to $\mathbb{P}(X > s + t) = \mathbb{P}(X > t)\mathbb{P}(X > s)$
- Letting $F(s) = \mathbb{P}(X > s)$, this is equivalent to the functional equation

$$F(s + t) = F(s)F(t)$$

Continuous and memoryless = exponential (II)

Theorem

The only (nontrivial) continuous functions $F(s)$ obeying

$$F(s + t) = F(s)F(t) \qquad \forall\, s, t \geqslant 0$$

are $F(s) = e^{-s\lambda}$, for some $\lambda > 0$.

Proof.

- $F(0) = F(0)^2$ implies that either $F(0) = 0$ or $F(0) = 1$
- $F(s) = F(s)F(0)$ implies $F(0) = 1$ since $F(s)$ is nontrivial
- For all $n \in \mathbb{N}$, $F(n+1) = F(n)F(1)$, whence $F(n) = F(1)^n$
- Since $0 < F(1) < 1$, we can write $F(1) = e^{-\lambda}$ for some $\lambda > 0$
- $F(\frac{k}{n})^n = F(k) = F(1)^k = e^{-k\lambda}$, whence $F(\frac{k}{n}) = e^{-k\lambda/n}$
- Two continuous functions agreeing on the (non-negative) rationals agree on the (non-negative) reals, so $F(s) = e^{-s\lambda}$

$\square$

Warning

Continuity is essential! There are discrete distributions (e.g., geometric) which are also memoryless. Of course, exponential distribution is the "continuous" limit of the geometric distribution.

We conclude with the observation that a continuous-time Markov chain is a stochastic process such that each time the system enters state i

1. the amount of time it spends in that state before making a transition into a different state is exponentially distributed with mean, say, $\frac{1}{\lambda_i}$, and
2. when the process leaves state i is enters states j with some probability p_{ij}, where the p_{ij} satisfy
   1. $p_{ii} = 0$ for all i, and
   2. $\sum_j p_{ij} = 1$ for all i.

In a Poisson process all the λ_i are equal.

Summary

- We have introduced **continuous-time Markov processes** $\{X(t) \mid t \geqslant 0\}$, satisfying the **Markov property** that for $0 \leqslant r < s < t$,

$$\mathbb{P}(X(t) = j \mid X(s) = i, X(r) = k) = \mathbb{P}(X(t) = j \mid X(s) = i)$$

- We focussed on **homogeneous** processes, for which $\mathbb{P}(X(t+s) = j \mid X(s) = i)$ does not depend on s
- Examples are the **counting processes** $\{N(t) \mid t \geqslant 0\}$, of which an important special case are the **Poisson processes**, where $N(t)$ is Poisson distributed with mean λt
- **Inter-arrival times** in a Poisson process are exponential, **waiting times** are "gamma" distributed and **time of occurrence** is uniformly distributed
- Continuous-time Markov chains are defined by
  1. a transition matrix $[p_{ij}]$ (as in the discrete case)
  2. the exponential transition rates λ_i