Correction Tutorial 4

1. A bit of statistics with Poisson Processes.

   (a) The random variable N_T^a obeys a Poisson distribution with parameter $\int_0^T \lambda_a(s)ds$. Similarly, N_T^b obeys a Poisson distribution with parameter $\int_0^T \lambda_b(s)ds$. Since those Poisson processes are independent, N_T^a and N_T^b are also independent.

   (b) It follows from question 1.a) that for all integer $k, l \geq 0$,

$$\begin{aligned}\mathbb{P}(N_T^a = k \text{ and } N_T^b = l) &= \mathbb{P}(N_T^a = k) \times \mathbb{P}(N_T^b = l) \\ &= \frac{\big(\int_0^T \lambda_a(s)ds\big)^k}{k!} e^{-\int_0^T \lambda_a(s)ds} \times \frac{\big(\int_0^T \lambda_b(s)ds\big)^l}{l!} e^{-\int_0^T \lambda_b(s)ds} \\ &= \frac{\big(\int_0^T \lambda_a(s)ds\big)^k}{k!} \times \frac{\big(\int_0^T \lambda_b(s)ds\big)^l}{l!} e^{-\int_0^T \big[\lambda_a(s)+\lambda_b(s)\big]ds}.\end{aligned}$$

   (c) For any integer $n \geq 0$, we have

$$\begin{aligned}\mathbb{P}(N_T = n) &= \sum_{k=0}^{n} \mathbb{P}(N_T^a = k \text{ and } N_T^b = n-k) \\ &= \sum_{k=0}^{n} \frac{\big(\int_0^T \lambda_a(s)ds\big)^k}{k!} \times \frac{\big(\int_0^T \lambda_b(s)ds\big)^{n-k}}{(n-k)!} e^{-\int_0^T \big[\lambda_a(s)+\lambda_b(s)\big]ds} \\ &= \frac{e^{-\int_0^T \big[\lambda_a(s)+\lambda_b(s)\big]ds}}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\int_0^T \lambda_a(s)ds\right)^k \times \left(\int_0^T \lambda_b(s)ds\right)^{n-k}\end{aligned}$$

   and we conclude thanks to the Binome formula that

$$\mathbb{P}(N_T = n) = \frac{\big(\int_0^T \big[\lambda_a(s) + \lambda_b(s)\big]ds\big)^n}{n!} e^{-\int_0^T \big[\lambda_a(s)+\lambda_b(s)\big]ds}.$$

   Then the variable N_T obeys a Poisson distribution with parameter $\int_0^T \big[\lambda_a(s) + \lambda_b(s)\big]ds$.

   (d) For all integers $n, k \geq 0$ such that $k \leq n$, we have

$$\begin{aligned}\mathbb{P}(N_T^a = k \,|\, N_T = n) &= \frac{\mathbb{P}(N_T^a = k \text{ and } N_T = n)}{\mathbb{P}(N_T = n)} \\ &= \frac{\mathbb{P}(N_T^a = k \text{ and } N_T^b = n-k)}{\mathbb{P}(N_T = n)} \\ &= \frac{\frac{(\int_0^T \lambda_a(s)ds)^k (\int_0^T \lambda_b(s)ds)^{n-k}}{k!(n-k)!} e^{-\int_0^T [\lambda_a(s)+\lambda_b(s)]ds}}{\frac{(\int_0^T [\lambda_a(s)+\lambda_b(s)]ds)^n}{n!} e^{-\int_0^T [\lambda_a(s)+\lambda_b(s)]ds}} \\ &= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}\end{aligned}$$

with

$$p = \frac{\int_0^T \lambda_a(s)ds}{\int_0^T [\lambda_a(s) + \lambda_b(s)]ds}.$$

Then, conditionally to N_T, the variable N_T^a obeys a Binomial distribution with parameters p and N_T.

(e) Under the null hypothesis $(\mathbf{H_0}) : \lambda_a(.) = \lambda_b(.)$, we have $p = 1/2$ and then $N_T^a \sim \mathcal{B}(1/2, N_T)$ conditionally to N_T. Under $(\mathbf{H_1}) : \lambda_a(.) \geq \lambda_b(.)$ the probability p is larger than 1/2 and N_T^a would be larger.
So we reject $(\mathbf{H_0})$ when $N_T^a \geq c$ where c is the $1-\alpha$ quantile of $\mathcal{B}(1/2, N_T)$. Notice that this value is random and depends on N_T. More precisely, we want c such that

$$\mathbb{P}(N_T^a \geq c \,|\, N_T) \leq \alpha,$$

that is

$$1 - F_{\mathcal{B}(1/2,N_T)}(c-1) \leq \alpha$$

where $F_{\mathcal{B}(1/2,N_T)}$ is the c.d.f of $\mathcal{B}(1/2, N_T)$. Indeed, we have

$$\begin{aligned}\mathbb{P}(N_T^a \geq c \,|\, N_T) &= 1 - \mathbb{P}(N_T^a < c \,|\, N_T) \\ &= 1 - \mathbb{P}(N_T^a \leq c-1 \,|\, N_T) \\ &= 1 - \mathbb{P}(N_T^a \leq c-1 \,|\, N_T)\end{aligned}$$

and then the p-value is given by $1 - F_{\mathcal{B}(1/2,N_T)}(N_T^a - 1)$.

(f) See the R code.

2. Simulation

(a) The intensity $h(t) = (1-t^2)\mathbb{1}_{\{t\in[0,1]\}}$ is bounded by 1. Then in the thinning strategy, one should simulate $\mathcal{E}(1)$ variables until their cumulative sum is larger than 1. This is the T_i's. We take the times that are smaller than 1. For them, we simulate the U_i's and compare with h to accept or reject it.

Generation of T_i

$t = 0$

While $t < 1$

  $\tau = \mathcal{E}(1)$

  $T_i = t + \tau$

  Append T_i to the set of points

  $t = T_i$

Remove the last point.

Generation of U_i For each T_i, simulate $U_i \sim \mathcal{U}([0,1])$ and accept the point T_i if and only if $U_i < h(T_i)$. Return the accepted points.

This algorithm corresponds to the function `MyPoisson()`.

(b) The Poisson process should produce in average

$$\int_0^1 (1-t^2)dt = \left[t - \frac{t^3}{3}\right]_{t=0}^{t=1} = \frac{2}{3} \text{ points}$$

(see the R code).

(c) Let us simulate parents and children on $[0, T_{max}]$:

Algorithm [see `Successive()`] ($T_{\max}$ has to be a parameter in entry as well as M)

- Generate a Poisson process with rate M for the parents on $[0, T_{max}]$. Let us denote by $T_{p_1}, ..., T_{p_n}$ those parents.
- For each of them, generate children thanks to the function `MyPoisson()`
- Remove the points that they are larger than T_{max}.

(d) In fact, it is a Hawkes process with two coordinates :

(i) Parents (mark 1) have intensity M, then $\nu_1 = M$ and $h_{1\to1} = h_{2\to1} = 0$.

(ii) Children (mark 2) have intensity $\int_0^t h(t-s)dN_s$, then $\nu_2 = 0$, $h_{1\to2} = h$ and $h_{2\to2} = 0$.

We want to simulate a Poisson process with intensity

$$\sum_{T_p} h(t - T_p) + M.$$

Since the support of h is $[0, 1]$, this quantity is bounded by

$$M' = M + N_{[T-1,T]}^{parents} \tag{1}$$

at time T, where $N_{[T-1,T]}^{parents}$ is the number of parents in $[T-1, T]$.

Algorithm [see `MyThinning()`]

- At time 0, the intensity is M. We then make a time step of length $\tau = \mathcal{E}(M)$ and the next time τ is always accepted as a parent.
- At time T, we make a time step of length $\tau' = \mathcal{E}(M')$ where M' is given by (1), and set $T_{next} = T + \tau'$. We also draw a mark $\mathcal{U}([0, M'])$:
  - If the mark is lower than M, we accept T as a parent.
  - If the mark is in the interval

    $$\left[M,\ M + \sum_{\substack{T_p \\ T_{next}-1 \le T_p < T_{next}}} h(T_{next} - T_p)\right[,$$

    then we accept T as a child.
  - Otherwise, T is rejected.

  We perform that until T is larger than T_{max}.