

nhppp: Simulating Nonhomogeneous Poisson Point Processes in R

Thomas A. Trikalinos 
Brown University

Yuliia Sereda 
Brown University

Abstract

We introduce the **nhppp** package for simulating events from one dimensional non-homogeneous Poisson point processes (NHPPs) in R. Its functions are based on three algorithms that provably sample from a target NHPP: the time-transformation of a homogeneous Poisson process (of intensity one) via the inverse of the integrated intensity function; the generation of a Poisson number of order statistics from a fixed density function; and the thinning of a majorizing NHPP via an acceptance-rejection scheme. We present a study of numerical accuracy and time performance of the algorithms and advice on which algorithm to prefer in each situation. Functions available in the package are illustrated with simple reproducible examples.

Keywords: stochastic point processes, counting processes, discrete event simulation, time-to-event simulation, R.

1. Introduction

It is often desirable to simulate series of events (stochastic point processes) so that the intensity of their occurrence varies over time. Examples include events such as the occurrence of death or occurrences of symptoms, infections, or tumors over a person's lifetime. The nonhomogeneous Poisson point process (NHPP), which generalizes the simpler homogeneous-Poisson, Weibull, and Gompertz point processes, is a widely used model for such series of events. NHPPs can model complicated event patterns given a suitable intensity function. They are therefore a very useful tool in statistical and mathematical model simulation.

A NHPP has the properties that the number of events in all non-overlapping time intervals are independent random variables and that, within each time interval, the number of events is Poisson distributed. Thus an NHPP is a memoryless point process. A large number of phenomena may reasonably conform with these properties.¹

The **nhppp** package in R contains functions for the simulation of NHPPs over a one-dimensional carrier space, which we will take to represent time.

We review NHPPs in Section 2 and algorithms for sampling from constant rate Poisson point processes in Section 3. We introduce the three sampling algorithms that are implemented

¹Other phenomena are not well described by a memoryless point process. An example is new infection events during an epidemic wave, where the number of new infections over time depends on the number of infections in the previous time interval. Thus, at least in its early phase, an epidemic wave may be better described by, e.g., self-exciting point processes that transiently increase the intensity function after a first event.

in the package in Section 4. We discuss special functional forms for the intensity function (constant, piecewise constant, linear, and log-linear) in Section 5. We describe **nhppp** versus other R packages that can simulate from one dimensional NHPPPs in Section 6 and present a numerical study in Section 7. We summarize in Section 8.

2. The Poisson point process

2.1. Definition

The Poisson point process is a stochastic series of events on the real line. For some sequence of events, let $N(t, t + \Delta t)$ be the number of events in the interval $(t, t + \Delta t]$. If for some positive intensity λ and, as $\Delta t \rightarrow 0$,

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

then that sequence of events is a Poisson point process. In Equation (1), the third statement demands that events occur one at a time. The fourth statement implies that the process is memoryless: For any time t_0 , the behavior of the process is independent to what happened before that time.

2.2. Homogeneous Poisson point process and counting process

Assume that the next event after time t_0 happens at time $t_0 + X$. It follows from the above definition (see Cox and Miller (1965, par. 4.1)) that, for a constant λ , X is exponentially distributed

$$X \sim \text{Exponential}(\lambda),\tag{2}$$

and that, over the compact interval $(a, b]$ the number of events is Poisson distributed

$$N(a, b) \sim \text{Poisson}(\lambda(b - a)).\tag{3}$$

Equation (2) generates the homogeneous Poisson point process $Z_1 = t_0 + X_1, Z_2 = Z_1 + X_2, \dots$, where Z_i is the time of arrival of event i and the inter-arrival times X_i follow (2). We will use $Z_{(i)}$ to denote the event i when event are ordered in increasing time. Equation (3) describes the corresponding (dual) counting process $N_1 = N(t_0, Z_1), N_2 = N(t_0, Z_2), \dots$, where N_i is the total number of events from time t_0 to time Z_i . The point process (the sequence $[Z_i]$ of event times) and the counting process (the sequence $[N_i]$ of cumulants) are two sides of the same coin.

Sampling from the constant rate point process in (2) is discussed in Section 3.

2.3. Non homogeneous Poisson point process and counting process

When the intensity function changes over time, the homogeneous Poisson point process generalizes to its nonstationary counterpart, an NHPPP, with intensity function $\lambda(t) > 0$. For

details see [Cox and Miller \(1965, par 4.2\)](#). Then the number of events over the interval $(a, b]$ becomes

$$N(a, b) \sim \text{Poisson}(\Lambda(a, b)), \quad (4)$$

where $\Lambda(a, b) = \int_a^b \lambda(t) dt$ is the integrated intensity or cumulative intensity of the NHPPP. Equation (4) describes the counting process of the NHPPP, which in turn implies a stochastic point process – a distribution of events over time.

Here the simulation task is to sample event times from the point process that corresponds to intensity function $\lambda(t)$, or equivalently, to the integrated intensity function $\Lambda(t) = \int_0^t \lambda(s) ds$ (Section 4).²

A note on zero intensity processes

In (1) λ is strictly positive but in **nhppp** we allow it to be nonnegative. If $\lambda = 0$, $\Pr[N(t, t + \Delta t) = 0] = 1$ and $\Pr[N(t, t + \Delta t) \geq 1] = 0$. This means that no events occur and the stochastic point process in the interval $(t, t + \Delta t]$ is denegate. Allowing $\lambda(t) \geq 0$ has no bearing on the results of simulations. If

$$\lambda(t) \begin{cases} > 0, & \text{for } t \in (a, b] \\ = 0, & \text{for } t \in (b, c] \\ > 0, & \text{for } t \in (c, d] \end{cases}$$

we can always ignore the middle interval in which no events happen.

2.4. Properties that are important for simulation

Composability and decomposability of NHPPs

The definition (1) implies that NHPPs are composable ([Cox and Miller 1965, par. 4.2](#)): merging two NHPPs with intensity functions $\lambda_1(t), \lambda_2(t)$ yields a new NHPPP with intensity function $\lambda(t) = \lambda_1(t) + \lambda_2(t)$. The reciprocal is also true: one can decompose an NHPPP with intensity function $\lambda(t)$ into two NHPPs, one with intensity function $\lambda_1(t) < \lambda(t)$ and one with intensity function $\lambda_2(t) = \lambda(t) - \lambda_1(t)$.³ An induction argument extends the above to merging and decomposing three or more processes.

The composability and decomposability properties are important for simulation because they

- give the flexibility to simulate several parallel NHPPs independently versus to merge them, simulate from the merged process, and then attribute the realized events to the component processes by assigning the i -th event to the j -th process with probability $\lambda_j(Z_i)/\lambda(Z_i)$, where $\lambda(t) = \sum \lambda_j(t)$.
- motivate a general sampling algorithm (Algorithm 4, “thinning” ([Lewis and Shedler 1979](#))) that simulates a target NHPPP with intensity $\lambda_1(t)$ by first drawing events from an easy-to-sample NHPPP with intensity $\lambda(t) > \lambda_1(t)$, and then accepts sample i with probability $\lambda_1(Z_i)/\lambda(Z_i)$.

²With some abuse of notation, we define $\Lambda(t) := \Lambda(0, t)$ when $a = 0$.

³The proof is omitted, but it involves showing that the composition and decomposition yields processes that conform with definition (1).

Transformations of the time axis

Strictly monotonic transformations of the carrier space of an NHPP yield an NHPP (Çinlar 1975). Consider an NHPP with intensity functions $\lambda(t)$ and a strictly monotonic transformation of the time axis $u : t \mapsto \tau$ that is differentiable once almost everywhere. On the transformed time axis the point process is an NHPP with intensity function

$$\rho(\tau) = \lambda(t) \left(\frac{du}{dt} \right)^{-1}. \quad (5)$$

This property is important for simulation because

- it motivates the use of another general sampling algorithm (Algorithm 5, “time transformation” or “inversion”, Çinlar (1975)): A smart choice for u yields an easy to sample point process. The event times in the original time scale can be obtained as $Z_i = u^{-1}(\zeta_i)$, where ζ_i is the i -th event in the transformed time axis and u^{-1} is the inverse function of u .
- given that at least i events have realized in the time interval $(a, b]$, it makes it possible to draw events $Z_{(j)}, j < i$ given event $Z_{(i)}$. This is useful for simulating earlier events conditional on the occurrence of a subsequent event. Choosing $u(t) := Z_{(i)} - t$ makes the time count backwards from $Z_{(i)}$. In this reversed clock we draw as if in forward time exactly $i - 1$ events $\zeta_{(1)}, \zeta_{(2)}, \dots, \zeta_{(i-1)}$. Back transforming yields all preceding events.

Table 1 summarizes the common simulation tasks, such as simulating single events (at most one, exactly one), a series of events (possibly demanding the occurrence of at least one event), or the occurrence of a prior (event $i - 1$ given $Z_{(i)}$). The **nhpp** package implements functions to simulate these tasks for general $\lambda(t)$ or $\Lambda(t)$.

3. Sampling the constant rate Poisson process

Sampling the constant rate Poisson process is straightforward. Algorithms 1 and 2 are two ways to sample event times in interval $(a, b]$ with constant intensity λ . Algorithm 3 describes sampling event times conditional on observing at least k events within the interval of interest.

3.1. Sequential sampling

Algorithm 1 samples events sequentially, using the fact that the inter-event times X_i are exponentially distributed with mean λ^{-1} (Cox and Miller 1965, par. 4.1). It involves generation only of exponential random variates, which is cheap on modern hardware. To sample at most k events, change the condition for the while loop in line 3 to

while $t < b$ & $|\mathcal{Z}| < k$ **do**.

The package’s `ppp_sequential()` function implements constant-rate sequential sampling that returns a vector with 0 or more event times in the interval $[a, b)$. The `range_t` argument is a two-values vector with the bounds a, b . The optional `tol` argument is used to get an upper bound on the realized number of events – using it speeds up the algorithm’s implementation

#	Sampling task	Sampled times	Number of sampled events	Example
I	Any next event	$\{\}$ or $\{Z_{(1)}\}$	0 or 1	Single event that may (or may not) occur in the interval: death, progression from Stage I to Stage II cancer.
II	Exactly one next event	$\{Z_{(1)}\}$	1	Single event which must occur in the interval: death from any cause in a lifetime-horizon simulation.
III	Any and all events	$\{\}$ or $\{Z_{(1)}, Z_{(2)}, \dots\}$	≥ 0	Zero, one, or more events: emergence of one or more bladder tumors.
IV	At least one next event	$\{Z_{(1)}, Z_{(2)}, \dots\}$	≥ 1	One or more events: emergence of bladder tumors when simulating only patients with bladder tumors.
V	Event $i - 1$ given $Z_{(i)}$	$\{Z_{(i-1)}\}$	1	Find the previous event when simulating conditional on a future event: time of symptom onset given the time of symptom-driven diagnosis; onset of Stage I cancer given progression from Stage I to Stage II cancer.

Table 1: Common simulation needs in discrete event simulation. All listed tasks involve sampling events over the interval $(a, b]$ with known $\lambda(t)$ or $\Lambda(t)$.

Algorithm 1 Sequential sampling of events in interval $(a, b]$ with constant intensity λ .

Require: $t \in (a, b]$

```

1:  $t \leftarrow a$ 
2:  $\mathcal{Z} \leftarrow \emptyset$ 
3: while  $t < b$  do
4:    $X \leftarrow X \sim \text{Exponential}(\lambda^{-1})$ 
5:    $t \leftarrow t + X$ 
6:   if  $t < b$  then
7:      $\mathcal{Z} \leftarrow \mathcal{Z} \cup \{t\}$ 
8:   end if
9: end while
10: return  $\mathcal{Z}$ 

```

in R. Setting the optional argument `only1` to `TRUE` from its default value of `FALSE` returns the first event or an empty vector, depending on whether at least one event is drawn in the

interval.

```
library("nhpp")
library("rstream")
ppp_sequential(range_t = c(0, 10), rate = 1, tol = 10^-6, only1 = FALSE)

## [1] 0.6738851 1.6505024 2.0112290 2.4075753 3.3888714 4.2948551
## [7] 4.3773730 4.7348883 5.5044973 5.5953475 8.8805621
```

All **nhpp** functions can accept a user provided random number stream object via the **rng_stream** option.

```
library("rstream")
S <- new("rstream.mrg32k3a")
ppp_sequential(
  range_t = c(0, 10), rate = 1, tol = 10^-6, only1 = FALSE,
  rng_stream = S
)

## [1] 0.7234174 0.8906377 1.2697303 2.0712702 3.5018199 5.1034342
## [7] 6.1769168 6.9603373 8.8045769 9.1952462
```

3.2. Sampling using order statistics

Algorithm 2 Sampling events in interval $(a, b]$ with constant intensity λ using order statistics.

Require: $t \in (a, b]$

```
1:  $N \leftarrow N \sim \text{Poisson}(\lambda(b - a))$ 
2:  $t \leftarrow a$ 
3:  $\mathcal{Z} \leftarrow \emptyset$  ▷  $\mathcal{Z}$  is an ordered set
4: if  $N > 0$  then
5:   for  $i \in [N]$  do:
6:      $U_i \leftarrow U_i \sim \text{Uniform}(0, 1)$  ▷ Generate order statistics
7:      $\mathcal{Z} \leftarrow \mathcal{Z} \cup \{a + (b - a)U_i\}$ 
8:   end for
9:    $\mathcal{Z} \leftarrow \text{sort}(\mathcal{Z})$ 
10: end if
11: return  $\mathcal{Z}$  ▷ Up to  $k$  earliest points: return  $\{Z_{(i)} \mid i \leq k, Z_{(i)} \in \mathcal{Z}\}$ 
```

Algorithm 2 first draws the number of events in $(a, b]$ from a Poisson distribution. Conditional on the number of events, the event times Z_i are uniformly distributed over $(a, b]$ (Cox and Miller 1965, par. 4.1). The algorithm returns the order statistics $[Z_{(i)}]$, obtained by sorting the event times $[Z_i]$ in ascending order. It is necessary to generate all event times to generate the order statistics. Thus, to sample at most k event times we should return the earliest k event times, and line 11 of the Algorithm would be changed to

return $\{Z_{(i)} \mid i \leq k, Z_{(i)} \in \mathcal{Z}\}.$

The `ppp_orderstat()` function implements constant-rate sampling via the order-statistics algorithm.

```
ppp_orderstat(range_t = c(3.14, 6.28), rate = 1, only1 = FALSE)

## [1] 3.202501 3.742638 4.573200 5.387799
```

3.3. Sampling conditional on observing at least m events

Algorithm 3 Sampling with constant intensity λ conditional that at least m events occurred in interval $(a, b]$. Relies on generating order statistics analogously to Algorithm 2.

Require: $t \in (a, b]$

```
1:  $N \leftarrow N \sim \text{TruncatedPoisson}_{N \geq m}(\lambda(b-a))$  ▷  $(m-1)$ -truncated Poisson
2:  $t \leftarrow a$ 
3:  $\mathcal{Z} \leftarrow \emptyset$  ▷  $\mathcal{Z}$  is an ordered set
4: if  $N > 0$  then
5:   for  $i \in [N]$  do:
6:      $U_i \leftarrow U_i \sim \text{Uniform}(0, 1)$  ▷ Generate order statistics
7:      $\mathcal{Z} \leftarrow \mathcal{Z} \cup \{a + (b-a)U_i\}$ 
8:   end for
9:    $\mathcal{Z} \leftarrow \text{sort}(\mathcal{Z})$ 
10: end if
11: return  $\mathcal{Z}$  ▷ Up to  $k$  earliest points: return  $\{Z_{(i)} \mid i \leq k, Z_{(i)} \in \mathcal{Z}\}$ 
```

Algorithm 3 is used to generate a point process conditional on observing at least m events. For example, if λ is the intensity of tumor generation, it can be used to simulate times of tumor emergence among patients with at least one ($m = 1$) tumor. To return the up to k earliest events, we modify line 11 the same way as for Algorithm 2. As an example, in a lifetime simulation we can sample the time of all-cause death by setting in Algorithm 3 $m = 1$, so that at least one event will happen in $(a, b]$, and $k = 1$, to sample only the time of the first event $Z_{(1)}$.

To sample exactly m events, change line 1 of Algorithm 3 to

$$N \leftarrow m.$$

Function `ztppp()` simulates times conditional on drawing at least one event - i.e., setting $m = 1$ in Algorithm 3 to sample from a zero truncated Poisson distribution in line 1.

```
ztppp(range_t = c(0, 10), rate = 0.001, only1 = FALSE)

## [1] 0.749629
```

Function `ppp_n()` simulates times conditional on drawing exactly m events.

```
ppp_n(size = 4, range_t = c(0, 10))

## [1] 3.066808 3.724638 4.265913 6.394102
```

4. The general sampling algorithms used in nhppp

The **nhppp** package uses three well known general sampling algorithms, namely thinning, time transformation or inversion, and order-statistics. These algorithms are efficiently combined to sample from special cases, including cases where the intensity function is a piecewise constant, linear, or log-linear function of time, as described in Section 5.2.

The thinning algorithm works with the intensity function $\lambda(t)$, which is commonly available. The inversion and order statistics algorithms have smaller computational cost than the thinning algorithm, but work with the integrated intensity function $\Lambda(t)$ and its inverse $\Lambda^{-1}(z)$, which may not be easily available. The generic function **draw()** is a wrapper function that dispatches to specialized functions depending on the provided arguments. It is useful for general use but the specialized functions are usually faster.

```
l <- function(t) t
L <- function(t) 0.5 * t^2
Li <- function(z) sqrt(2 * z)

draw(
  lambda = l, lambda_maj = l(10), range_t = c(5, 10),
  only1 = FALSE, atleast1 = FALSE
) |> head(n = 20)

## [1] 5.088108 5.197488 5.578401 5.760426 5.841001 5.916301 6.109621
## [8] 6.226782 6.239840 6.303556 6.420387 6.436196 6.448025 6.544486
## [15] 6.561641 7.027645 7.220043 7.414564 7.435004 7.533949

draw(
  Lambda = L, Lambda_inv = Li, range_t = c(5, 10),
  only1 = FALSE, atleast1 = FALSE
) |> head(n = 20)

## [1] 5.031636 5.467510 5.556208 5.889774 5.971826 6.041089 6.267354
## [8] 6.833528 7.330811 7.613152 7.782012 7.784715 7.788021 7.955658
## [15] 8.291194 8.372332 8.407120 8.441751 8.465296 8.537526
```

4.1. The thinning algorithm

The thinning algorithm relies on the decomposability of NHPs (Section 2.4) and is described in Lewis and Shedler (1979). Let the target NHP have intensity function $\lambda(t)$ and $\lambda_*(t) \geq \lambda(t)$ for all $t \in (a, b]$ be a majorizing intensity function. Think of the majorizing

function as an easy-to-sample function which is the sum of the intensity of the target point process $\lambda(t)$ and the intensity $\lambda_{reject}(t)$ of its complementary point-process,

$$\lambda_*(t) = \lambda(t) + \lambda_{reject}(t).$$

The acceptance-rejection scheme in Algorithm 4 generates proposal samples with intensity function $\lambda_*(t)$ and stochastically attributes them to the target process (to keep, with probability $\lambda(Z)/\lambda_*(Z)$) or its complement.

Algorithm 4 The thinning algorithm for sampling from $\lambda(t)$.

Require:

$\lambda_*(t) \geq \lambda(t) \ \forall t \in (a, \infty)$	▷ majorizing intensity function
$\mathcal{Z}_* = \{Z_i^* \mid Z_i^* \text{ are samples from } \lambda_*(t)\}$	▷ \mathcal{Z}_* is an ordered set
1: $N \leftarrow \mathcal{Z}_* $	
2: $\mathcal{Z} \leftarrow \emptyset$	▷ \mathcal{Z} is an ordered set
3: if $N > 0$ then	
4: for $i \in [N]$ do:	
5: $U_i \leftarrow U_i \sim \text{Uniform}(0, 1)$	
6: if $U_i < \lambda(Z_{(i)}^*)/\lambda_*(Z_{(i)}^*)$ then	
7: $\mathcal{Z} \leftarrow \mathcal{Z} \cup \{Z_{(i)}^*\}$	
8: end if	
9: end for	
10: end if	
11: return \mathcal{Z}	▷ Up to k earliest points: return $\{Z_{(i)} \mid i \leq k, Z_{(i)} \in \mathcal{Z}\}$

To sample the earliest k points, one can exit the for loop in lines 4-9 when k events have been sampled in line 7, or, alternatively, return the first up to k points in line 11.

A measure of the efficiency of Algorithm 4 is the proportion of samples that are accepted, which is

$$\frac{\int_a^b \lambda(t) \, dt}{\int_a^b \lambda_*(t) \, dt} \quad (6)$$

on average. Thus, the closer $\lambda_*(t)$ is to $\lambda(t)$, the more efficient the algorithm.

In practice, $\lambda_*(t)$ can be chosen as one of the special cases in Section 5, for which we have fast sampling algorithms. For example, it can be a piecewise constant majorizer. Algorithm 7 in Appendix A can be automatically generate a piecewise constant majorizer function for the general case of a K -Lipschitz continuous intensity function.

The **nhppp** package has functions that sample from time-varying intensity functions. The first function, **draw_intensity()**, expects a user-provided linear ($\lambda_*(t) = \alpha + \beta t$) or log-linear ($\lambda_*(t) = e^{\alpha + \beta t}$) majorizer function.

```
lambda_fun <- function(t) exp(0.02 * t)

draw_intensity(
  lambda = lambda_fun, # linear majorizer
  lambda_maj = c(intercept = 1.01, slope = 0.03),
```

```

    exp_maj = FALSE, range_t = c(0, 10), only1 = FALSE
  )

## [1] 0.2581614 1.3983756 3.7431439 3.8796828 4.4741099

draw_intensity(
  lambda = lambda_fun, # log-linear majorizer
  lambda_maj = c(intercept = 0.01, slope = 0.03),
  exp_maj = TRUE, range_t = c(0, 10), only1 = FALSE
)

## [1] 0.3562703 0.8397360 1.4451052 2.1072677 2.2543554 3.2642303
## [7] 3.6010669 4.0846178 4.1001892 5.0753928 5.3104299 5.8952167
## [13] 8.6209589 8.9008607

```

The second function, `draw_intensity_step()`, expects a user-provided piecewise linear majorizer

$$\lambda_*(t) = \begin{cases} \lambda_1 & \text{for } t \in [a_1, b_1) = [a, b_1), \\ \dots & \\ \lambda_m & \text{for } t \in [a_m, b_m) \text{ with } a_m = b_{m-1}, \\ \dots & \\ \lambda_M & \text{for } t \in [a_M, b_M) = [a_M, b), \end{cases}$$

which is specified as a vector of length $M + 1$ including the points $(a, [b_m]_{m=1}^M)$ and a vector of length M with the values $[\lambda_m]_{m=1}^M$ in each subinterval of $(a, b]$. For example, the following code splits the interval $(0, 10]$ into $M = 10$ subintervals of length one. Because `lambda_fun()` is strictly increasing, its value at the upper bound of each subinterval is the supremum of the interval (Appendix A, Algorithm 7).

```

draw_intensity_step(
  lambda = lambda_fun,
  lambda_maj_vector = lambda_fun(1:10), # 1:10 (10 intensity values)
  times_vector = 0:10, # 0:10 (11 interval bounds)
  only1 = FALSE
)

## [1] 0.03343148 2.61586529 2.63676775 3.07384008 3.38281065 6.77061576
## [7] 7.31728627 7.58168410 8.33557589 9.86127604 9.87429700 9.93368963

```

4.2. The time transformation or inversion algorithm

Algorithm 5 implements the time transformation or inversion algorithm from Çinlar (1975) and also Cox and Miller (1965, par. 4.2). As mentioned in Section 2.4, strictly monotonic transformations of the carrier space (here, time) of a Poisson Point Process yield another

Poisson Point Process. In equation (5), choosing the transformation $\tau = u(t) = \Lambda(t)$, so that $\frac{du(t)}{dt} = \lambda(t)$, results in $\rho(\tau) = 1$.

This means (proof sketched in Cox and Miller (1965, par. 4.2)) that we can sample points from a Poisson point process with intensity one over the interval $(\tau_a, \tau_b] = (\Lambda(a), \Lambda(b)]$. Via a similar argument, we transform event times sampled on the transformed scale back to the original scale using $g(t) = \Lambda^{-1}(\tau)$. The transformations $u(\cdot), g(\cdot)$ are not unique – at least up to the group of affine transformations.

Function `draw_cumulative_intensity_inversion()` works with a cumulative intensity function $\Lambda(t)$ and its inverse $\Lambda^{-1}(z)$, if available. If the inverse function is not available (argument `Lambda_inv = NULL`), the Brent bisection algorithm is used to invert $\Lambda(t)$ numerically, at a performance cost (Press, Teukolsky, Vetterling, and Flannery 2007).

```
Lambda_fun <- function(t) 50 * exp(0.02 * t) - 50
Lambda_inv_fun <- function(z) 50 * log((z + 50) / 50)

draw_cumulative_intensity_inversion(
  Lambda = Lambda_fun,
  Lambda_inv = Lambda_inv_fun,
  range_t = c(5, 10.5),
  range_L = Lambda_fun(c(5, 10.5)),
  only1 = FALSE
)

## [1] 5.096645 5.390107 5.408923 6.705197 6.870329 7.329849
## [7] 8.088067 9.244232 10.304507
```

Algorithm 5 The time transformation or inversion algorithm for sampling given $\Lambda(t), \Lambda^{-1}(z)$ (Çinlar 1975; Cox and Miller 1965). The notation `PoissonProcess1` indicates sampling event times from a constant rate one Poisson point process.

Require: $\Lambda(t), \Lambda^{-1}(z), t \in (a, b]$ $\triangleright \Lambda^{-1}(z)$ possibly numerically

- 1: $\tau_a \leftarrow \Lambda(a), \tau_b \leftarrow \Lambda(b)$
- 2: $\mathcal{C} \leftarrow \mathcal{C} \sim \text{PoissonProcess}_1(\tau_a, \tau_b)$ \triangleright From Algorithm 1 (or 3 for conditional sampling)
- 3: $\mathcal{Z} \leftarrow \Lambda^{-1}(\mathcal{C})$ $\triangleright \Lambda^{-1}(\cdot)$ as set function, meant elementwise
- 4: **return** \mathcal{Z}

4.3. The order statistics algorithm

The general order statistics algorithm (Algorithm 6) is a direct generalization of Algorithm 2. It first draws the number N of realized events. Conditional on N

$$\begin{aligned} U_{(i)} &= \frac{\Lambda(Z_{(i)}) - \Lambda(a)}{\Lambda(b) - \Lambda(a)} \sim \text{Uniform}(0, 1), \\ Z_{(i)} &= \Lambda^{-1}\left(\Lambda(a) + U_{(i)}(\Lambda(b) - \Lambda(a))\right), \end{aligned} \tag{7}$$

as discussed in Lewis and Shedler (1979). Algorithm 6 makes the above explicit.

Algorithm 6 The order statistics algorithm for sampling from an NHPPP given $\Lambda(t), \Lambda^{-1}(z)$.

Require: $\Lambda(t), \Lambda^{-1}(z), t \in (a, b]$ $\triangleright \Lambda^{-1}(z)$ possibly numerically

- 1: $N \leftarrow N \sim \text{Poisson}(\Lambda(b) - \Lambda(a))$
- 2: $t \leftarrow a$
- 3: $\mathcal{Z} \leftarrow \emptyset$ $\triangleright \mathcal{Z}$ is an ordered set
- 4: **if** $N > 0$ **then**
- 5: **for** $i \in [N]$ **do**:
- 6: $U_i \leftarrow U_i \sim \text{Uniform}(0, 1)$ \triangleright Generate order statistics
- 7: $\mathcal{Z} \leftarrow \mathcal{Z} \cup \{\Lambda^{-1}(\Lambda(a) + U_i(\Lambda(b) - \Lambda(a)))\}$
- 8: **end for**
- 9: $\mathcal{Z} \leftarrow \text{sort}(\mathcal{Z})$
- 10: **end if**
- 11: **return** \mathcal{Z} \triangleright Up to k earliest points: **return** $\{Z_{(i)} \mid i \leq k, Z_{(i)} \in \mathcal{Z}\}$

Sampling up to k earliest points means returning the up to k earliest event times. If $\Lambda(t)$ is a positive linear function of time, λ is constant and Algorithm 6 becomes Algorithm 2.

To sample conditional on observing at least m events in the interval $(a, b]$, modify line 1 of Algorithm 6 to sample from a $(m - 1)$ -truncated Poisson distribution (Appendix B, Algorithm 8).

$$N \leftarrow N \sim \text{TruncatedPoisson}_{N \geq m}(\Lambda(b) - \Lambda(a)).$$

Function `draw_cumulative_intensity_orderstats()` works with a cumulative intensity function $\Lambda(t)$ and its inverse $\Lambda^{-1}(z)$, if available. Function `ztdraw_cumulative_intensity()` conditions that at least one event is sampled in the interval. If the inverse function is not available (argument `Lambda_inv = NULL`), the Brent bisection algorithm is used to invert $\Lambda(t)$ numerically, at a performance cost.

```
draw_cumulative_intensity_orderstats(
  Lambda = Lambda_fun,
  Lambda_inv = Lambda_inv_fun,
  range_t = c(4.1, 7.6),
  only1 = FALSE
)

## [1] 4.553782 5.441297 6.180854 6.856586 7.133086 7.308007
```

```
ztdraw_cumulative_intensity(
  Lambda = Lambda_fun,
  Lambda_inv = Lambda_inv_fun,
  range_t = c(4.1, 7.6),
  only1 = FALSE
)

## [1] 4.753191 5.126775 5.644922 5.820607 7.313658 7.500967
```

5. Special cases

The **nhppp** package implements several special cases where the intensity function $\lambda(\cdot)$, the integrated intensity function $\Lambda(\cdot)$, and its inverse $\Lambda^{-1}(\cdot)$ have straightforward analytical expressions.

5.1. Sampling a piecewise constant NHPPP

Functions `draw_sc_step()` and `draw_sc_step_regular()` sample piecewise constant intensity functions based on Algorithm 5. The first can work with unequal-length subintervals $(a_m, b_m]$. The second results in a small computational time improvement when all subintervals are of equal length.

```
draw_sc_step(
  lambda_vector = 1:5, times_vector = c(0.5, 1, 2.4, 3.1, 4.9, 5.9),
  only1 = FALSE, atleast1 = FALSE
)

## [1] 1.108874 1.257993 1.537622 1.902278 2.187478 3.074243 3.468703
## [8] 3.663942 3.668432 3.770789 4.208698 4.217727 4.251112 4.421734
## [15] 5.207486 5.337123 5.363360 5.390178 5.391736 5.412233 5.752539

draw_sc_step_regular(
  lambda_vector = 1:5, range_t = c(0.5, 5.9), only1 = FALSE,
  atleast1 = FALSE
)

## [1] 0.5895937 2.0516267 2.7937694 2.9536086 3.3331037 3.5542259
## [7] 3.7629171 4.4524616 4.5786990 4.9336974 4.9527614 5.0204750
## [13] 5.1448457 5.2911578 5.4356068 5.7565030 5.7599498
```

Function `vdraw_sc_step_regular()` is a vectorized version of `draw_sc_step_regular()`. It returns a matrix with one event series per row, and as many columns as the maximum number of events across all draws.

```
vdraw_sc_step_regular(
  lambda_matrix = matrix(runif(20), ncol = 5), range_t = c(1,4),
  only1 = FALSE
)

##           [,1]      [,2]      [,3]      [,4]      [,5]
## [1,] 1.662662 1.732405 3.929535         NA         NA
## [2,] 2.946469 3.105219 3.638962         NA         NA
## [3,] 1.890595 2.195388 2.741042 3.671363 3.826785
## [4,] 2.143551 2.242584         NA         NA         NA
```

5.2. Sampling NHPPs with linear and log-linear intensities

Functions `draw_sc_linear()` and `ztdraw_sc_linear()` sample zero or more and at least one event, respectively, from NHPPs with linear intensity functions. An optional argument (`only1`) returns the first event only.

$$\lambda(t) = \begin{cases} \alpha + \beta t & \text{for } t \in [a, b), t > -\frac{\alpha}{\beta} \\ 0 & \text{otherwise} \end{cases}.$$

```
draw_sc_linear(
  alpha = 3, beta = -0.5, range_t = c(0, 10),
  only1 = FALSE
)

## [1] 0.9329605 1.6433187 1.9847112 2.0635383 2.3302896 2.7281810

ztdraw_sc_linear(
  alpha = 0.5, beta = 0.2, range_t = c(0, 10),
  only1 = FALSE
)

## [1] 0.7922714 1.4901395 1.9336521 2.4992449 2.5149976 2.5233687
## [7] 3.4027855 3.8154668 4.8094813 5.0791610 6.8615724 7.3394189
## [13] 7.9330570 8.5009305 8.6498476 9.4013359 9.6029110 9.6178765
```

An analogous set of functions (`[nhppp|ztnhppp]_sc_loglinear()`) samples from log-linear intensity functions

$$\lambda(t) = \begin{cases} e^{\alpha + \beta t} & \text{for } t \in [a, b) \\ 0 & \text{otherwise} \end{cases}.$$

The sampling algorithm is a variation of Algorithm 5, as described in [Lewis and Shedler \(1976\)](#). Example usage follows.

```
draw_sc_loglinear(
  alpha = 1, beta = -0.02, range_t = c(8, 10),
  only1 = FALSE
)

## [1] 8.445952

ztdraw_sc_loglinear(
  alpha = 1, beta = -0.02, range_t = c(9.8, 10),
  only1 = FALSE
)

## [1] 9.820388
```

6. Comparisons with other R packages

Table 2 lists five R packages that simulate from NHPPPs, including **nhppp**.

Package **reda** (Wang, Fu, and Yan 2022) focuses on recurrent event data analysis and can simulate NHPPPs with the inversion and the thinning algorithms using the `simEvent()` function. It can take function object arguments for $\lambda(t)$. When using the thinning algorithm, it takes a constant majorizer. For the inversion algorithm, it approximates $\Lambda(t)$ and its inverse numerically, at a computational cost.

Package **simEd** (Lawson, Leemis, and Kudlay 2023) includes various functions for simulation education. Function `thinning()` implements the homonymous algorithm for drawing points from an NHPPP. Users can specify the intensity function and a piecewise constant or linear majorizer function.

Package **IndTestPP** (Cebrián 2020a) provides a framework for exploring the dependence between two or more realizations of point processes. It includes the ancillary function `simNHPc()` for simulating NHPPPs with the inversion or thinning algorithms. The function's argument is a piecewise constant approximation of the intensity function via a vector of evaluations, each corresponding to unit length subintervals. This resolution may not be adequate to simulate processes that change fast over a unit time interval.

Package **NHPoisson** (Cebrián, Abaurrea, and Asín 2015; Cebrián 2020b) fits NHPPP models to data and is not really geared towards mathematical simulation. Its `simNHP.fun()` function provides the ability for simulation-based inference via an implementation of the inversion algorithm. This function is designed to work with the package's inference machinery and is not practical to use for simulation. This is because the user has no direct control over the function's rescaling of the time axis.

The claimed advantage of **nhppp** over the existing packages is that

- it samples from the target NHPPP and not from a numerical approximation thereof, e.g., as **IndTestPP** does.
- It can sample conditional on observing at least one event in the interval, which no other package implement.
- It accepts user-provided random number stream objects, which is useful for implementing simulation variance reduction techniques such as common random numbers (Wright and Ramsay Jr 1979) and antithetic variates (Hammersley and Mauldon 1956).
- It is fast, even though it is currently implemented primarily in R-only code. It has specialized functions to leverage additional information about the process, such as $\Lambda(t)$, $\Lambda^{-1}(z)$, when available, which can result in faster simulation use the cumulative intensity function and its inverse, often at a computational speed advantage. It also includes a vectorized function, `vdraw_sc_step_regular()`, which draws from piecewise constant intensity functions.⁴

⁴More functions will be vectorized in future releases.

R package	Function	Algorithms (inputs)			Sample only earliest event	Custom RNG	Simulate given $N > 0$	Vectorized functions
		Thinning	Inversion	Order statistics				
nhppp	[see text]	$\lambda(t), \lambda_*(t)$	$\Lambda(t), \Lambda^{-1}(z)$	$\Lambda(t), \Lambda^{-1}(z)$	Yes	rstream objects	Yes	For piecewise constant intensity
reda	<code>simEvent()</code>	$\lambda(t), \lambda_*$ constant	$\lambda(t)$ (no $\Lambda(t), \Lambda^{-1}(z)$)	No	Yes	No	No	No
simEd	<code>thinning()</code>	$\lambda(t), [\lambda_{**m}]_{m=1}^M$	No	No	No	No	No	No
IndTestPP	<code>simNHPPec()</code>	$[\lambda_m]_{m=1}^M, \lambda_*$ constant	$[\lambda_m]_{m=1}^M$ (no $\Lambda(t), \Lambda^{-1}(z)$)	No	No	No	No	No
NHPPoisson	<code>simNHPP.fun()</code>	No	$\lambda(t)$, (no $\Lambda(t), \Lambda^{-1}(z)$)	No	No	No	No	No

Table 2: NHPPP generation in R packages. RNG: random number generator object.

7. Illustrations

Depending on the application, we may have access to the intensity function or the integrated intensity function. We compared the R packages in Table 2 for sampling from a non-monotone and highly nonlinear intensity function for which the integrated intensity function is known analytically.

7.1. The target NHPPP to be simulated

Consider the example

$$\begin{aligned}\lambda(t) &= e^{rt}(1 + \sin wt), \\ \Lambda(t) &= \frac{e^{rt}(r \sin wt - w \cos wt) + w}{r^2 + w^2} + \frac{e^{rt} - 1}{r}\end{aligned}\tag{8}$$

of a sinusoidal intensity function $\lambda(t)$ scaled to have an exponential amplitude and one of its antiderivatives $\Lambda(t)$, with such a constant term that $\Lambda(0) = 0$. For the numerical study we set $r = 0.2$, $w = 1$, and $t \in (0, 6\pi]$. There is no analytic inverse function for this example. However, we can precompute $\text{Li}()$, a good numerical approximation to $\Lambda^{-1}(z)$. We will use it in Section 7.5 to compare the time performance of functions that use the inversion and order statistics algorithms when Λ^{-1} is available versus not.

```
l <- function(t) (1 + sin(t)) * exp(0.2 * t)
L <- function(t) {
  exp(0.2 * t) * (0.2 * sin(t) - cos(t)) / 1.04 +
  exp(0.2 * t) / 0.2 - 4.038462
}
Li <- approxfun(
  x = L(seq(0, 6 * pi, 10^-3)),
  y = seq(0, 6 * pi, 10^-3), rule = 2
)
```

Figure 1 graphs the intensity function and three majorizing functions over the interval of interest, which will be needed for the thinning algorithm.

The first, $\lambda_{*a}(t) = 43.38$, shown as a dashed blue line, is a constant majorizer equal to the maximum of the intensity function. A constant majorizer may be a practical choice when only an upper bound is known for $\lambda(t)$. From (6), the efficiency of the thinning algorithm using this majorizer is 0.209.

The second, $\lambda_{*b}(t)$, shown as a thin black line, is a piecewise constant envelope generated automatically from Algorithm 7 (Appendix A) with 20 equal-length subintervals and $K = 52.05$, where the Lipschitz cone coefficient K is equal to the maximum value of $|\frac{d\lambda(t)}{dt}|$ in the interval, attained at 6π . The efficiency of the thinning algorithm using this majorizer is 0.245.

The third, $\lambda_{*c}(t)$, shown as a thicker black line, is a tighter piecewise constant majorizer with the same 20 equal-length subintervals that is constructed by finding a least upper bound in each subinterval. The efficiency of the thinning algorithm with the third majorizer is 0.718.

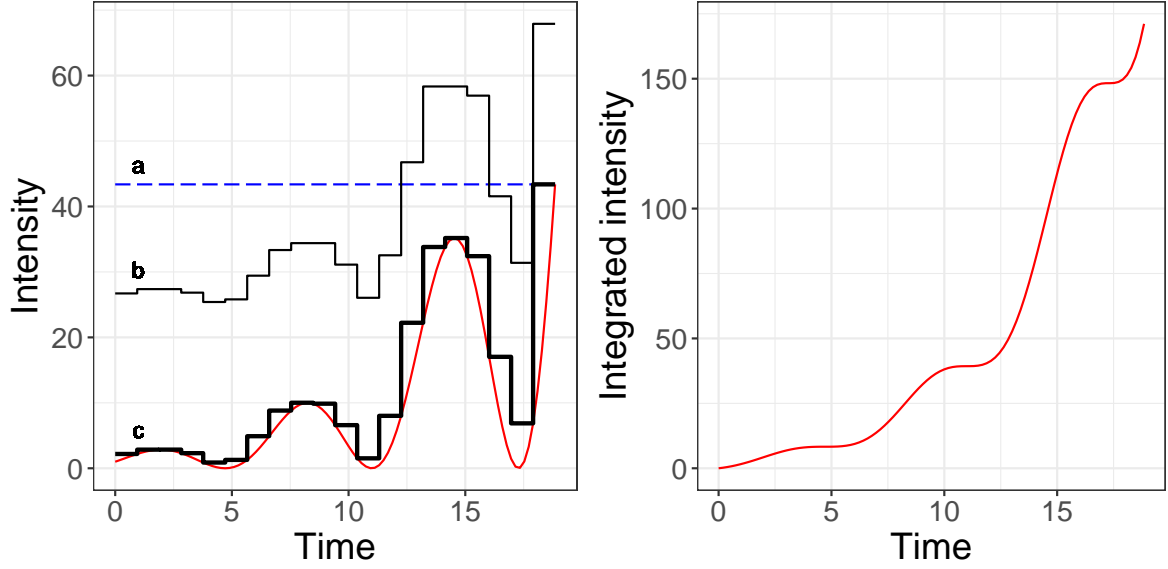


Figure 1: The $\lambda(t)$ (left) and $\Lambda(t)$ used in the illustration. Also shown three majorizing functions (left panel, marked a, b, c) that are used with the thinning algorithm in the analyses.

7.2. Simulation functions and algorithms

We sampled series of events from the target NHPPP using the packages and functions listed in Table 2. We repeated the sampling 100 times, recording all simulated points (event times). We also recorded the median computation time for drawing one series of events with single-threaded computation on modern hardware (Apple M1 Max laptop with 64 megabytes of random access memory).

From the **nhppp** package we use

1. two functions that take as argument the intensity function and are based on Algorithm 4 (thinning): `draw_intensity()`, which uses linear majorizers such as λ_{*a} , and `draw_intensity_step()`, which uses piecewise constant majorizers such as λ_{*b} and λ_{*c} in the example.
2. Function `draw_cumulative_intensity_inversion()`, which takes as argument the cumulative intensity function $\Lambda(t)$ and is based on Algorithm 5 (time transformation/inversion), and
3. function `draw_cumulative_intensity_orderstats()`, which also uses $\Lambda(t)$ and is based on Algorithm 6 (order statistics).

Regarding the other R packages in Table 2, we used all except for **NHPoisson**, whose simulation function is tailored to supporting simulation based inference for data analysis and is not practical to use as a standalone function.⁵ However, its source code/algorithm is very similar to that of the **IndTestPP** simulation function, which is developed by the same authors.

⁵The implementation does not allow the user to control the scaling of the time axis in a practical way.

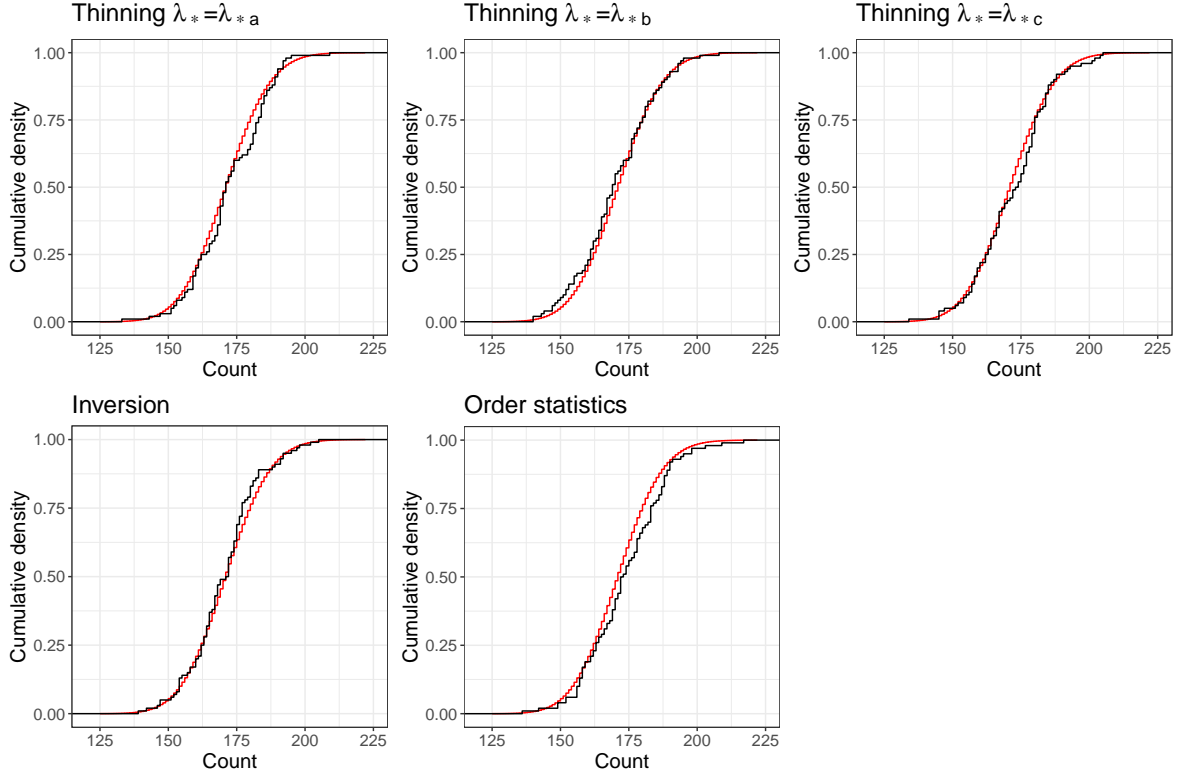


Figure 2: Theoretical (red) and empirical (black) cumulative distribution functions for event counts in the illustration example with **nhppp** functions. The unsigned area between the theoretical and empirical curves equals the Wasserstein-1 distance in Table 4.

We used the metrics in Table 3 to assess simulation performance with each function. We compared the empirical versus the simulated distributions of number of events and event times over $J = 100$ simulation runs.

7.3. Simulation performance with respect to number of events

We calculated the absolute and relative bias in the first two moments of the empirical distribution in the counts of events, the bounds of equal-tailed confidence intervals at the 95, 90, 75, and 50 percent levels, a χ^2 -distributed goodness of fit statistic and its p -value, and the Wasserstein-1 distance W_1 between the empirical and the theoretical count distributions and the asymptotic one sided p value to reject whether $W_1 = 0$ according to [Sommerfeld and Munk \(2018\)](#). W_1 is the smallest mass that has to be redistributed so that one distribution matches the other. W_1 is equal to the unsigned area between the cumulative distribution functions of the compared distributions. For example, $W_1 = 5.25$ means that the mass that must be moved to transform one density to the other is no less than 5.25 counts and a $W_1 = 0$ implies perfect fit.

The results for the **nhppp** functions in Table 4 and Figure 2 suggest excellent simulation performance.

Metric	Definition	Description
Bias in mean	$B_\mu = \frac{1}{J} \sum_j n_j - N$	Mean difference from target in the number of counts.
Relative bias in mean	$B_{\mu,rel} = \frac{B_\mu}{N}$	Mean proportional difference from target in the number of counts.
Bias in variance	$B_V = \frac{1}{J} \sum_j (n_j - \frac{1}{J} \sum_j n_j)^2 - V$	Mean difference from target in variance of counts.
Relative bias in variance	$B_{V,rel} = \frac{B_V}{V}$	Mean proportional difference from target in variance of counts.
Equal-tailed $p\%$ confidence interval bounds	$n_{[p/2]}, n_{[1-p/2]}$	Quantiles of the empirical distribution of counts.
Goodness of fit p value	Statistic $\sum_x \frac{(O_x - E_x)^2}{E_x} \sim \chi_{U-L+1}^2$	Left-tail p value. p values near 1 imply good fit.
Wasserstein-1 distance	W_1 , the smallest rearrangement of probability mass so that one distribution matches the other.	$W_1 = 0$ implies good fit
p value for $W_1 \neq 0$	Asymptotic theory p value	Two-sided p value. p values near 1 imply good fit.

Table 3: Simulation metrics for the number of counts. In the Table, $j \in [J]$ indexes simulations, n_j is the number of counts in simulation j , $N = \Lambda(6\pi) - \Lambda(0)$ is the theoretical mean number of counts, and $V = \Lambda(6\pi) - \Lambda(0) = N$ the theoretical variance. The lower and upper bounds of an equal-tailed $p\%$ confidence interval, $p \in \{95, 90, 75, 50\}$, are denoted with $n_{[p/2]}, n_{[1-p/2]}$, respectively. For the goodness of fit, we created bins $[0, L), [L, L+1), \dots, [U, \infty)$, where L, U are the 0.001 and 0.999 percentiles of the Poisson distribution with parameter $\Lambda(6\pi) - \Lambda(0)$. We indexed bins with $x \in \{1, \dots, U - L + 2\}$. The goodness of fit statistic contrasts the observed (O_x) versus expected (E_x) numbers of events over the bins and it is compared with a χ_{U-L+1}^2 distribution to obtain a p value.

	Thinning $\lambda_*=\lambda_{*a}$	Thinning $\lambda_*=\lambda_{*b}$	Thinning $\lambda_*=\lambda_{*c}$	Inversion	Order statistics
Sample mean	172.440	170.030	172.020	170.350	173.680
B_μ	1.305	-1.105	0.885	-0.785	2.545
$B_{\mu,rel}$	0.763	-0.646	0.517	-0.459	1.487
Sample variance	171.865	196.938	187.333	165.745	202.947
B_V	0.730	25.804	16.198	-5.390	31.812
$B_{V,rel}$	0.427	15.078	9.465	-3.149	18.589
Goodness of fit, χ^2 [p value]	4.090 [1.000]	4.263 [1.000]	3.563 [1.000]	4.269 [1.000]	5.572 [1.000]
W_1 [p value]	1.759 [0.204]	1.427 [0.356]	1.325 [0.074]	1.244 [0.066]	2.631 [0.114]
Equal tail 95% CI = [146, 197]	[148, 192]	[143, 194]	[145, 201]	[146, 197]	[149, 200]
Equal tail 90% CI = [150, 193]	[151, 192]	[147, 193]	[150, 193]	[150, 192]	[152, 195]
Equal tail 75% CI = [156, 186]	[159, 187]	[153, 186]	[157, 185]	[154, 183]	[157, 188]
Equal tail 50% CI = [162, 180]	[163, 183]	[161, 180]	[162, 180]	[162, 177]	[163, 183]

Table 4: Simulated total number of events with **nhppp** functions for the illustration example. Equal tail $p\%$ CI: a confidence interval whose bounds are the $p/2$ and $(1 - p/2)$ count percentiles of the respective cumulative distribution function.

The respective results for the R packages are in Figure 3 and Table 5. The simulation performance with the **reda** functions is excellent. Performance with **simEd** and **IndTestPP** functions depends on the adequacy with which they approximate the target density. In this example, the approximation accuracy is not ideal for either package, but is somewhat worse for **IndTestPP**.

	reda thinning, $\lambda_*=\lambda_{*a}$	reda inversion	simEd thin- ning, $\lambda_*=\lambda_{*a}$	IndTestPP thinning, no λ_*	IndTestPP in- version
Sample mean	169.780	173.620	179.520	189.300	190.720
B_μ	-1.355	2.485	8.385	18.165	19.585
$B_{\mu,rel}$	-0.792	1.452	4.900	10.615	11.444
Sample variance	177.103	148.097	135.949	164.798	216.325
B_V	5.968	-23.038	-35.186	-6.337	45.190
$B_{V,rel}$	3.487	-13.462	-20.560	-3.703	26.406
Goodness of fit, χ^2 [p value]	5.807 [1.000]	10.794 [1.000]	59.829 [0.999]	201.416 [<0.001]	229.035 [<0.001]
W_1 [p value]	1.437 [0.106]	2.536 [0.066]	8.714 [0.006]	18.167 [0.024]	19.585 [0.206]
Equal tail 95% CI = [146, 197]	[144, 194]	[154, 196]	[158, 196]	[163, 212]	[167, 223]
Equal tail 90% CI = [150, 193]	[147, 192]	[156, 193]	[158, 196]	[165, 210]	[170, 215]
Equal tail 75% CI = [156, 186]	[153, 184]	[160, 187]	[165, 192]	[174, 202]	[174, 206]
Equal tail 50% CI = [162, 180]	[161, 179]	[163, 182]	[173, 190]	[181, 197]	[179, 202]

Table 5: Simulated total number of events with the R packages of Table 2 for the illustration example. Equal tail $p\%$ CI: a confidence interval whose bounds are the $p/2$ and $(1 - p/2)$ count percentiles of the respective cumulative distribution function.

7.4. Event times

We compared the theoretical and empirical distribution of event times for all $J = 100$ event time draws. We calculated a goodness of fit statistic by binning realized times in 70 bins and its p value, by comparing the statistic against the χ^2_{69} distribution. We also calculated the W_1 distance between these distributions and its associated p value.

Figure 4 and Table 6 indicate excellent simulation performance with the **nhppp** functions.

Figure 5 and Table 7 indicate excellent simulation performance with the **reda** functions. The

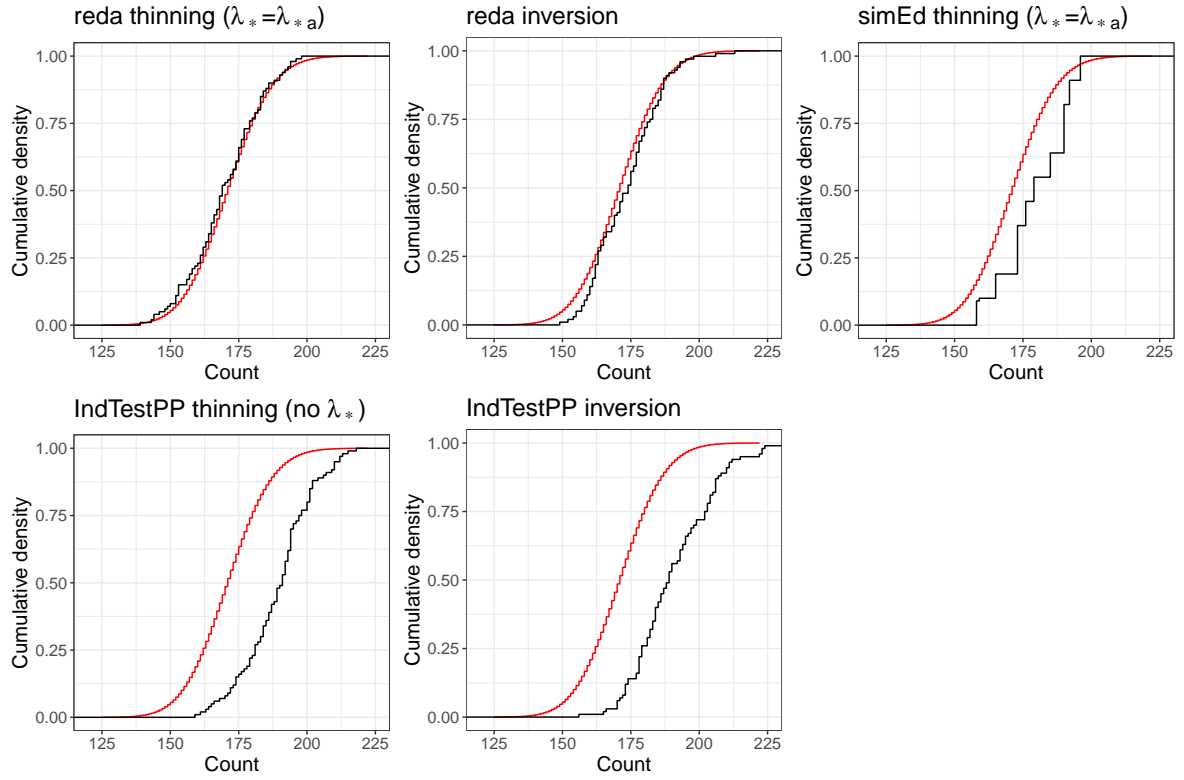


Figure 3: Theoretical (red) and empirical (black) cumulative distribution functions for event counts in the illustration example with the R packages in Table 2. The unsigned area between the theoretical and empirical curves equals the Wasserstein-1 distance in Table 4.

simulation performance with the **simEd** and **IndTestPP** functions, which rely on approximations, is not as good.

	Goodness of fit, χ^2 [p value]	W_1 [p value]
Thinning $\lambda_* = \lambda_{*a}$	0.008 [1.000]	0.561 [1.000]
Thinning $\lambda_* = \lambda_{*b}$	0.009 [1.000]	0.378 [0.970]
Thinning $\lambda_* = \lambda_{*c}$	0.011 [1.000]	0.452 [1.000]
Inversion	0.009 [1.000]	0.374 [0.996]
Order statistics	0.009 [1.000]	0.407 [1.000]

Table 6: Goodness of fit of simulated event times with **nhppp** functions for the example.

	Goodness of fit, χ^2 [p value]	W_1 [p value]
reda thinning ($\lambda_* = \lambda_{*a}$)	0.010 [1.000]	0.279 [1.000]
reda inversion	0.010 [1.000]	0.379 [1.000]
simEd thinning ($\lambda_* = \lambda_{*a}$)	0.054 [1.000]	0.520 [0.766]
IndTestPP thinning (no λ_*)	0.391 [1.000]	2.288 [0.870]
IndTestPP inversion	0.515 [1.000]	2.136 [0.874]

Table 7: Goodness of fit of simulated event times with R functions in Table 2.

7.5. Time performance

Time performance of nonvectorized functions

To indicate time performance, we benchmarked functions by recording execution times when drawing a series of points (Figure 6). We also benchmarked functions for drawing the first-occurring event, because **nhppp** functions can sample the first time more efficiently when the inversion algorithm is used (Figure 7).

We provided functions with the arguments they need to run fastest. For example, functions that use the inversion or order statistics algorithm execute faster when the inverse function $\Lambda^{-1}(z)$ is provided, rather than numerically calculated, as shown in both Figures for the **nhppp** package. (Functions in other packages do not take $\Lambda(t)$ and $\Lambda^{-1}(z)$ arguments.) The fastest functions are **nhppp** functions that rely on the inversion or order statistics algorithms given $\Lambda^{-1}(z)$.

According to (6), the thinning algorithm has higher efficiency, and is expected to execute faster, for majorizer functions that envelop the intensity function more closely. Observe that $\lambda_{*a} \succ \lambda_{*c}$ and $\lambda_{*b} \succ \lambda_{*c}$ in Figure 1. As expected, the execution times are indeed shorter for majorizer ‘c’ compared to ‘b’ in Figures 6 and 7. However, the execution times are longer with majorizer ‘c’ compared to ‘a’ because **draw_intensity()**, the function that uses constant majorizers, has a different implementation than **draw_intensity_step()**, the function that use piecewise constant majorizers. **draw_intensity()** happens to be faster in this example, but this is not always the case.

In **nhppp**, functions that use the inversion or order statistics algorithms can exit earlier when only the first event is requested. This is not possible, however, for the thinning algorithm. This efficiency does not appear to be implemented in the other packages.

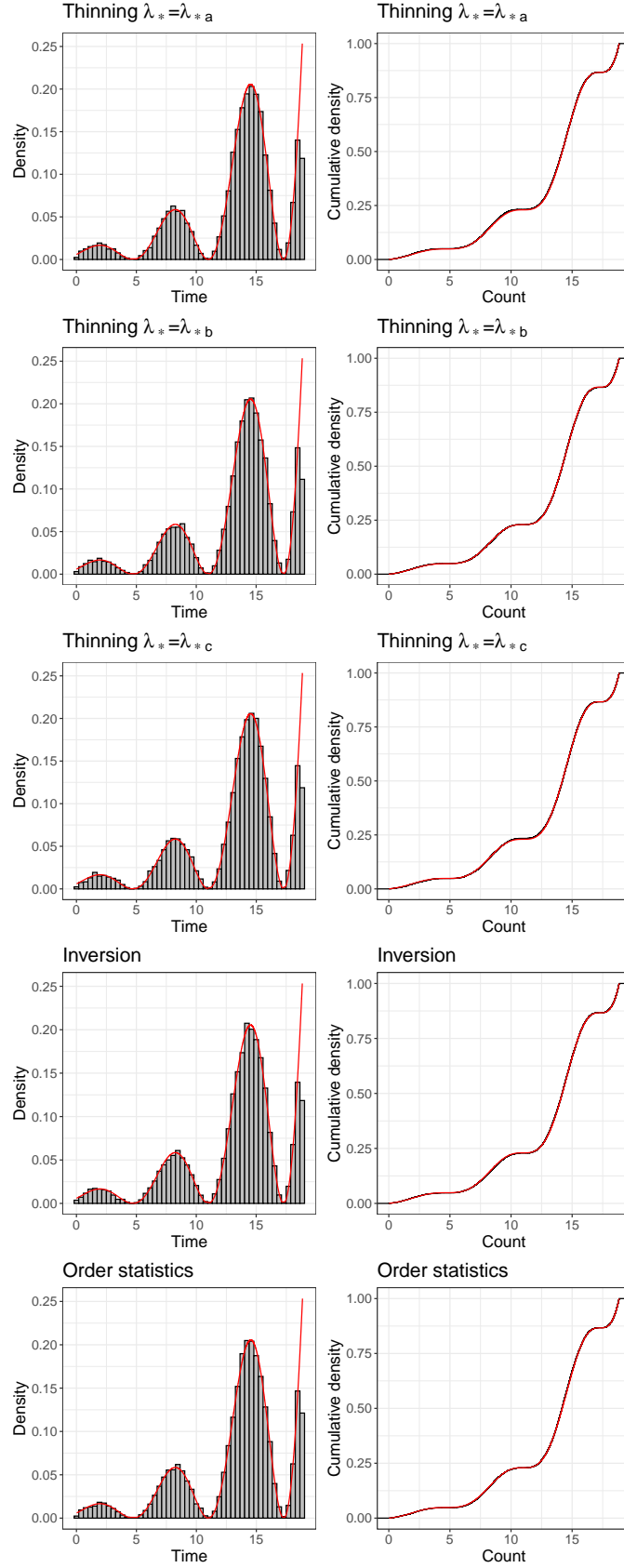


Figure 4: Simulated event times with **nhppp**. Left column: histogram (gray) and theoretical distribution (red) of event times; right column: empirical (black) and theoretical (red) cumulative distribution function. The unsigned area between the empirical and cumulative distribution functions is the W_1 distance in Table 6.

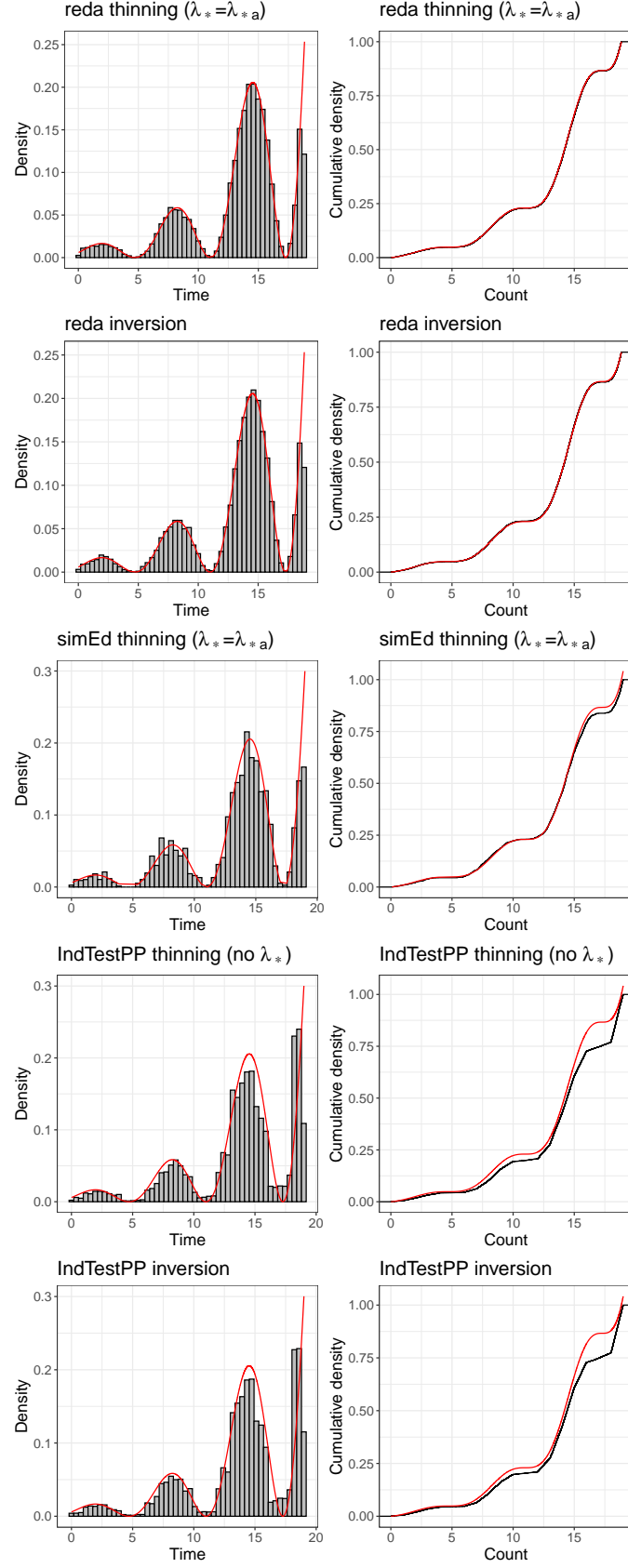


Figure 5: Simulated event times with the R packages in Table 2. Left column: histogram (gray) and theoretical distribution (red) of event times; right column: empirical (black) and theoretical (red) cumulative distribution function. The unsigned area between the empirical and cumulative distribution functions is the W_1 distance in Table 7.

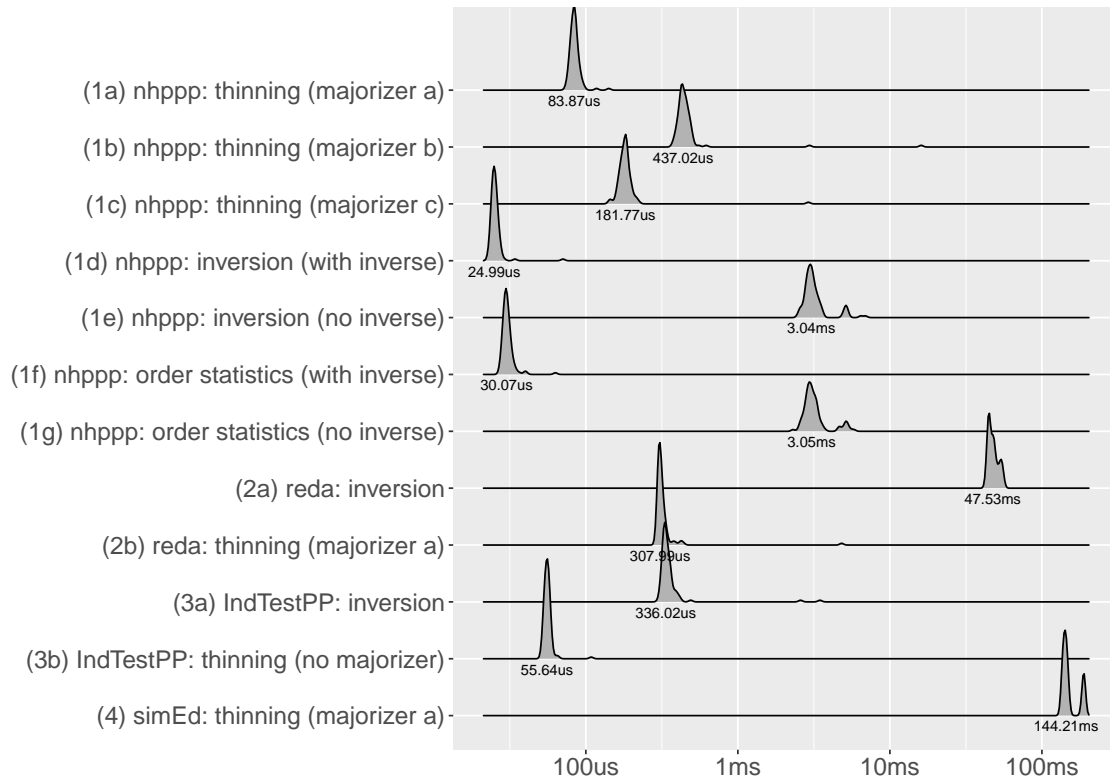


Figure 6: Computation times when drawing all events in interval. Unit **us** is μs (microsecond).

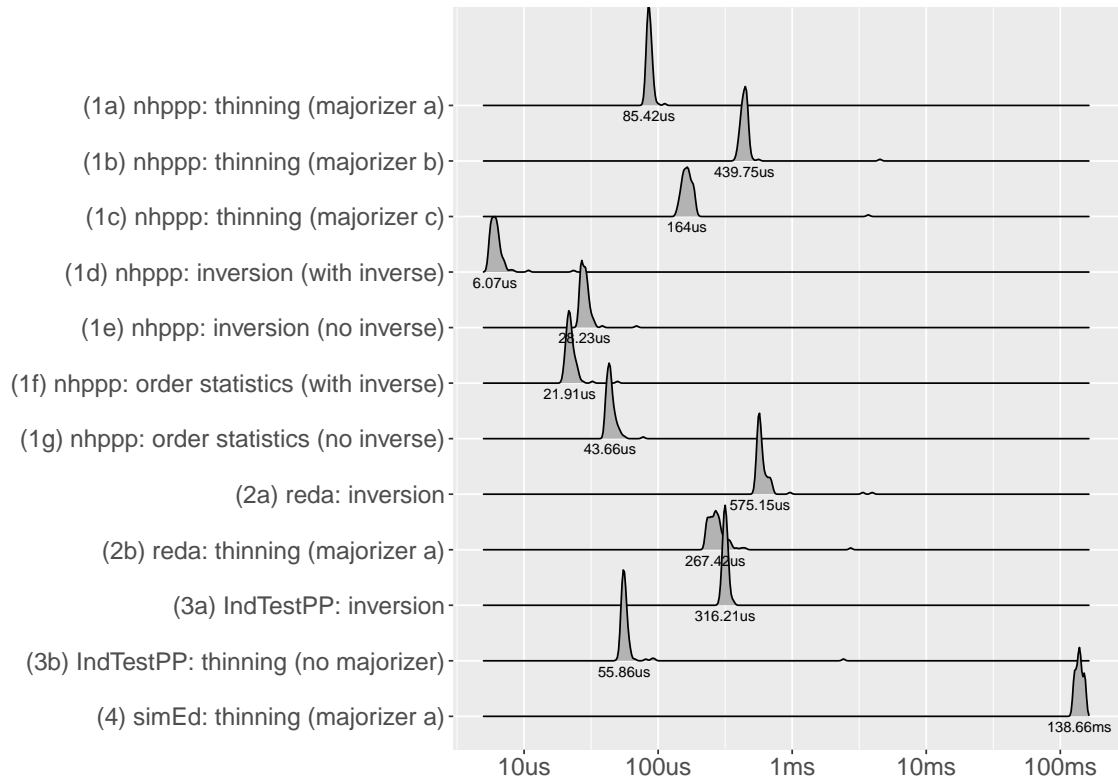


Figure 7: Computation times when drawing the first event in interval. Unit **us** is μs (microsecond).

Time performance of vectorized functions

In R, ‘vectorized’ computation, where operations are done in columns, is faster than using `for` loops or `apply()` functions. As of this writing, the only vectorized **nhppp** function is `vdraw_sc_step_regular()`, which samples from an array of piecewise constant intensity functions provided that all constant-rate time intervals have equal length (are ‘regular’). It is the vectorized analogue of `draw_sc_step_regular()`.

We compared the execution speed of non-vectorized and vectorized functions for sampling 100 times from the piecewise constant ‘b’ majorizer (λ_{*b}) in Figure 1. When drawing only the earliest event, the vectorized function is approximately 51 times faster than the non-vectorized function ($1.15\mu s$ versus $59.01\mu s$ per sample drawn.) However, when drawing all events, the order changes: the vectorized function is approximately 2.5 times slower than the non-vectorized function ($1.23ms$ versus $0.50ms$ per set of event times drawn.) The reversal in apparent performance is because the vectorized function is not optimized for large numbers of events per draw. (The expected number of events with λ_{*b} in $(0, 6\pi]$ is 741.97.)

8. Summary and next developments

The **nhppp** is developed to aid the simulation of NHPPPs from general intensity or cumulative intensity functions. It contains specialized functions to ensure speed.

The next steps are a vectorized implementation of some functions, to allow fast simulation over vectors of arguments.

A C++ implementation is also forthcoming.

Computational details and credits

R 4.3.1 (R Core Team 2023) was used for all analyses. Packages **xtable** 1.8.4 (Dahl, Scott, Roosen, Magnusson, and Swinton 2019) and **knitr** 1.45 (Xie 2014) were used for automatic report generation. Packages **ggplot2** 3.4.4 (Wickham 2016), **ggridges** 0.5.5 (Wilke 2023), and **latex2exp** 0.9.6 (Meschiari 2022) were used for plot generation and L^AT_EX formatting. Packages **bench** 1.1.3 (Hester and Vaughan 2023), **rstream** 1.3.7 (Leydold 2022), **otinferrence** 0.1.0 (Sommerfeld 2017), and **parallel** 4.3.1 were used in the examples and the analyses. R itself and all aforementioned packages are available from the Comprehensive R Archive Network (CRAN) at <https://CRAN.R-project.org/>.

Acknowledgments

This work was funded from grant U01CA265750 from the National Cancer Institute. We thank the investigators of the Cancer Incidence and Surveillance Modeling Network (CISNET) Bladder Cancer Site Stavroula Chrysanthopoulou, Jonah Popp, Fernando Alarid-Escudero, Hawre Jalal, and David Garibay for useful discussions.

References

- Çinlar E (1975). “Introduction to stochastic processes Prentice-Hall.” *Englewood Cliffs, New Jersey* (420p).
- Cebrián AC (2020a). **IndTestPP**: *Tests of Independence and Analysis of Dependence Between Point Processes in Time*. R package version 3.0, URL <https://CRAN.R-project.org/package=IndTestPP>.
- Cebrián AC (2020b). **NHPoisson**: *Modelling and Validation of Non-Homogeneous Poisson Processes*. R package version 3.3, URL <https://CRAN.R-project.org/package=NHPoisson>.
- Cebrián AC, Abaurrea J, Asín J (2015). “**NHPoisson**: An R Package for Fitting and Validating Nonhomogeneous Poisson Processes.” *Journal of Statistical Software*, **64**(6), 125. doi: 10.18637/jss.v064.i06. URL <https://www.jstatsoft.org/index.php/jss/article/view/v064i06>.
- Cox D, Miller H (1965). “The Poisson Process.” In *The Theory of Stochastic Processes*, p. 147. Chapman and Hall.
- Dahl DB, Scott D, Roosen C, Magnusson A, Swinton J (2019). **xtable**: *Export Tables to L^AT_EX or HTML*. R package version 1.8.4, URL <https://CRAN.R-project.org/package=xtable>.
- Hammersley JM, Mauldon JG (1956). “General principles of antithetic variates.” In *Mathematical proceedings of the Cambridge philosophical society*, volume 52, pp. 476–481. Cambridge University Press.
- Hester J, Vaughan D (2023). **bench**: *High Precision Timing of R Expressions*. R package version 1.1.3, URL <https://CRAN.R-project.org/package=bench>.
- Lawson B, Leemis L, Kudlay V (2023). **simEd**: *Simulation Education*. R package version 2.0.1, URL <https://CRAN.R-project.org/package=simEd>.
- Lewis P, Shedler G (1976). “Simulation of nonhomogeneous Poisson processes with log linear rate function.” *Biometrika*, **63**(3), 501–505.
- Lewis PW, Shedler GS (1979). “Simulation of nonhomogeneous Poisson processes by thinning.” *Naval Research Logistics Quarterly*, **26**(3), 403–413.
- Leydold J (2022). **rstream**: *Streams of Random Numbers*. R package version 1.3.7, URL <https://CRAN.R-project.org/package=rstream>.
- Meschiari S (2022). **latex2exp**: *Use L^AT_EX Expressions in Plots*. R package version 0.9.6, URL <https://CRAN.R-project.org/package=latex2exp>.
- Press W, Teukolsky S, Vetterling W, Flannery B (2007). *Section 9.3. Van Wijngaarden-Dekker-Brent Method. Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, New York.
- R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

- Sommerfeld M (2017). **otinference**: *Inference for Optimal Transport*. R package version 0.1.0, URL <https://CRAN.R-project.org/package=otinference>.
- Sommerfeld M, Munk A (2018). “Inference for empirical Wasserstein distances on finite spaces.” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **80**(1), 219–238.
- Wang W, Fu H, Yan J (2022). **reda**: *Recurrent Event Data Analysis*. R package version 0.5.4, URL <https://github.com/wenjie2wang/reda>.
- Wickham H (2016). **ggplot2**: *Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4. URL <https://ggplot2.tidyverse.org>.
- Wilke CO (2023). **ggridges**: *Ridgeline Plots in ggplot2*. R package version 0.5.5, URL <https://CRAN.R-project.org/package=ggridges>.
- Wright R, Ramsay Jr T (1979). “On the effectiveness of common random numbers.” *Management Science*, **25**(7), 649–656.
- Xie Y (2014). “**knitr**: A Comprehensive Tool for Reproducible Research in **R**.” In V Stodden, F Leisch, RD Peng (eds.), *Implementing Reproducible Computational Research*. Chapman and Hall/CRC. ISBN 978-1466561595.

A. Piecewise constant majorizer functions

Let $\lambda(t)$ be a K -Lipschitz continuous intensity function, i.e., an intensity function where $|(\lambda(b) - \lambda(a))| \leq K|b - a|$, with K known. Algorithm 7 finds a piecewise constant majorizing function $\lambda_*(t)$. Starting from a partition of the time interval in time steps (not necessarily equal) it finds an upper bound for λ within the each partition.

If $\lambda(t)$ is monotonic, the least upper bound (supremum) is always found at the extremes of the interval and no knowledge of K is required.

The algorithm should be started with a good partitioning of the time interval. In practice, it is generally easy to specify equispaced intervals that are fine enough and impose little computational penalty for the application.

The **nhppp** function `get_step_majorizer()` implements Algorithm 7. Functions `ppp_t_step()` and `draw_intensity_step()` expect the majorizer function values as an argument.

```
get_step_majorizer(
  fun = abs, breaks = -5:5, is_monotone = FALSE,
  K = 1
)

## [1] 5.5 4.5 3.5 2.5 1.5 1.5 2.5 3.5 4.5 5.5
```

Algorithm 7 Pick a majorizing piecewise constant function $\lambda_*(t)$. Partition the interval and find an upper bound for $\lambda(t)$ in each partition.

Require:

$\lambda(t)$ is K -Lipschitz in $(a, b]$

Partition interval: $(a, b] = \bigcup_{m=1}^M (a_m, b_m]$ $\triangleright a = a_1, b_M = b, a_m = b_{m-1} \ (m > 1)$

1: $c \leftarrow K$ \triangleright Fastest possible slope

2: **if** $\lambda(t)$ is monotonic **then** \triangleright Then $\sup_{t \in (a_m, b_m]} (\lambda(t)) = \max(\lambda(a_m), \lambda(b_m))$

3: $c \leftarrow 0$

4: **end if**

5: **for** $m \in [M]$ **do**:

6: $\lambda_m^* \leftarrow \max(\lambda(a_m), \lambda(b_m)) + c(b_m - a_m)/2$ \triangleright Upper bound for $\lambda(t)$ in $(a_m, b_m]$

7: **end for**

8: $\lambda_*(t) \leftarrow \bigcup_{m=1}^M \{(a_m, b_m], \lambda_m^*\}$ \triangleright Piecewise constant map: $\lambda : (a_m, b_m] \mapsto \lambda_m$

9: **return** $\lambda_*(t)$

B. Conditional sampling from NHPPs

Algorithm 8 is a direct modification of the order statistics Algorithm 6 (Line 1 modified, in red font). To sample exactly m points, change line 1 of Algorithm 8 to

$$N \leftarrow m.$$

To sample up to k earliest points, replace line 11 with in Algorithm 8 with

$$\mathbf{return} \{Z_{(i)} \mid i \leq k, Z_{(i)} \in \mathcal{Z}\}.$$

Algorithm 8 Modified order statistics algorithm for sampling at least m events from an NHPP given $\Lambda(t), \Lambda^{-1}(z)$.

Require: $\Lambda(t), \Lambda^{-1}(z), t \in (a, b]$ $\triangleright \Lambda^{-1}(z)$ possibly numerically
1: $N \leftarrow N \sim \text{TruncatedPoisson}_{N \geq m}(\Lambda(b) - \Lambda(a))$ $\triangleright (m-1)\text{-truncated Poisson}$
2: $t \leftarrow a$
3: $\mathcal{Z} \leftarrow \emptyset$ $\triangleright \mathcal{Z}$ is an ordered set
4: **if** $N > 0$ **then**
5: **for** $i \in [N]$ **do**:
6: $U_i \leftarrow U_i \sim \text{Uniform}(0, 1)$ \triangleright Generate order statistics
7: $\mathcal{Z} \leftarrow \mathcal{Z} \cup \{\Lambda^{-1}(\Lambda(a) + U_i(\Lambda(b) - \Lambda(a)))\}$
8: **end for**
9: $\mathcal{Z} \leftarrow \text{sort}(\mathcal{Z})$
10: **end if**
11: **return** \mathcal{Z} \triangleright Up to k earliest points: **return** $\{Z_{(i)} \mid i \leq k, Z_{(i)} \in \mathcal{Z}\}$

Affiliation:

TA Trikalinos
Department of Health Services, Policy & Practice
and
Department of Biostatistics
School of Public Health
Brown University
RI 02912, USA
E-mail: thomas_trikalinos@brown.edu