

Jump-point PLP (JPLP) simulation

Miao Cai miao.cai@slu.edu

2020-04-08

Contents

1	Power law process (PLP)	1
1.1	PLP intensity function	1
1.2	PLP simulation	1
2	Jump Power Law Process (JPLP)	4
2.1	JPLP intensity function	4
2.2	JPLP mean function	4
2.3	Simulation JPLP events in one shift	7
2.4	Simulation JPLP events in multiple shifts	12

1 Power law process (PLP)

1.1 PLP intensity function

Power law process (PLP): When the intensity function of a NHPP is:

$$\lambda(t) = \frac{\beta}{\theta} \left(\frac{t}{\theta} \right)^{\beta-1} = \beta \theta^{-\beta} t^{\beta},$$

where $\beta > 0$ and $\theta > 0$, the process is called the power law process (PLP). The mean function $\Lambda(t)$ is the integral of the intensity function:

$$\Lambda(t) = \int_0^t \lambda(t) dt = \int_0^t \frac{\beta}{\theta} \left(\frac{t}{\theta} \right)^{\beta-1} = \left(\frac{t}{\theta} \right)^{\beta}$$

1.2 PLP simulation

```
# simulating PLP - time truncated case
sim_plp_tau = function(tau = 30,
                        beta = 1.5,
                        theta = 10){
  # initialization
  s = 0; t = 0
  while (max(t) <= tau) {
    u <- runif(1)
    s <- s - log(u)
```

```

    t_new <- theta*s^(1/beta)
    t <- c(t, t_new)
  }
  t = t[c(-1, -length(t))]

  return(t)
}

# simulate multiple NHPPs - time truncated case
sim_mul_plp_tau = function(n_shift = 20,
                           shift_len_mean = 20, shift_len_sd = 5,
                           theta = 10, beta = 2, mean_n = 5){
  tau_vector = rnorm(n_shift, shift_len_mean, shift_len_sd) #difference1

  t_list = list()
  len_list = list()
  # end_time1 = list() # not needed for time truncated case

  for (i in 1:n_shift) {
    t_list[[i]] = sim_plp_tau(tau_vector[i], beta, theta)
    len_list[[i]] = length(t_list[[i]])
  }

  event_dat = data.frame(
    shift_id = rep(1:n_shift, unlist(len_list)),
    event_time = Reduce(c, t_list)
  )

  start_end_dat = data.frame(
    shift_id = 1:n_shift,
    start_time = rep(0, n_shift),
    end_time = tau_vector #difference2
  )

  return(list(event_dat = event_dat,
              start_end_dat = start_end_dat,
              shift_length = unlist(len_list)))
}

sim_hier_plp_tau = function(N, beta = 1.5, theta){
  t_list = list()
  len_list = list()
  tau_vector = rnorm(N, 10, 1.3)

  for (i in 1:N) {

```

```

    t_list[[i]] = sim_plp_tau(tau_vector[i], beta = beta, theta = theta[i])
    len_list[[i]] = length(t_list[[i]])
  }

  event_dat = data.frame(
    shift_id = rep(1:N, unlist(len_list)),
    event_time = Reduce(c, t_list)
  )

  start_end_dat = data.frame(
    shift_id = 1:N,
    start_time = rep(0, N),
    end_time = tau_vector #difference2
  )

  return(list(event_dat = event_dat,
             start_end_dat = start_end_dat,
             shift_length = unlist(len_list)))
}

```

```

plot_est = function(data, var = "beta", hline_var = 1.5){
  p = data %>%
    filter(term == var) %>%
    ggplot(aes(id, est_mean)) +
    geom_point() +
    geom_line(linetype = "dashed", color = "red")+
    geom_errorbar(aes(ymax = est_mean + 1.96*est_sd,
                     ymin = est_mean - 1.96*est_sd),
                 width = 1)+
    geom_segment(aes(x = 10, xend = 100,
                     y = hline_var, yend = hline_var),
                 color = "green")+
    scale_x_continuous(breaks = c(0, 10, 25, 50, 75, 100),
                      labels = c("0", "10", "25", "50", "75", "100")) +
    labs(x = "The number of drivers (random effects)",
         y = var) +
    theme_bw()
  return(p)
}

```

```

# plot events
plot_events = function(event_dat, start_end_dat, cross_size = 2){
  p = event_dat %>%
    ggplot(aes(x = event_time, y = shift_id)) +
    geom_point(alpha = 0.8, shape = 4, color = 'red', size = cross_size) +
    scale_y_continuous("shift ID",
                      labels = as.character(start_end_dat$shift_id),

```

```

        breaks = start_end_dat$shift_id)+
xlab('Time to event (minutes)') +
geom_segment(data = start_end_dat,
             aes(x = start_time, xend = end_time,
                 y = shift_id, yend = shift_id),
             lineend = 'butt',
             arrow = arrow(length = unit(0.2, "cm")) +
theme_classic()
return(p)
}

```

2 Jump Power Law Process (JPLP)

2.1 JPLP intensity function

A Bayesian hierarchical JPLP has the following intensity function:

$$\lambda_{\text{JPLP}}(t|d, s, r, \beta, \gamma_{0,d}, \gamma, \mathbf{X}_d, \mathbf{W}) = \begin{cases} \kappa^0 \lambda(t|\beta, \gamma_{0,d}, \gamma, \mathbf{X}_d, \mathbf{W}) & 0 \leq t \leq a_{d,s,1} \\ \kappa^1 \lambda(t|\beta, \gamma_{0,d}, \gamma, \mathbf{X}_d, \mathbf{W}) & a_{d,s,1} \leq t \leq a_{d,s,2} \\ \dots & \dots \\ \kappa^{R-1} \lambda(t|\beta, \gamma_{0,d}, \gamma, \mathbf{X}_d, \mathbf{W}) & a_{d,s,R-1} \leq t \leq a_{d,s,R} \end{cases} \quad (1)$$

$$= \kappa^{r-1} \lambda(t|d, s, r, \kappa, \beta, \gamma_{0,d}, \gamma, \mathbf{X}_d, \mathbf{W}) \quad a_{d,s,r-1} \leq t \leq a_{d,s,r},$$

where the introduced parameter κ is the percent of intensity function recovery once the driver takes a break. We assume that this κ is constant across drivers and shifts.

2.2 JPLP mean function

Let T_1, T_2, \dots be random variables representing the event times of a nonhomogeneous Poisson process with continuous expectation function $\Lambda(t)$, and let N_t represent the total number of events occurring before time t in the process. Then, conditional on the number of events $N_{t_0} = n$, the event times T_1, T_2, \dots, T_n are distributed as order statistics from a sample with distribution function $F(t) = \Lambda(t)/\Lambda(t_0)$ for $t \in [0, t_0]$.

This is a generalization of the result for homogeneous Poisson processes. It naturally gives rise to the following algorithm for generating random variates from a nonhomogeneous Poisson process with expectation function $\Lambda(t)$ in a fixed interval $[0, t_0]$.

- (1) Generate $n \sim \text{Poisson}(\Lambda(t_0))$.
- (2) Independently generate n random variates t'_1, t'_2, \dots, t'_n from the cdf $F(t) = \Lambda(t)/\Lambda(t_0)$.
- (3) Order t'_1, t'_2, \dots, t'_n to obtain $t_1 = t'_{(1)}, t_2 = t'_{(2)}, \dots, t_n = t'_{(n)}$.
- (4) Deliver t_1, t_2, \dots, t_n .

2.2.1 Mean function $\Lambda(t)$

```

# Mean function Lambda for JPLP
L_plp = function(t, beta = 1.5, theta = 4) return((t/theta)^beta)
Lambda = function(t,

```

```

        tau = 12,
        kappa = 0.8,
        t_trip = c(3.5, 6.2, 9),
        beta = 1.5,
        theta = 4)
{
  t_trip1 = c(0, t_trip)
  n_trip = length(t_trip1)
  comp = L_plp(t_trip, beta, theta)
  kappa_vec0 = rep(kappa, n_trip - 1)^(0:(n_trip - 2))
  kappa_vec1 = rep(kappa, n_trip - 1)^(1:(n_trip - 1))
  cum_comp0 = comp*kappa_vec0
  cum_comp1 = comp*kappa_vec1
  index_trip = max(cumsum(t > t_trip1)) - 1

  if(index_trip == 0){
    return((t/theta)^beta)
  }else{
    return(sum(cum_comp0[1:index_trip]) - sum(cum_comp1[1:index_trip]) +
           kappa^index_trip*(t/theta)^beta)
  }
}

```

2.2.2 Test the mean function $\Lambda(t)$

```

# test Lambda
kappa = 0.8
t_trip = c(3.5, 6.2, 9)
beta = 1.5
theta = 4

Lambda(3.1)

## [1] 0.6822642

L_plp(3.1)

## [1] 0.6822642

Lambda(4.1)

## [1] 0.9938842
kappa^0*L_plp(t_trip[1]) +
  kappa^1*L_plp(4.1) - kappa^1*L_plp(t_trip[1])

## [1] 0.9938842

Lambda(8.9)

## [1] 2.596555

```

```
kappa^0*L_plp(t_trip[1]) +
  kappa^1*L_plp(t_trip[2]) - kappa^1*L_plp(t_trip[1]) +
  kappa^2*L_plp(8.9) - kappa^2*L_plp(t_trip[2])
```

```
## [1] 2.596555
```

```
Lambda(12)
```

```
## [1] 3.564885
```

```
kappa^0*L_plp(t_trip[1]) +
  kappa^1*L_plp(t_trip[2]) - kappa^1*L_plp(t_trip[1]) +
  kappa^2*L_plp(t_trip[3]) - kappa^2*L_plp(t_trip[2]) +
  kappa^3*L_plp(12) - kappa^3*L_plp(t_trip[3])
```

```
## [1] 3.564885
```

2.2.3 Plot the mean function of PLP and JPLP

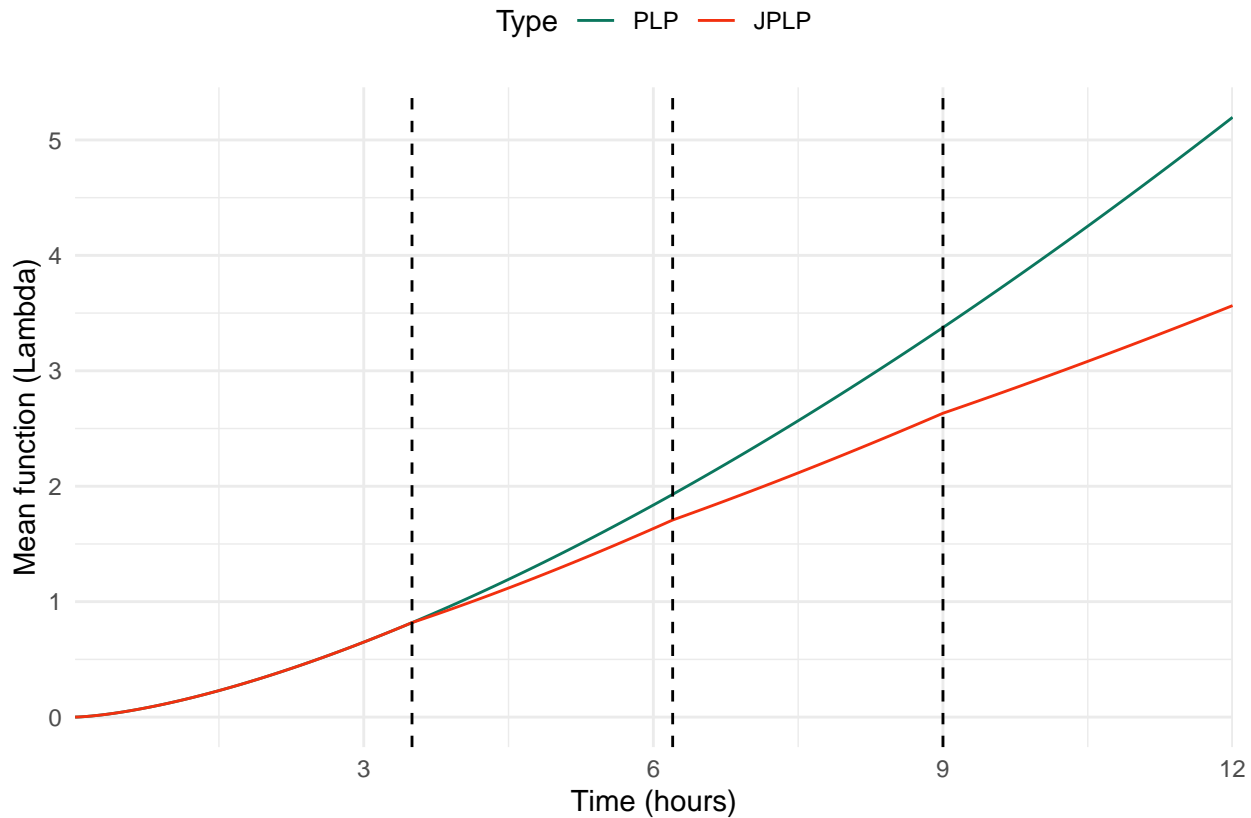
```
pacman::p_load(ggplot2, dplyr, tidyr)

t_trip = c(3.5, 6.2, 9)
beta = 1.5
theta = 4

x = seq(0.01, 12, 0.01)
y0 = (x/theta)^beta
y1 = rep(NA_real_, length(x))

for (i in 1:length(x)) {
  y1[i] = Lambda(t = x[i], kappa = 0.8)
}

data.frame(x, y0, y1) %>%
  tidyr::pivot_longer(cols = c('y0', 'y1'),
    names_to = "Type",
    values_to = 'y') %>%
  mutate(Type = case_when(Type == 'y0' ~ 'PLP',
    Type == 'y1' ~ 'JPLP')) %>%
  mutate(Type = factor(Type, levels = c('PLP', 'JPLP'))) %>%
  ggplot(aes(x = x, y = y, group = Type, color = Type)) +
  scale_color_manual(values = c(PLP = "#0B775E", JPLP = "#F2300F")) +
  geom_line() +
  geom_vline(xintercept = t_trip, linetype = "dashed") +
  labs(x = 'Time (hours)',
    y = 'Mean function (Lambda)') +
  theme_minimal() +
  theme(legend.position = "top") +
  scale_x_continuous(expand = c(0, 0))
```



2.3 Simulation JPLP events in one shift

2.3.1 Simulate JPLP events in one shift

```
# Define a inverse function for mean function Lambda
inverse = function (f, lower = 0.0001, upper = 100) {
  function (y) uniroot((function (x) f(x) - y), lower = lower, upper = upper)[1]
}
```

```
# Test the inverse function
```

```
Lambda1 = function(t, tau = 12, kappa = 0.8, t_trip = c(3.5, 6.2, 9),
  beta = 1.5, theta = 4){
  return(Lambda(t, tau = tau, kappa = kappa, t_trip = t_trip,
    beta = beta, theta = theta))
}
```

```
inv_Lambda = inverse(Lambda1, 0.0001, 100)
inv_Lambda(3.564885)$root
```

```
## [1] 12
```

```
# simulating JPLP - time truncated case
```

```
sim_jplp = function(tau0 = 12,
  kappa0 = 0.8,
  t_trip0 = c(3.5, 6.2, 9),
  beta0 = 1.5,
```

```

        theta0 = 0.5)
{ s = 0; t = 0
  Lambda1 = function(t, tau1 = tau0, kappa1 = kappa0, t_trip1 = t_trip0,
                     beta1 = beta0, theta1 = theta0){
    return(Lambda(t, tau = tau1, kappa = kappa1, t_trip = t_trip1,
                 beta = beta1, theta = theta1))
  }
  inv_Lambda = inverse(Lambda1, 0.0001, 100)

  while (max(t) <= tau0) {
    u <- runif(1)
    s <- s - log(u)
    t_new <- inv_Lambda(s)$root
    t <- c(t, t_new)
  }
  t = t[c(-1, -length(t))]

  return(t)
}

sim_jplp(theta0 = 1)

## [1] 1.116532 1.378210 1.632287 2.591005 3.186471 3.807230 3.826825
## [8] 3.975958 4.064756 4.320142 5.280300 5.509483 5.867686 6.463746
## [15] 7.115613 7.196707 7.309012 7.456360 7.616563 8.032883 8.359083
## [22] 8.731068 8.839096 8.908454 9.696974 10.209258 10.525626 10.942322
## [29] 11.548983 11.791985 11.984025

sim_jplp(theta0 = 2)

## [1] 0.7915436 1.4447692 1.8000240 2.2861354 2.2919278 2.9363789
## [7] 2.9710548 4.3172914 5.5247106 10.2626141 11.6184992

```

2.3.2 Plot JPLP events in one shift

```

set.seed(123)

tauX = 12
kappaX = 0.8
t_tripX = c(3.5, 6.2, 9)
betaX = 1.5
thetaX = 2

t_events = sim_jplp(tau0 = tauX,
                    kappa0 = kappaX,
                    t_trip0 = t_tripX,
                    beta0 = betaX,
                    theta0 = thetaX)

```



```

x = seq(0.01, 12, 0.01)
y0 = (x/theta)^beta
y1 = rep(NA_real_, length(x))

for (i in 1:length(x)) {
  y1[i] = Lambda(t = x[i], kappa = 0.8)
}

dLambda = data.frame(x, y0, y1) %>%
  tidyr::pivot_longer(cols = c('y0', 'y1'),
    names_to = "Type",
    values_to = 'y') %>%
  mutate(Type = case_when(Type == 'y0' ~ 'PLP',
    Type == 'y1' ~ 'JPLP')) %>%
  mutate(Type = factor(Type, levels = c('PLP', 'JPLP')))

d_event = data.frame(t_events = t_events, y = 0)
d_shift = data.frame(start_x = 0,
  end_x = tauX,
  start_y = 0,
  end_y = 0)

ggplot(data = dLambda) +
  geom_line(aes(x = x, y = y, group = Type, color = Type)) +
  scale_color_manual(values = c(PLP = "#0B775E", JPLP = "#F2300F")) +
  geom_vline(xintercept = t_trip, linetype = "dashed") +
  labs(x = 'Time (hours)',
    y = 'Mean function (Lambda)') +
  theme_minimal() +
  theme(legend.position = "top") +
  geom_point(data = d_event, aes(x = t_events, y = y),
    alpha = 0.8, shape = 4, color = 'red', size = 2) +
  geom_segment(data = d_shift,
    aes(x = start_x, xend = end_x,
      y = start_y, yend = end_y),
    lineend = 'butt',
    arrow = arrow(length = unit(0.2, "cm")))) +
  scale_x_continuous(expand = c(0, 0))

set.seed(666)

tauX = 12
kappaX = 0.8
t_tripX = c(3.5, 6.2, 9)
betaX = 1.5
thetaX = 1

```

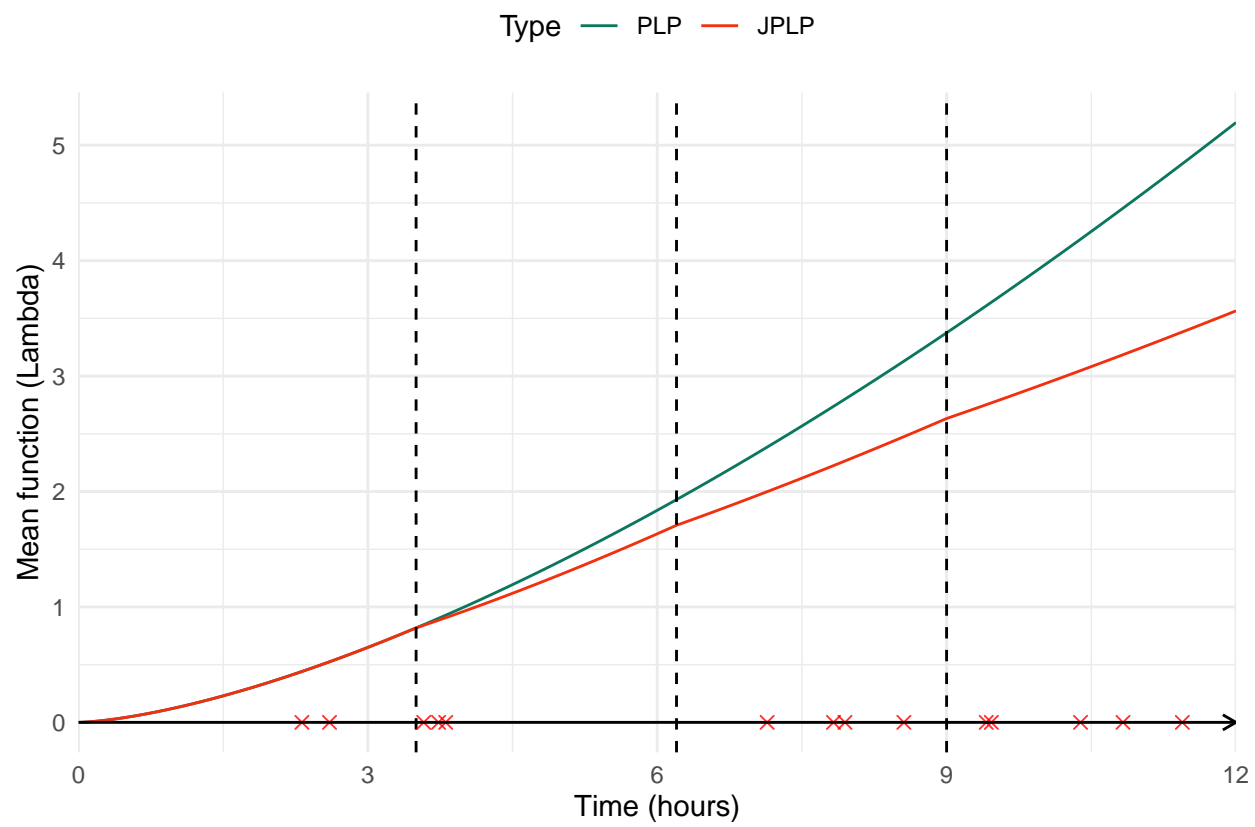


Figure 1: Event times (red cross) and mean function Λ (curves)

```

t_events = sim_jplp(tau0 = tauX,
                    kappa0 = kappaX,
                    t_trip0 = t_tripX,
                    beta0 = betaX,
                    theta0 = thetaX)

lambda_plp = function(t, beta, theta) return(beta*theta^(-beta)*t^(beta - 1))
lambda_jplp = function(t, beta, theta, kappa = 0.8, t_trip){
  index_trip = max(cumsum(t > t_trip))
  return((kappa^index_trip)*beta*theta^(-beta)*t^(beta - 1))
}

x = seq(0.01, 12, 0.01)
y_plp = lambda_plp(x, beta = betaX, theta = thetaX)
y_jplp = rep(NA_real_, length(lambda_plp))

for (i in 1:length(x)) {
  y_jplp[i] = lambda_jplp(x[i],
                          beta = betaX, theta = thetaX,
                          kappa = 0.8, t_trip = t_tripX)
}

dlambda = tibble::tibble(x, y_plp, y_jplp) %>%
  tidyr::pivot_longer(cols = c('y_plp', 'y_jplp'),
                      names_to = "Type",
                      values_to = 'y') %>%
  mutate(Type = case_when(Type == 'y_plp' ~ 'PLP',
                          Type == 'y_jplp' ~ 'JPLP')) %>%
  mutate(Type = factor(Type, levels = c('PLP', 'JPLP')))

d_event = data.frame(t_events = t_events, y = 0)
d_shift = data.frame(start_x = 0,
                      end_x = tauX,
                      start_y = 0,
                      end_y = 0)

ggplot(data = dlambda) +
  geom_line(aes(x = x, y = y, group = Type, color = Type)) +
  scale_color_manual(values = c(PLP = "#0B775E", JPLP = "#F2300F")) +
  geom_vline(xintercept = t_trip, linetype = "dashed") +
  labs(x = 'Time (hours)',
       y = 'Mean function (Lambda)') +
  theme_minimal() +
  theme(legend.position = "top") +
  geom_point(data = d_event, aes(x = t_events, y = y),

```

```

alpha = 0.8, shape = 4, color = 'red', size = 2) +
geom_segment(data = d_shift,
  aes(x = start_x, xend = end_x,
    y = start_y, yend = end_y),
  lineend = 'butt',
  arrow = arrow(length = unit(0.2, "cm")))) +
scale_x_continuous(expand = c(0, 0))

```

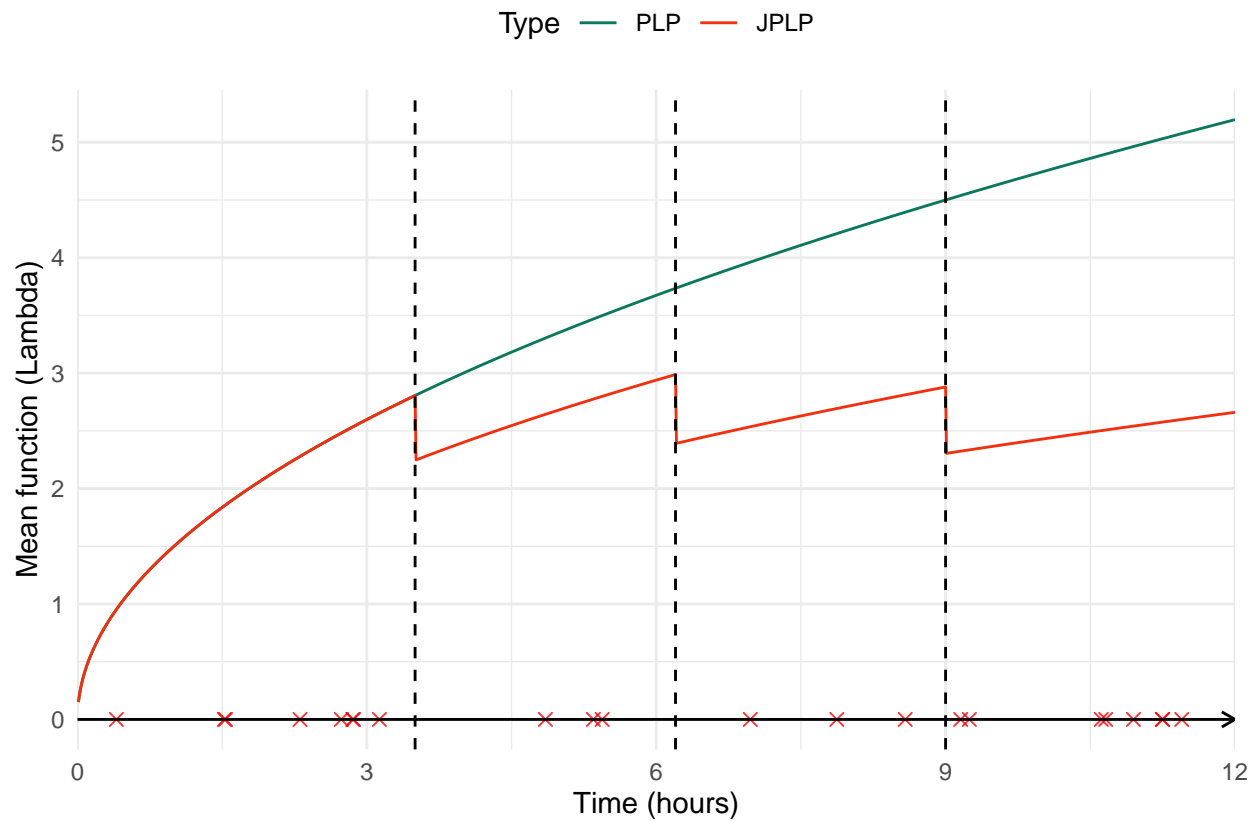


Figure 2: Event times (red cross) and intensity function lambda (curves)

2.4 Simulation JPLP events in multiple shifts

```

# Function to get the number of stops
get_n_stop = function() sample(1:4, 1, TRUE)

# Function to simulate event time for multiple shifts
sim_mul_jplp = function(kappa = 0.8, beta = 1.5, theta = 2, n_shift = 10)
{
  t_shift_vec = list()
  n_stop_vec = list()
  t_stop_vec = list()
  n_event_vec = list()
  t_event_vec = list()

```

```

for (i in 1:n_shift) {
  sim_tau = round(runif(1, min = 6, max = 12), 1)
  n_stop = get_n_stop()
  sim_t_trip = runif(n_stop, min = 1, max = sim_tau - 1)
  t_events = sim_jplp(tau0 = sim_tau,
                      kappa0 = kappa,
                      t_trip0 = sim_t_trip,
                      beta0 = beta,
                      theta0 = theta)

  t_shift_vec[[i]] = sim_tau
  n_stop_vec[[i]] = n_stop
  t_stop_vec[[i]] = sim_t_trip
  n_event_vec[[i]] = length(t_events)
  t_event_vec[[i]] = t_events
}

event_dt = data.frame(
  shift_id = rep(1:n_shift, unlist(n_event_vec)),
  event_time = Reduce(c, t_event_vec)
)

trip_dt = data.frame(
  shift_id = rep(1:n_shift, unlist(n_stop_vec)),
  trip_time = Reduce(c, t_stop_vec)
)

shift_dt = data.frame(
  shift_id = 1:n_shift,
  start_time = rep(0, n_shift),
  end_time = Reduce(c, t_shift_vec)
)

return(list(event_time = event_dt,
            trip_time = trip_dt,
            shift_time = shift_dt))
}

z = sim_mul_jplp()
str(z)

## List of 3
## $ event_time:'data.frame': 78 obs. of 2 variables:
## ..$ shift_id : int [1:78] 1 1 1 1 1 1 2 2 2 2 ...
## ..$ event_time: num [1:78] 0.586 1.778 3.557 3.876 5.374 ...
## $ trip_time :'data.frame': 24 obs. of 2 variables:
## ..$ shift_id : int [1:24] 1 1 1 1 2 2 3 3 3 4 ...

```

```
## ..$ trip_time: num [1:24] 2.28 1.2 3.94 3.66 5.47 ...
## $ shift_time:'data.frame': 10 obs. of 3 variables:
## ..$ shift_id : int [1:10] 1 2 3 4 5 6 7 8 9 10
## ..$ start_time: num [1:10] 0 0 0 0 0 0 0 0 0 0
## ..$ end_time : num [1:10] 6.8 7.7 11.9 10.6 11.9 10.4 8.4 9.8 8.3 6.5
```