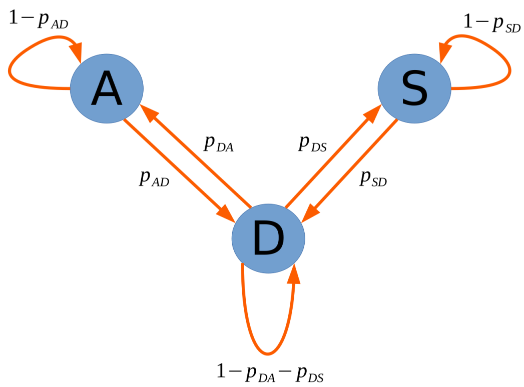
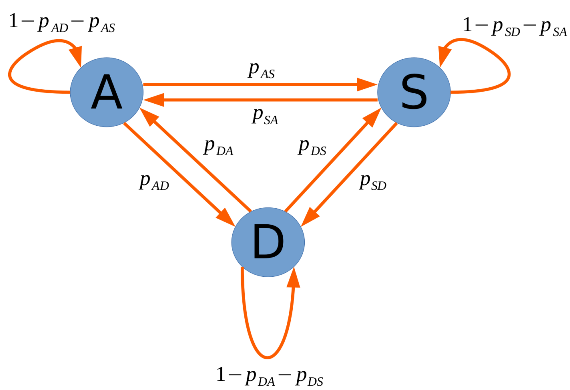
# Correction Tutorial 5

1. (a) Let us denote the different states by :
   $A$ for "Fully Awake",
   $D$ for "Drowsy",
   $S$ for "Sleeping".
   Each model need an initialization. The initial probabilities are then denoted $p_A$ for the state $A$, $p_D$ for the state $D$ and $p_S = 1 - p_A - p_D$ for the state $S$.

   **Model 1**




   **Model 2**




   (b) Take $U \sim \mathcal{U}([0,1])$, and then

   **For the case** $A$

   - if $U \in [0, p_{AD}[$, you go to $D$.
   - otherwise, you stay in $A$.

**For the case $S$**

- if $U \in [0, p_{SD}[$, you go to $D$.
- otherwise, you stay in $S$.

**For the case $D$**

- if $U \in [0, p_{DS}[$, you go to $S$.
- if $U \in [p_{DS}, p_{DS} + p_{DA}[$, you go to $A$.
- otherwise, you stay in $D$.

(See the R code).

(c) The likelihood of a sequence $x_1, ..., x_n$ is

$$\mathbb{P}(X_0 = x_0, ..., X_n = x_n) = \mathbb{P}(X_n = x_n | \text{ Past})\mathbb{P}(X_0 = x_0, ..., X_{n-1} = x_{n-1})$$

and thanks to the Markov property,

$$\mathbb{P}(X_0 = x_0, ..., X_n = x_n) = \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1})\mathbb{P}(X_0 = x_0, ..., X_{n-1} = x_{n-1}).$$

By recursion, we deduce that

$$\begin{aligned}
\mathbb{P}(X_0 = x_0, ..., X_n = x_n) &= p_{x_n x_{n-1}} \times ... \times p_{x_1 x_0} \times p_{x_0} \\
&= p_A^{1_{x_0=A}} \times p_D^{1_{x_0=D}} \times \left(1 - p_A - p_D\right)^{1_{x_0=S}} \\
&\quad \times p_{AD}^{N_{A\to D}} \times \left(1 - p_{AD}\right)^{N_{A\to A}} \\
&\quad \times p_{DS}^{N_{D\to S}} \times p_{DA}^{N_{D\to A}} \times \left(1 - p_{DS} - p_{DA}\right)^{N_{D\to D}} \\
&\quad \times p_{SD}^{N_{S\to D}} \times \left(1 - p_{SD}\right)^{N_{S\to S}}
\end{aligned}$$

where $N_{i\to j}$ is the number of times that you have seen the transition $i \to j$ in the sequence $x_0, ..., x_n$. So the log-likelihood is, with $p = (p_{AD}, p_{DS}, p_{DA}, p_{SD})$,

$$\begin{aligned}
\ell_{p_A, p_D, p} = &\, 1_{x_0=A} \log(p_A) + 1_{x_0=D} \log(p_D) + 1_{x_0=S} \log\left(1 - p_A - p_D\right) \\
&+ N_{A\to D} \log(p_{AD}) + N_{A\to A} \log\left(1 - p_{AD}\right) \\
&+ N_{D\to S} \log(p_{DS}) + N_{D\to A} \log(p_{DA}) + N_{D\to D} \log\left(1 - p_{DS} - p_{DA}\right) \\
&+ N_{S\to D} \log(p_{SD}) + N_{S\to S} \log\left(1 - p_{SD}\right)
\end{aligned}$$

The maximum likelihood estimator $\hat{p} = (\hat{p}_{AD}, \hat{p}_{DS}, \hat{p}_{DA}, \hat{p}_{SD})$ staisfies

$$\frac{\partial \ell_{p_A, p_D, p}}{\partial p_{DS}}(\hat{p}) = \frac{\partial \ell_{p_A, p_D, p}}{\partial p_{DA}}(\hat{p}) = 0,$$

that is

$$\frac{N_{D\to S}}{\hat{p}_{DS}} - \frac{N_{D\to D}}{1 - \hat{p}_{DS} - \hat{p}_{DA}} = 0 \quad \text{and} \quad \frac{N_{D\to A}}{\hat{p}_{DA}} - \frac{N_{D\to D}}{1 - \hat{p}_{DS} - \hat{p}_{DA}} = 0$$

which is equivalent to the system

$$\begin{cases} \hat{p}_{DA} N_{D\to D} = (1 - \hat{p}_{DS} - \hat{p}_{DA}) N_{D\to A} \\ \hat{p}_{DS} N_{D\to D} = (1 - \hat{p}_{DS} - \hat{p}_{DA}) N_{D\to S}. \end{cases} \tag{1}$$

We then have

$$(1) \iff \begin{cases} \hat{p}_{DA}(N_{D\to D} + N_{D\to A}) = (1 - \hat{p}_{DS})N_{D\to A} \\ \hat{p}_{DS}(N_{D\to D} + N_{D\to S}) = (1 - \hat{p}_{DA})N_{D\to S} \end{cases}$$

$$\iff \begin{cases} \hat{p}_{DA} = (1 - \hat{p}_{DS})\dfrac{N_{D\to A}}{N_{D\to D} + N_{D\to A}} \\ \hat{p}_{DS} = (1 - \hat{p}_{DA})\dfrac{N_{D\to S}}{N_{D\to D} + N_{D\to S}} \end{cases}$$

and by introducing the first equation in the second one, we get

$$\begin{aligned} \hat{p}_{DS} &= \left(1 - (1 - \hat{p}_{DS})\frac{N_{D\to A}}{N_{D\to D} + N_{D\to A}}\right)\frac{N_{D\to S}}{N_{D\to D} + N_{D\to S}} \\ &= \left(1 - \frac{N_{D\to A}}{N_{D\to D} + N_{D\to A}}\right)\frac{N_{D\to S}}{N_{D\to D} + N_{D\to S}} + \hat{p}_{DS}\frac{N_{D\to A}}{N_{D\to D} + N_{D\to A}}\frac{N_{D\to S}}{N_{D\to D} + N_{D\to S}}. \end{aligned}$$

That implies that

$$\left(1 - \frac{N_{D\to A}}{N_{D\to D} + N_{D\to A}}\frac{N_{D\to S}}{N_{D\to D} + N_{D\to S}}\right)\hat{p}_{DS} = \frac{N_{D\to D}}{N_{D\to D} + N_{D\to A}}\frac{N_{D\to S}}{N_{D\to D} + N_{D\to S}}$$

and then

$$\Big(\underbrace{\left[N_{D\to D} + N_{D\to A}\right]\left[N_{D\to D} + N_{D\to S}\right] - N_{D\to A}N_{D\to S}}_{N_{D\to D}\left(N_{D\to S}+N_{D\to A}+N_{D\to D}\right)}\Big)\hat{p}_{DS} = N_{D\to D}N_{D\to S}.$$

We conclude that

$$\hat{p}_{DS} = \frac{N_{D\to S}}{N_{D\to S} + N_{D\to A} + N_{D\to D}}.$$

If we write the log-likelihood for the second model, the part which depends on $\hat{p}_{DS}$ is exactly the same. Then we would make the same computation and get the same MLE in the second model.

(d) See the R code.

(e) If we never see the transitions $A \to S$ and $S \to A$ then the model 1 is good, and even if we were thinking about the model 2, the transitions are estimated by 0 because $N_{A\to S} = N_{S\to A} = 0$.

If we see the transitions $A \to S$ or $S \to A$, we know for sure that model 1 is wrong.

2. (a) Let $d = 1$, then $\lambda(t) = a_1^1 1_{0\le t<T_{max}}$. The log-likelihood of a point process is

$$\ell = \int_0^{T_{max}} \log \lambda(t) dN_t - \int_0^{T_{max}} \lambda(t)dt,$$

that is in our case

$$\ell = log(a_1^1)N_{[0,T_{max}]} - a_1^1 T_{max}.$$

If we derive it with respect to $a_1^1$, we get

$$\frac{\partial \ell}{\partial a_1^1} = \frac{1}{a_1^1}N_{[0,T_{max}]} - T_{max}$$

and then the maximum likelihood estimator is

$$\hat{a}_1^1 = \frac{N_{[0,T_{max}]}}{T_{max}}.$$

(b) More generally, $\lambda(t) = \sum_{i=1}^{d} a_i^d 1_{b_{i-1}^d \leq t < b_i^d}$ and the log-likelihood is given by

$$\begin{aligned}\ell &= \int_0^{Tmax} \log \lambda(t) dN_t - \int_0^{Tmax} \lambda(t) dt \\ &= \sum_{i=1}^{d} \left\{ \log(a_i^d) N_{\left[b_{i-1}^d, b_i^d\right)} - a_i^d \underbrace{\left(b_i^d - b_{i-1}^d\right)}_{=T_{max}/d} \right\}\end{aligned}$$

We derive with respect to $a_i^d$ and the same computation gives the maximum likelihood estimator

$$\hat{a}_i^d = \frac{N_{\left[b_{i-1}^d, b_i^d\right)}}{T_{max}/d}.$$

(c) <u>To simulate by thinning</u> :

Take in entry $T_{max}$ and $n$ and type 1 (for exp) and 2 (for sin).
The bound $M$ is $n$ for type 1 and $2n$ for type 2.
We make a step $\mathcal{E}(M)$ and mark $\mathcal{U}([0, M])$ and :

- if the mark is lower than the intensity, then the point is accepted.
- otherwise, the point is rejected.

<u>To compute the MLE</u> :

Take in entry the observed process, $d$ and $T_{max}$.
Compute $\hat{a}_i^d$ for $i = 1, ..., d$, then return the vector $(a_1^d, ..., a_d^d)$ and the likelihood.

(d) See the R code.

(e) See the R code.
**N.B :** Neuron 1 respond to the odor puff. It is not that clear for neuron 2 which seems more oscillatory.

3. (a) The easiest way to do this is by thinning. We want to simulate the process on $[0, T_{max}]$.

- At time 0, the intensity is $\nu$. So one make one step $\mathcal{E}(\nu)$ which is always accepted. This gives $myT$, that we put in the set of points.
- As long as $myT < T_{max}$,

  - Compute an upperbound $M$ of the intensity. Here the bound is

  $$M = \nu + N_{[myT-1, myT]} \tag{2}$$

  since $h$ has support $[0, 1]$.
  - $myT$ become the old $myT$ added to $\mathcal{E}(M)$. Then we pick a mark $myU \sim \mathcal{U}([0, M])$, and $myT$ is accepted if and only if

  $$myU < \nu + \sum_{myT-1 \leq T < myT} h(myT - T). \tag{3}$$

At the end of the above While loop, it may appen that the last accepted $myT$ is larger that $T_{max}$. In that case, we need to remove it before returning the set of accepted points as the result of the function.
**N.B :** The main difference with Tutorial 4 is that in Tuto 4 we were counting only the parents in (2) and (3). Now we are counting everybody accepted.

(b)

$$
\begin{aligned}
\Lambda(t) &= \int_0^t \lambda(s)ds \\
&= \int_0^t \Big(\nu + \sum_{\substack{T\in N\\ T<s}} h(s-T)\Big) ds \\
&= \nu t + \int_0^t \sum_{T\in N} h(s-T) 1_{\{s-T>0\}} ds \\
&= \nu t + \sum_{T\in N} \int_0^t \big(1-(s-T)^2\big) 1_{\{0<s-T<1\}} ds.
\end{aligned}
$$

The integral here above being null if $T \geq t$, we deduce that

$$
\begin{aligned}
\Lambda(t) &= \nu t + \sum_{\substack{T\in N\\ T<t}} \int_0^t \big(1-(s-T)^2\big) 1_{\{0<s-T<1\}} ds \\
&= \nu t + \sum_{\substack{T\in N\\ T<t}} \int_T^{\min(T+1,t)} \big(1-(s-T)^2\big) ds \\
&= \nu t + \sum_{\substack{T\in N\\ T\leq t-1}} \int_T^{T+1} \big(1-(s-T)^2\big) ds + \sum_{\substack{T\in N\\ t-1<T<t}} \int_T^{t} \big(1-(s-T)^2\big) ds \\
&= \nu t + \sum_{\substack{T\in N\\ T\leq t-1}} \Big[s - \frac{1}{3}(s-T)^3\Big]_{s=T}^{s=T+1} + \sum_{\substack{T\in N\\ t-1<T<t}} \Big[s - \frac{1}{3}(s-T)^3\Big]_{s=T}^{s=t} \\
&= \nu t + \frac{2}{3} N_{t-1} + \sum_{\substack{T\in N\\ t-1<T<t}} \Big[t - T - \frac{1}{3}(t-T)^3\Big].
\end{aligned}
$$

(c) See the R code. Just for illustration, you will see that the counting process is not far from $\Lambda(t)$. In fact their difference is a martingale. Run it several time, you should see that the expectation of their difference is zero.

(d) See the R code.

(e) See the R code.