Sunday 27th of October 8:30 to 12:00

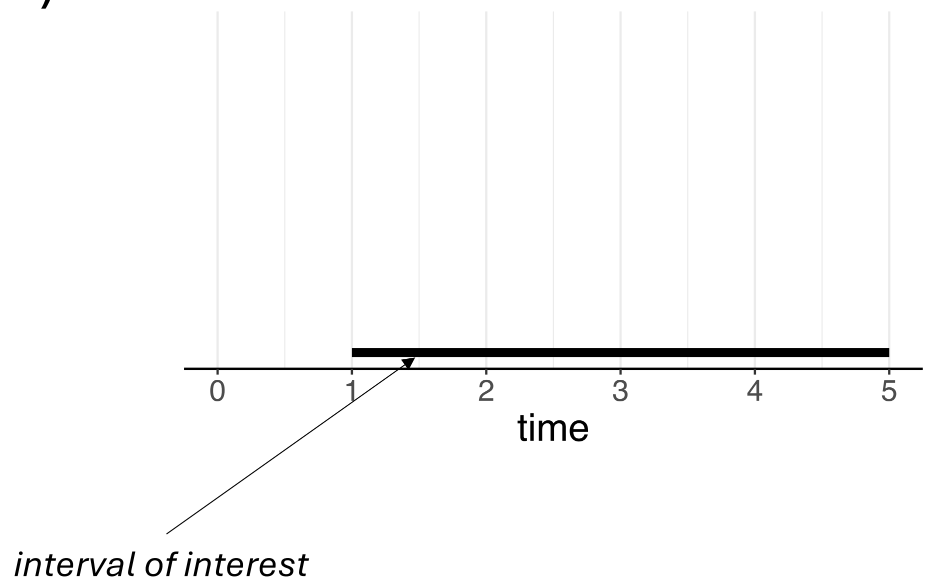
Time	Description	Discussant
[15 min]	(0) Introductions and administrivia	Trikalinos
[25 min]	(1) DES as a composition of point processes	Alarid-Escudero
[30 min]	(2) NHPPPs – key properties	Trikalinos
[30 min]	(3) Sampling from NHPPPs	Sereda
[15 min]	**Break**	
[80 min]	(4) Guided exercise: - Implement a simple cancer natural history DES for one person - The many-person case - Packaging	[All] Chrysanthopoulou Sereda/Alarid-Escudero Trikalinos
[10 min]	(5) Advanced Topic Teaser on self-excitatory processes: point processes that are not NHPPPs and when you may need them	Trikalinos
[15 min]	General Q & A	All

Section 2: Theory

The building block

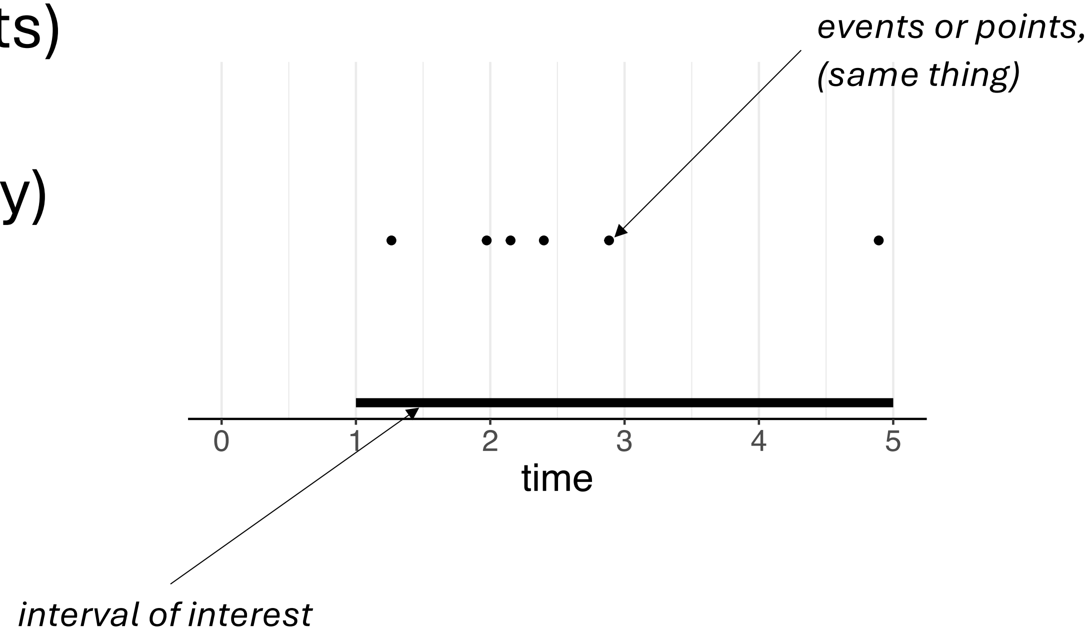
The point process

- A scheme that generates a sequence of events (points) over a time interval



The point process

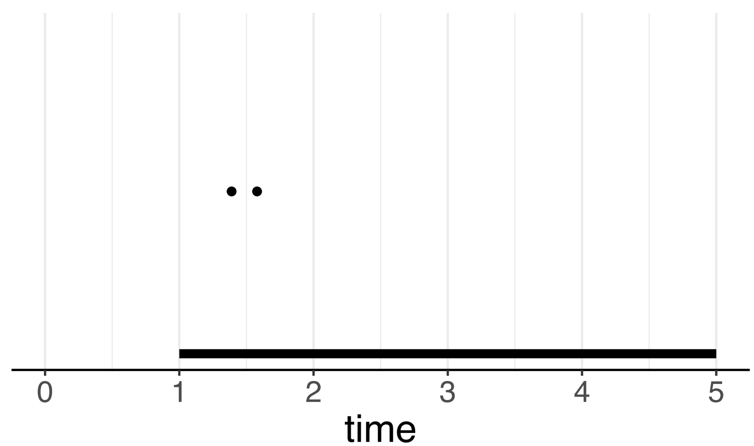
- A scheme that generates a sequence of events (points) over time
- An instantiation (trajectory) of the process is a sequence of 0, 1 or more events in the interval, but none outside it



The point process

- A scheme that generates a sequence of events (points) over time
- Each instantiation is random

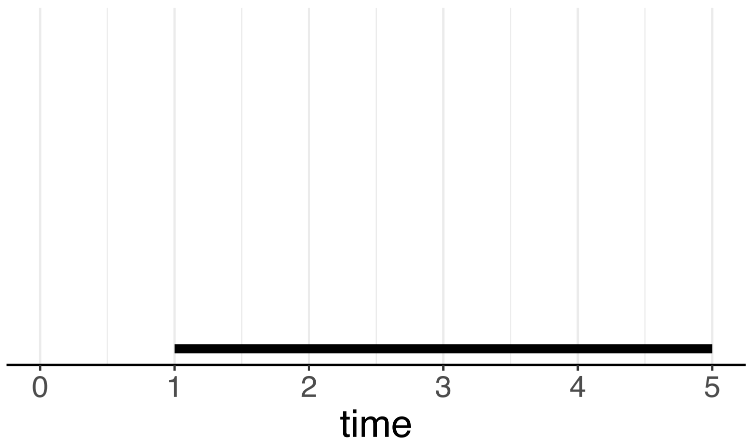
another instantiation (trajectory)



The point process

- A scheme that generates a sequence of events (points) over time
- Each instantiation is random

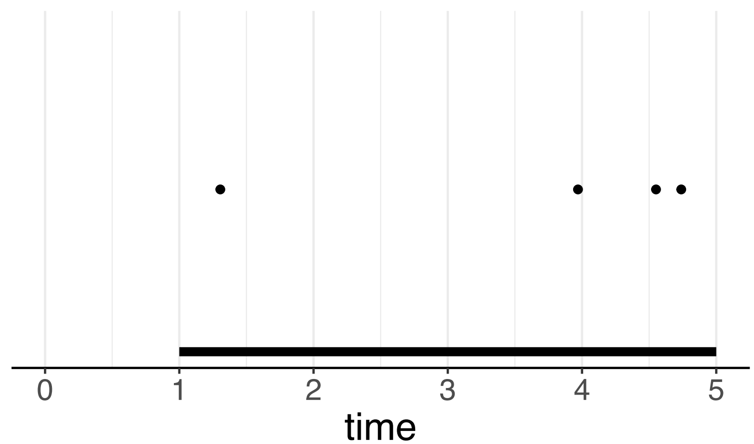
another instantiation (with no events)



The point process

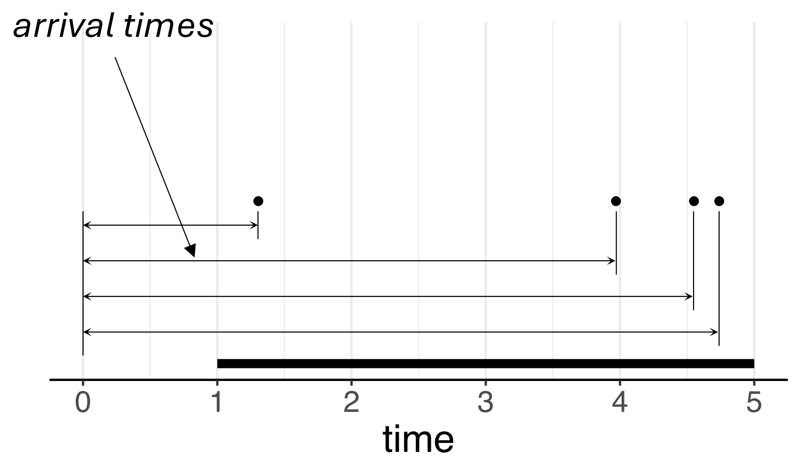
- A scheme that generates a sequence of events (points) over time
- Each instantiation is random

yet another instantiation



The point process

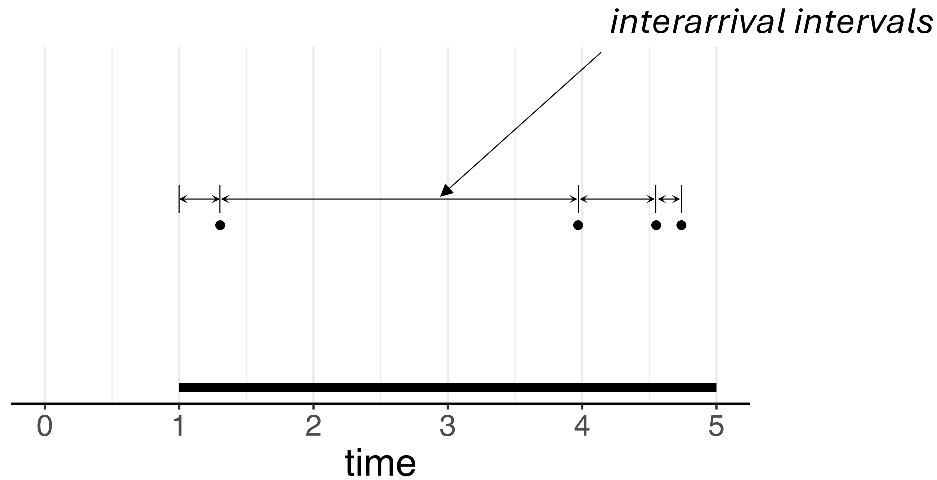
- The *arrival times* (times of the events) are random
- They start from whenever we zeroed the clock



The point process

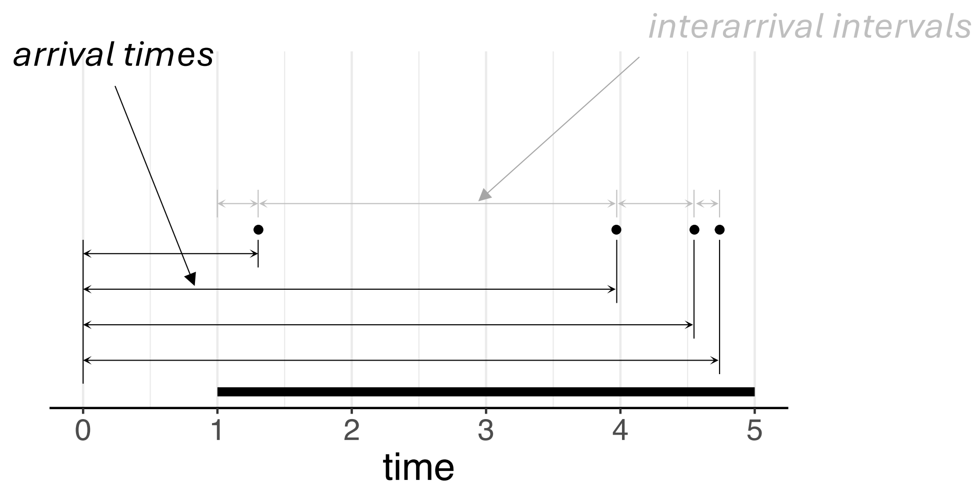
- The interarrival times are the lengths of the interarrival time intervals
- The arrival times and interarrival times give the same information

  (… thus, the interarrival times are random)



The point process

Hereon, we refer only to arrival times

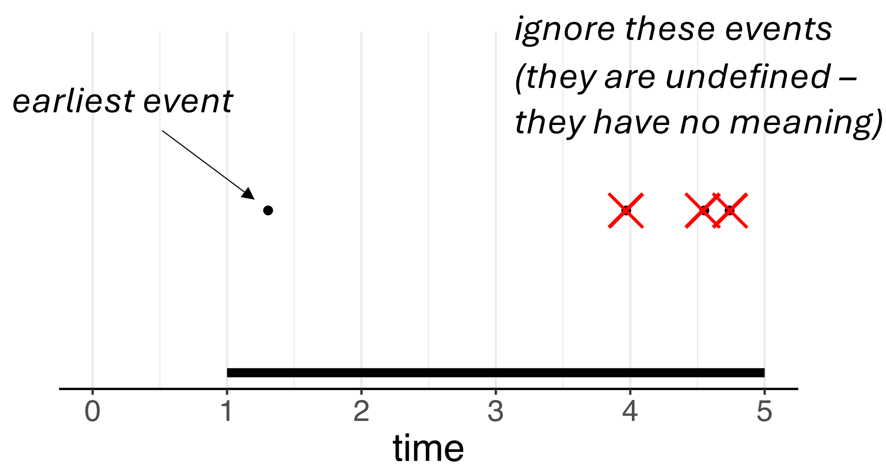


Modeling non-repeatable events

If the point process models a *nonrepeatable* event, we care only about the **earliest event.**

Will it occur in the interval, and, and if so, when?

Example: model a cause of death

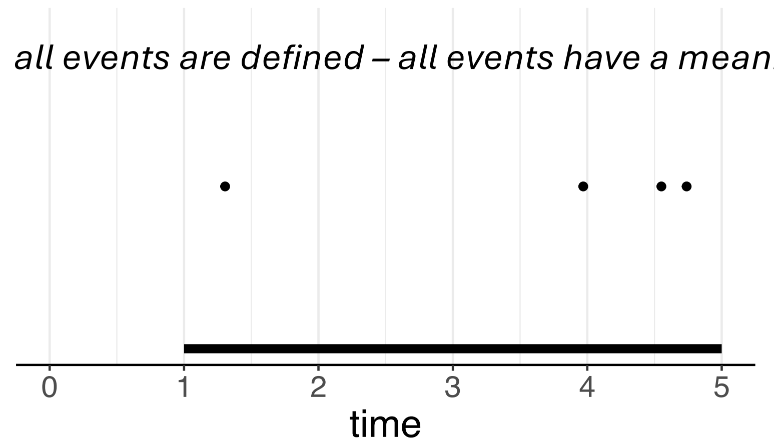


Modeling repeatable events

If the point process models a *repeatable* event, we care are about **all events**.

Will any occur in the interval, and, and if so, when?

Example: model the emergence of tumors, or the start of symptomatic episodes



The Poisson point process

- There are many types of point processes
- We will consider only a one type – the Poisson point process

The Poisson point process

Number of events between t and t + Δt

If for a sequence of events

$o(\Delta t)$ becomes 0 ***very fast***

$$
\begin{aligned}
\Pr[\boxed{N(t, t+\Delta t)} = 0] &= 1 - \lambda \Delta t + \cancel{o(\Delta t)}, \\
\Pr[N(t, t+\Delta t) = 1] &= \lambda \Delta t + \cancel{o(\Delta t)}, \\
\Pr[N(t, t+\Delta t) > 1] &= \cancel{o(\Delta t)}, \text{ and} \\
N(t, t+\Delta t) &\perp\!\!\!\perp N(0, t),
\end{aligned}
$$

for some $\lambda > 0$ and as $\Delta t \to 0$, then that sequence is a Poisson point process

The Poisson point process (in English)

If for a sequence of events

$$\Pr[N(t, t + \Delta t) = 0] = 1 - \lambda \Delta t$$
$$\boxed{\Pr[N(t, t + \Delta t) = 1] = \lambda \Delta t}$$
$$\Pr[N(t, t + \Delta t) > 1] = 0 \quad \text{and}$$
$$N(t, t + \Delta t) \perp\!\!\!\perp N(0, t),$$

for some $\lambda > 0$ and as $\Delta t \to 0$, then that sequence is a Poisson point process

Over a vanishingly small interval

- *you may get 1 event with probability $\lambda \Delta t$ …*

The Poisson point process (in English)

If for a sequence of events

$$\boxed{\Pr[N(t, t+\Delta t) = 0] = 1 - \lambda \Delta t}$$

$$\Pr[N(t, t+\Delta t) = 1] = \lambda \Delta t$$

$$\Pr[N(t, t+\Delta t) > 1] = 0 \qquad \text{and}$$

$$N(t, t+\Delta t) \perp\!\!\!\perp N(0, t),$$

for some $\lambda > 0$ and as $\Delta t \to 0$, then that sequence is a Poisson point process

Over a vanishingly small interval

- *you may get 1 event with probability $\lambda \Delta t$ …*
- *otherwise, you'll get 0 events;*

The Poisson point process (in English)

If for a sequence of events

$$\Pr[N(t, t + \Delta t) = 0] = 1 - \lambda \Delta t$$
$$\Pr[N(t, t + \Delta t) = 1] = \lambda \Delta t$$
$$\boxed{\Pr[N(t, t + \Delta t) > 1] = 0 \quad \text{and}}$$
$$N(t, t + \Delta t) \perp\!\!\!\perp N(0, t),$$

for some $\lambda > 0$ and as $\Delta t \to 0$, then that sequence is a Poisson point process

Over a vanishingly small interval

- *you may get 1 event with probability $\lambda \Delta t$ …*
- *otherwise, you'll get 0 events;*
- *you'll never get many concurrent events*

The Poisson point process (in English)

If for a sequence of events

$$\Pr[N(t, t+\Delta t) = 0] = 1 - \lambda \Delta t$$
$$\Pr[N(t, t+\Delta t) = 1] = \lambda \Delta t$$
$$\Pr[N(t, t+\Delta t) > 1] = 0 \qquad \text{and}$$
$$\boxed{N(t, t+\Delta t) \perp\!\!\!\perp N(0, t),}$$

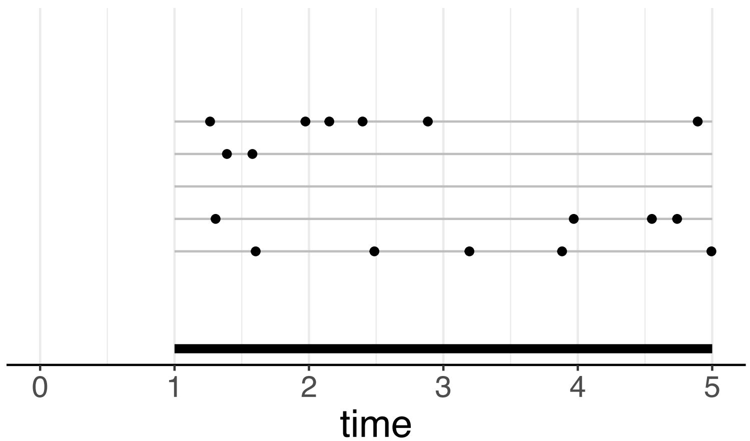
for some $\lambda > 0$ and as $\Delta t \to 0$, then that sequence is a Poisson point process

Over a vanishingly small interval

- *you may get 1 event with probability $\lambda\Delta t$ …*
- *otherwise, you'll get 0 events;*
- *you'll never get many concurrent events*
- *and it does not matter what happened in the past*

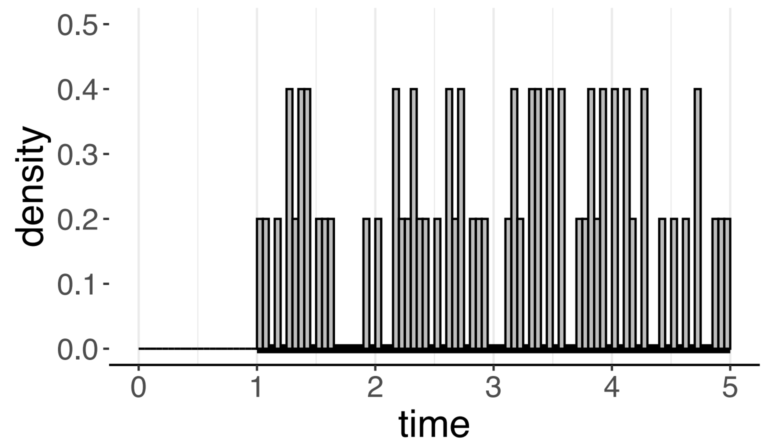
The intensity function λ in the example

Event times for five instantiations



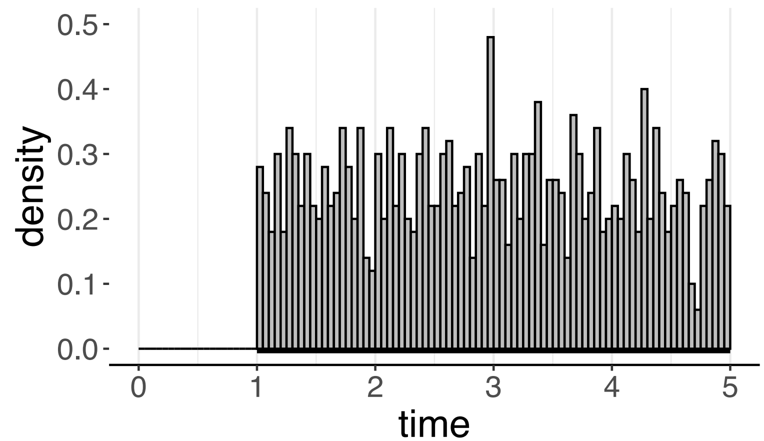
The intensity function λ in the example

A histogram of the event times
for 100 instantiations



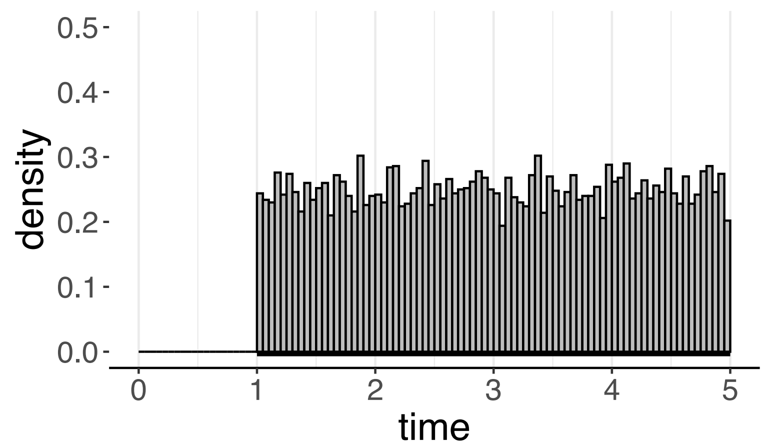
The intensity function λ in the example

… for 1000 instantiations



The intensity function λ in the example

... and for 10000 instantiations

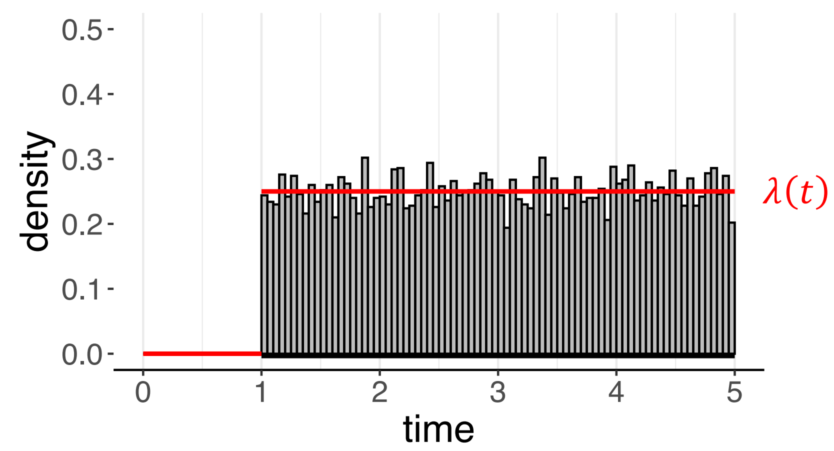


The intensity function λ in the example

As the number of instantiations goes to infinity, the histogram approaches the shape of the intensity function $\lambda(t)$.

The intensity function governs event occurrence.

(It is the same quantity as the hazard function in survival analysis)



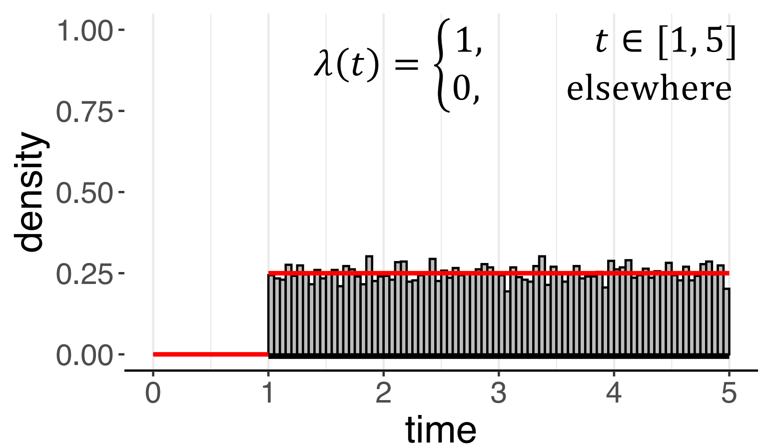
The intensity function is scaled by the expected number of events in the interval to be on the same plot

Time-homogeneous and non-homogeneous

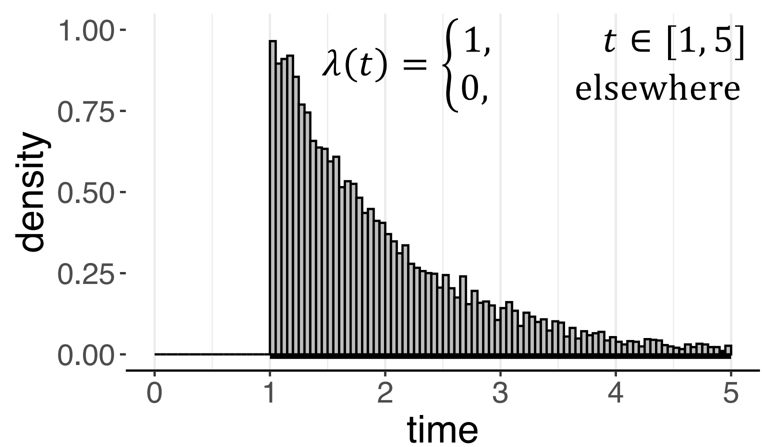
- $\lambda(\mathrm{t}) = \mathrm{constant}$: the Poisson point process (PPP) is called time-homogeneous
- Otherwise, it is called a non-homogeneous PPP (NHPPP)

All events vs earliest event in the example

All events, 10K instantiations



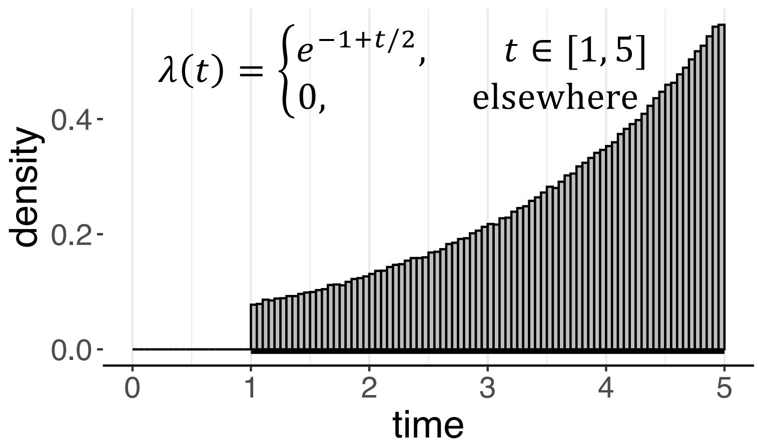
Earliest event, 10K instantiations



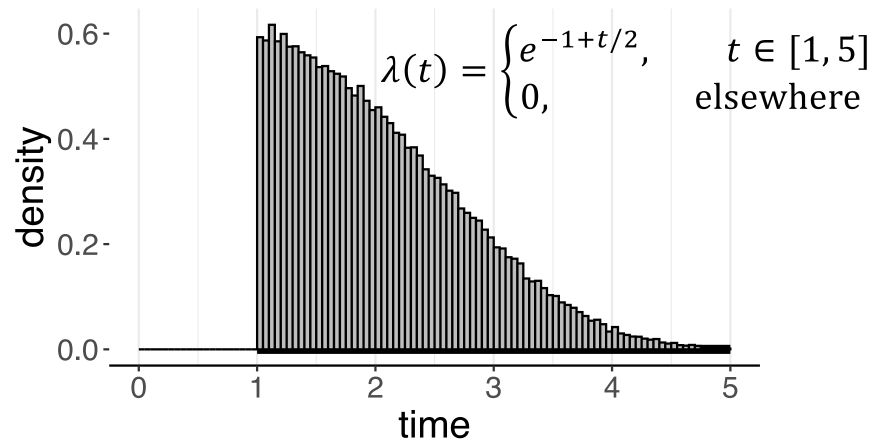
The histogram of the earliest event times does not approach the shape of the intensity function

All events vs earliest event, different example

All events, 100K instantiations



Earliest event, 100K instantiations



The three important functions

- Intensity function $\lambda(t)$

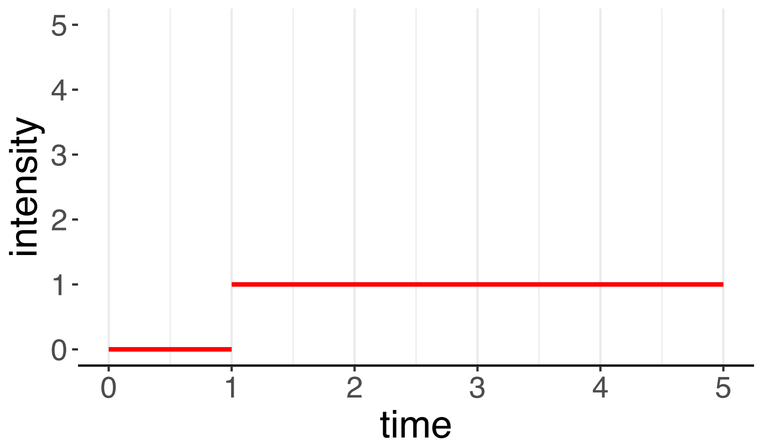
- *Always available*
- *Sufficient to sample from any NHPPP efficiently and accurately*

- Cumulative intensity function $\Lambda(t) = \int_0^t \lambda(s)\,\mathrm{d}s$
- Inverse cumulative intensity function $\Lambda^{-1}(z)$, defined so that $\Lambda^{-1}(\Lambda(t)) = t$

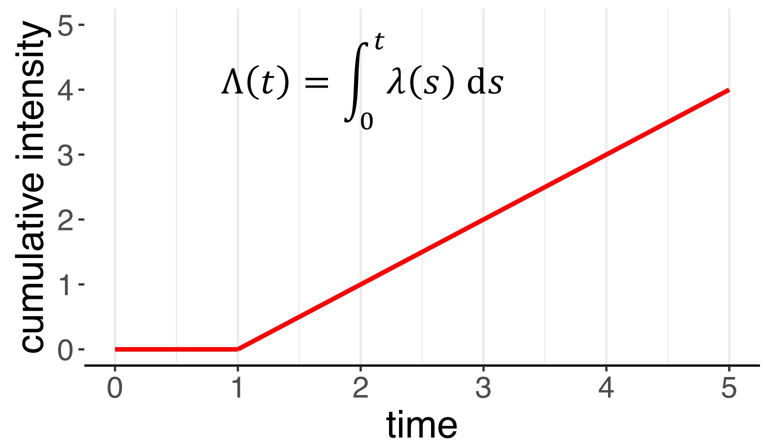
- *Not always available*
- *If available, you accelerate sampling by several times*

Intensity and cumulative intensity functions

Intensity function $\lambda(t)$

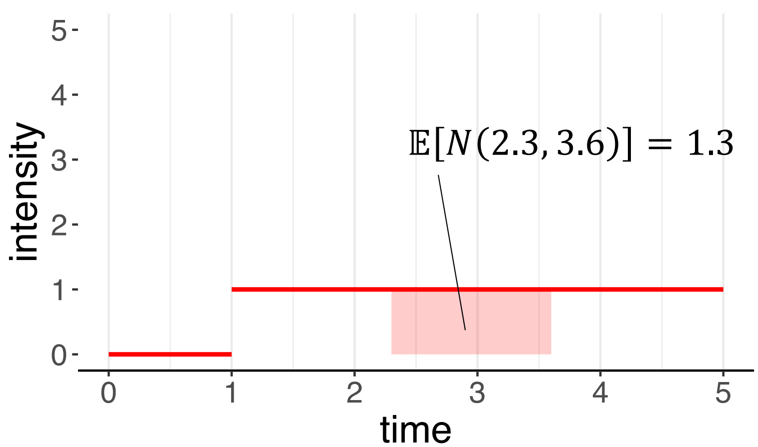


Cumulative intensity function $\Lambda(\mathrm{t})$

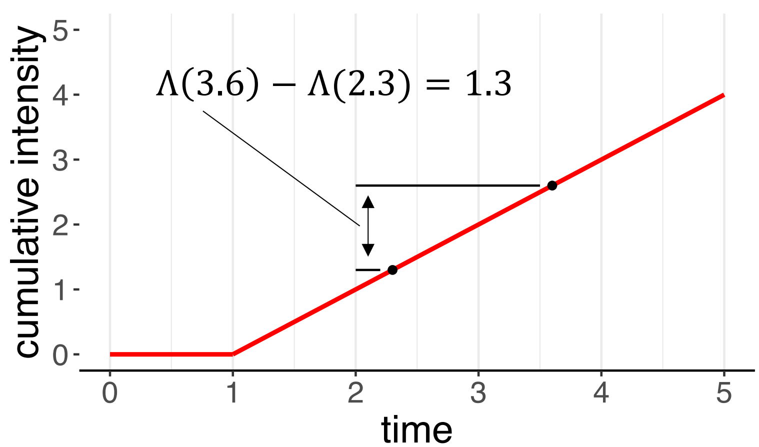


Intensity and cumulative intensity functions

Intensity function $\lambda(t)$

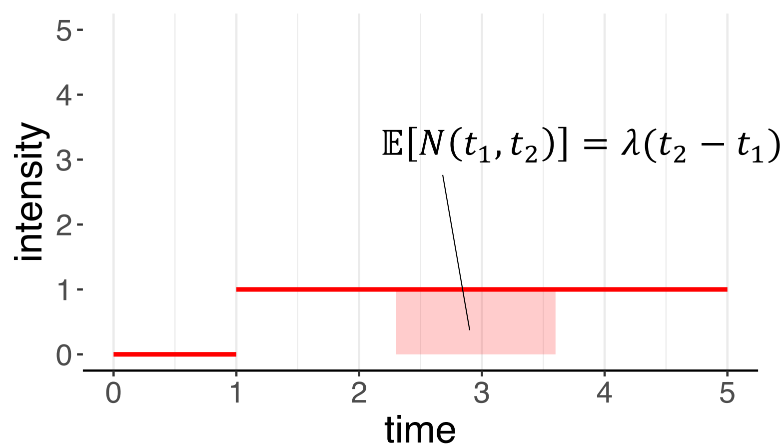


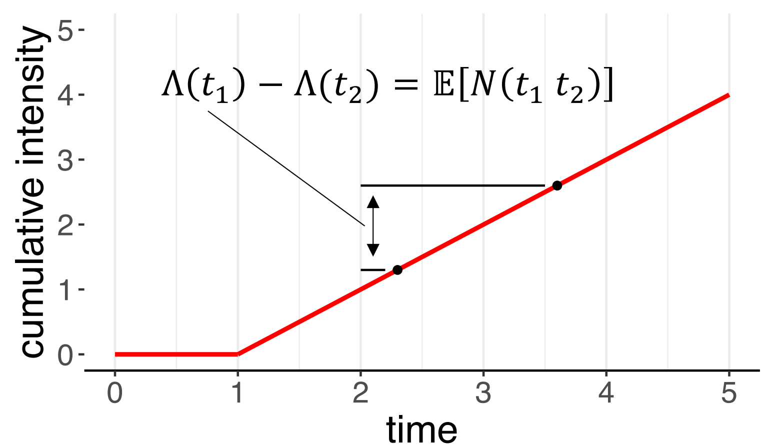
Cumulative intensity function $\Lambda(\mathrm{t})$



Intensity and cumulative intensity functions

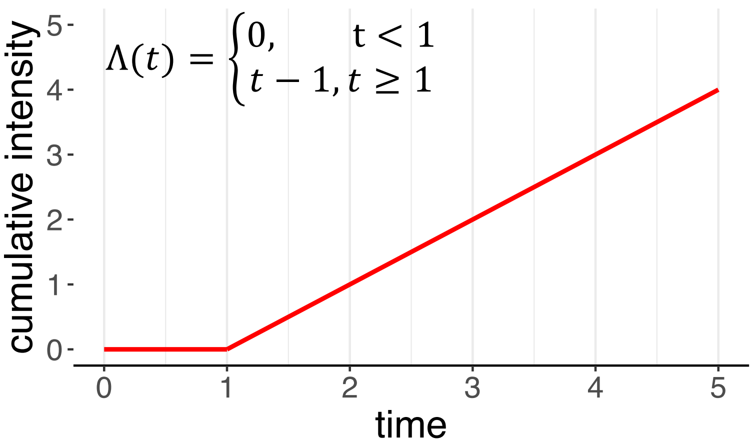
$$N(t_1, t_2) \sim \text{Poisson}(\Lambda(t_2) - \Lambda(t_1)),$$
irrespective of the form of $\lambda(t)$



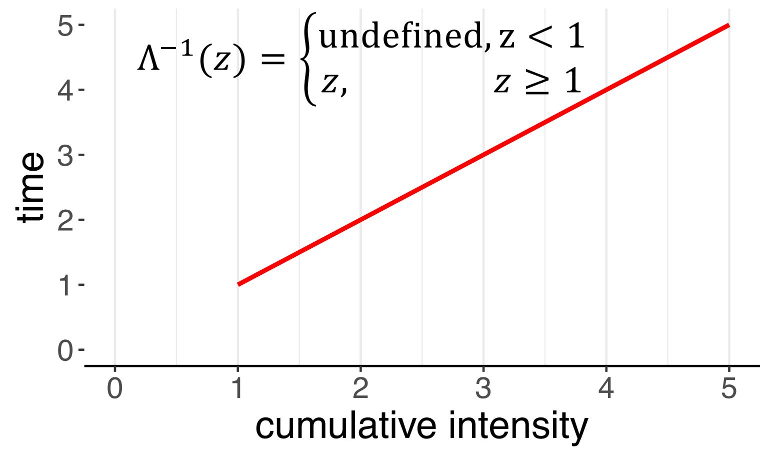


Cumulative intensity function and its inverse

Cumulative intensity function $\Lambda(\mathrm{t})$



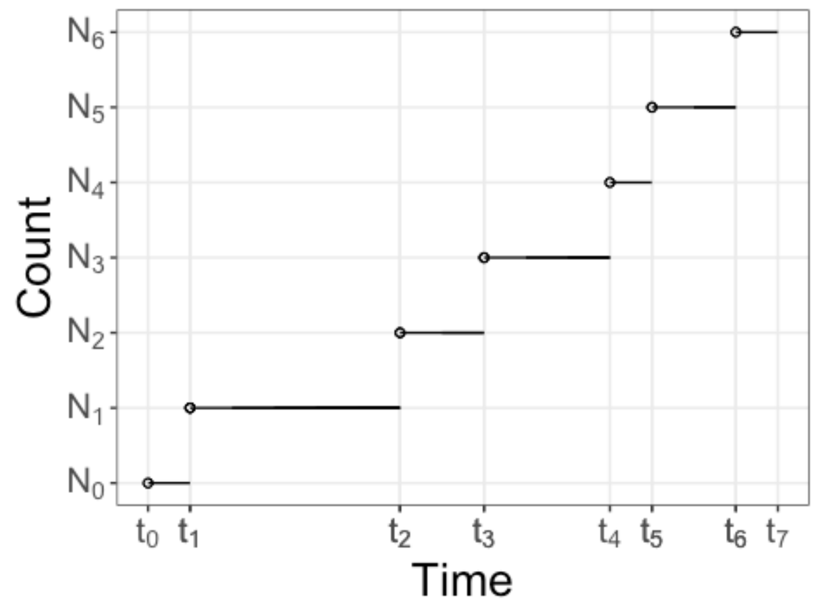
Inverse cumulative intensity function $\Lambda^{-1}(\mathrm{z})$



Duality with the Poisson counting process

Poisson *counting process*

$N_0, N_1, \ldots$
Cumulative number events over time



Poisson *point process*

$t_0, t_1, \ldots$

How does this connect to survival analysis?

- The theory of point processes is the foundation of survival analyses
- Often survival analysis is about the time to the earliest event
- The intensity function is the same hazard function
- The cumulative intensity function is the same as the integrated or cumulative hazard function
- All the tools that we describe here can be used for statistical simulations for survival analysis

Next … Section 3: Sampling

Sunday 27th of October 8:30 to 12:00

Time	Description	Discussant
[15 min]	(0) Introductions and administrivia	Trikalinos
[25 min]	(1) DES as a composition of point processes	Alarid-Escudero
[30 min]	(2) NHPPPs – key properties	Trikalinos
[30 min]	(3) Sampling from NHPPPs	Sereda
[15 min]	**Break**	
[80 min]	(4) Guided exercise: - Implement a simple cancer natural history DES for one person - The many-person case - Packaging	[All] Chrysanthopoulou Sereda/Alarid-Escudero Trikalinos
[10 min]	(5) Advanced Topic Teaser on self-excitatory processes: point processes that are not NHPPPs and when you may need them	Trikalinos
[15 min]	General Q & A	All